NASA Contractor Report 185637

National Aeronautics and Space Administration (NASA)/ American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program—1990

Volume 1

Richard B. Bannerot, Editor
University of Houston — University Park
Houston, Texas

Stanley H. Goldstein, Editor
University Programs Office
Lyndon B. Johnson Space Center
Houston, Texas

Grant NGT 44-005-803
December 1990

NASA

PREFACE

The 1990 Johnson Space Center (JSC) National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program was conducted by the University of Houston—University Park and JSC. The 10-week program was operated under the auspices of the ASEE. The program at JSC, as well as the programs at other NASA Centers, was funded by the Office of University Affairs, NASA Headquarters, Washington, D.C. The objectives of the program, which began nationally in 1964 and at JSC in 1965, are

1. To further the professional knowledge of qualified engineering and science faculty members

2. To stimulate an exchange of ideas between participants and NASA

3. To enrich and refresh the research and teaching activities of participant's institutions

4. To contribute to the research objectives of the NASA Centers

Each faculty fellow spent at least 10 weeks at JSC engaged in a research project commensurate with his/her interests and background and worked in collaboration with a NASA/JSC colleague. This document is a compilation of the final reports on the research projects done by the faculty fellows during the summer of 1990. Volume 1 contains reports 1 through 14, and volume 2 contains reports 15 through 28.

CONTENTS

Volume 1

Volume 2

LIF AND EMISSION STUDIES OF COPPER AND NITROGEN

Final Report

NASA/ASEE Summer Faculty Fellowship Program -- 1990

Johnson Space Center

Prepared BY: Murty A. Akundi, Ph.D.

Academic Rank: Associate Professor

University & Department: Xavier University
Physics/Engineering Department
New Orleans, Louisiana 71245

NASA/JSC

Directorate: Engineering

Division: Structures and Mechanics Division

Branch: Thermal

JSC Colleague: James Milhoan

Date Submitted: August 15, 1990

Contract Number: NGT 44-005-803

ABSTRACT

A technique is developed to determine the rotational temperature of Nitrogen Molecular Ion (N_2^+) from the emission spectra of B-X transition, when P and R branches are not resolved. Its validity is tested on simulated spectra of the 0-1 band of N_2^+ produced under low resolution. The method is applied to experimental spectra of N_2^+ taken in the shock layer of a blunt body at distances of 1.91 cm, 2.54 cm, and 3.18 cm from the body.

The laser induced fluorescence (LIF) spectra of copper atoms is analyzed to obtain the free stream velocities and temperatures. The only broadening mechanism considered is Doppler broadening. The temperatures are obtained by manual curve fitting; and the results are compared with least square fits. The agreement on the average is within 10%.

INTRODUCTION

There is a strong need for non-intrusive diagnostic techniques such as passive radiation techniques which utilize radiation emitted by the plasma or active radiation techniques such as Laser Induced fluorescence (LIF) measurements. These techniques are essential to the arc jet flow studies in order to fully understand the nonequilibrium flow conditions spacecraft encounter during reentry. The arc jet facilities simulate the spacecraft reentry conditions and are used for testing thermal protection materials. The identification, characterization, velocity and temperature determination of different atomic, molecular and ionic species makes an important contribution to the arc jet flow studies. The major species of interest in the spacecraft reentry flow environment are N_2, O_2, NO, N_2^+, N, O, O_2^+, NO^+, N^+, O^+ and electrons. Concentration of these species change when the enthalpy of the flow changes. This report is divided into two parts, one on the analyses of the emission spectra and the other on the analysis of the LIF data.

Emission studies: During the last couple of years the arc jet flow diagnostic program at NASA, Johnson Space Center reported emission studies (Ref.1 and 2) of N2 and N_2^+ for vibrational and rotational temperature determination across the shock layer. This technique involves calculating the spectrum for a number of cases and obtain integrals over the wavelength regions of the spectrum as a function of temperature. Ratios of these integrals were then related to the temperatures used to generate the spectra. Spectral integrals from the measured spectra are then compared with the calculated values to determine the temperature. This technique has some limitations and requires numerous parameters to produce the calculated spectra. This method requires repetitive calculations be performed each time a comparison is made. In this report a simpler technique is presented to find the rotational temperatures of N_2^+ in the shock layer. This can be applied to other molecules as well.

LIF studies: The LIF measurements reported here probe different species with a known excitation mechanism using selective excitation with a narrow band laser. Preliminary LIF studies of copper are used to calculate the velocity and temperature are presented here.

MEASUREMENTS

Emission spectra: Details of the experimental set up to record the emission spectrum of nitrogen in the wavelength region of 340 nm to 480 nm are reported earlier (Ref.3).

LIF studies on copper: Laser fluorescence spectra of copper recorded earlier using the experimental setup shown in Fig.1, are analyzed. The setup essentially consists of a YAG pumped dye laser system with intracavity etalons in the dye as well as in the pump laser reducing the dye fundamental band width to less than 0.08 cm^{-1}. To probe the copper atoms, Dicynomethylene (DCM) dye is used together with a KD*P doubling crystal to produce laser light at 327.5 nm. Details of the laser fluorescence detection and control electronics along with laser pulse energy used for copper are reported earlier (Ref.4). Velocity measurements of the free stream for various mass flow rates and arc currents have been carried out using copper as the tracer material. The excitation wavelength is about 327.5 nm (Figs. 2) and the 578 nm fluorescence is collected at right angles to the flow using cutoff and narrow band interference filters. Two separate measurements were taken for Doppler shift studies one with the laser beam perpendicular and the other at an angle of 60^{o} with respect to the flow direction.

RESULTS AND ANALYSIS

Emission spectra of N_2^+: The rotational structure of the 0-1 band of B-X transition of N_2^+ is shown in Fig. 3. Due to the limited resolution the P and R branches are not resolved. In order to analyze and determine the rotational temperature, we have developed a method described in an earlier report (Ref.3). For completeness of this report the equations are shown in Appendix 1 and 2. In order to test the accuracy of these equations, simulated spectra of N_2^+ using NEQAIR program under low resolution are produced at various temperatures. A sample spectra produced at 8000 K is shown in Fig. 4. We have now used the areas and corrected them for P and R branches using equations (7) and (8) shown in Appendix 2., at various temperatures (T). Graphs are now plotted: ln $[I_{PC}/(k'+k"+1)]$ against $k'(k'+1)$ for P-branch and ln $[I_{RC}/(k'+k"+1)]$ against $k'(k'+1)$ for R-branch. According to equation (4) shown in Appendix 1, as given by Herzberg (Ref.5), the slope of the line gives $B_o hC/KT$ from which the rotational temperature can be determined. The temperatures (T_P and T_R) are now calculated using the slopes of the straight lines obtained from the graphs for both P and R branches. The particular set for which the temperature $T_P = T_R = T$ is considered as the rotational temperature. Fig. 5 is a graph showing the agreement between P and R branches at 8000 K. Table 1 shows the temperatures calculated using the slopes of the lines at three different temperatures.

This method is now applied to the spectra taken on February ' 89 at 1.91 cm, 2.54 cm and 3.18 cm positions in the shock layer of a blunt body . Figs. 6 and 7 shows the graphs plotted for corrected P and R branch intensities taking the areas and the results are tabulated in Table - 2.

The calculated temperatures are slightly less than the temperatures obtained by Blackwell etal. (Ref. 4) using computer simulation techniques. This low temperature seems to be due to the overlap of intensity of some other bands of the same species or due to the emissions of the species present in the shock layer in the region analyzed. If an account can be made for this background intensity rotational temperatures can be more precisely determined even in unresolved bands. Work in this direction will be continued in the future.

LIF studies on copper: A representative spectrum of Doppler shifted copper spectrum is shown in Fig. 8. The measured Doppler profiles are used to determine the velocity and temperature. The velocity is obtained using the relation

$$V = C \Delta W / W_o \qquad (1)$$

where V is the velocity of the free stream, C is the speed of light , W is the measured Doppler shift and W_o is the laser excitation frequency. A graph of the measured velocity against bulk enthalpy is shown in Fig. 9 and it demonstrates the expected increase in velocity with increasing enthalpy.

To determine the temperature, a linear curve fit of the quadruplet of Gaussian profiles was fit to the measured profiles. The laser line width was accounted for by folding its profile with the Doppler profile. Both were taken to be Gaussian.
The intensity of Gaussian line shape is given by

$$I(W) = I_o \exp[-4 \ln 2 \{ W_o - W / \Delta W_D \}^2] \qquad (2)$$

and the Doppler Width ΔW_D is given by

$$\Delta W_D = W_o / C \ (8 RT \ln 2 / M)^{1/2}$$

where

I_o = the peak intensity of the Doppler profile
W_o = the frequency at peak intensity
R = gas constant
T = Temperature in Kelvin
M = Mass of copper atom.

Fitted curves at different temperatures are developed using equation (2) and were overlapped on the experimental curve. The one that best fits to the experimental curve is taken to be the temperature at that flow rate. A sample curve fit using one of the two peaks is shown in Fig. 10. The first peak is saturated and so only one peak is considered for an accurate determination of the temperature. Temperatures are determined using this method at various flow rates and the results are presented in Table - 3. The results do not

compare particularly well with the least square fit, indicating the assumptions used in the simple calculations are not adequate. More systematic study has to be carried out and using least square fit Gaussian profiles may yield a more precise temperature determination. It may be noted that a Voigt profile (a combination of Gaussian and Lorentzian profile) may be the appropriate choice for future studies.

CONCLUSIONS

On the basis of the analysis made in both the projects the following conclusions are made:

1. Rotational temperature measurements can be made within 5% even in the case of unresolved bands due to a single species if there is no overlap of other bands.

2. This method can be extended to the observed experimental spectra, where other species are present provided the contribution due to the background is accounted for.

3. The method used to find the temperature of Doppler shifted curves in copper agrees on the average with in 10% with that obtained using least square fit. The velocity can be calculated to an accuracy of 10%.

ACKNOWLEDGMENTS

I wish to take this opportunity to thank NASA/ASEE summer faculty fellowship program for giving me an opportunity to work at NASA, Johnson Space center. My sincere thanks to Dr. Carl Scott and Dr. Sivaram Arepalli for their constant help and guidance throughout the project period. My special thanks to Mr. Eric Yuen for providing the manual curve fit program. I wish to express my thanks to Mr. James Milhoan, my NASA colleague for providing the facilities and for continued support. Last but not the least, I wish to thank all NASA and the Lockheed personnel of the arc jet facility who made my stay here very pleasant and rewarding.

REFERENCES

1. Blackwell, H.E., Wierum, F.A., Arepalli, S.and Scott, C.D., " vibrational measurements of N_2 and N_2^+ shock layer radiation." 27th Aerospace Sciences Meeting, AIAA-89-0248, (1989).

2. Blackwell, H.E., Yuen, E., Arepalli, S., and Scott, C.D., " Nonequilibrium shock layer temperature profiles from arc jet radiation measurements." 24th Thermophysics conference, AIAA-89-1679, (1989).

3. Akundi, M.A., " Temperature determination of shock layer using spectroscopic techniques." NASA summer faculty report, NGT 44-001-800, PP 1- 17, (1989)

4. Arepalli, S., Yuen, E.H., and Scott, C.D., " Application of Laser induced fluorescence for flow diagnostics in arc jets." 5th joint Thermophysics and Heat transfer conference, AIAA-90-1763,(1990)

5. Herzberg, G., " Molecular Spectra and Molecular structure, Vol. 1 Spectra of Diatomic Molecules", D. Van Nostrand Co., Inc., New York, N.Y. (1950).

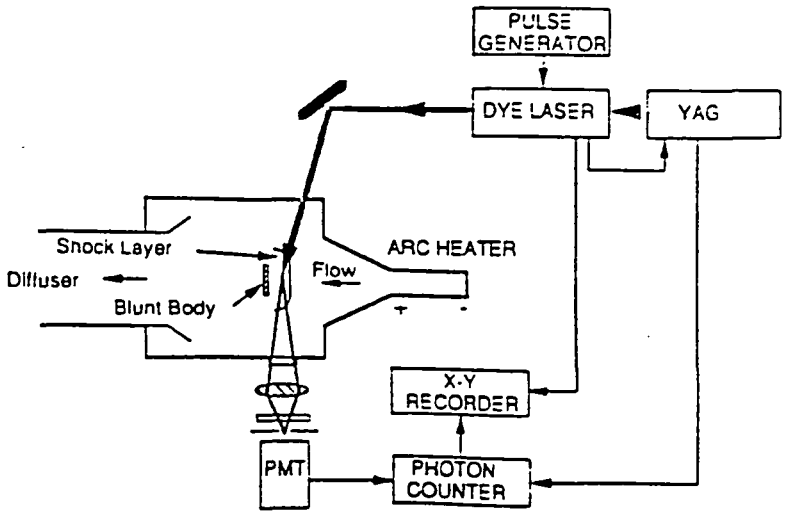


Fig. 1 Experimental setup

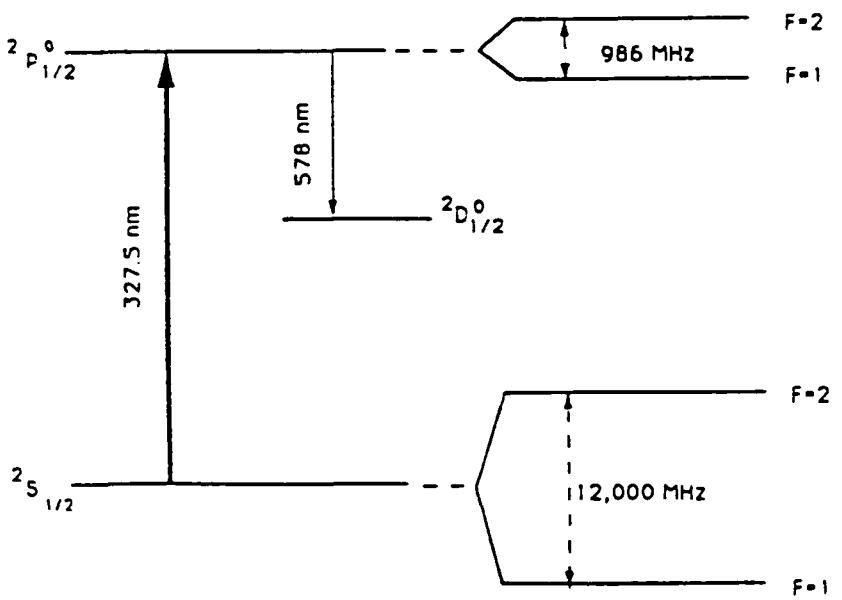


Fig. 2. Detailed energy level diagram of copper atom

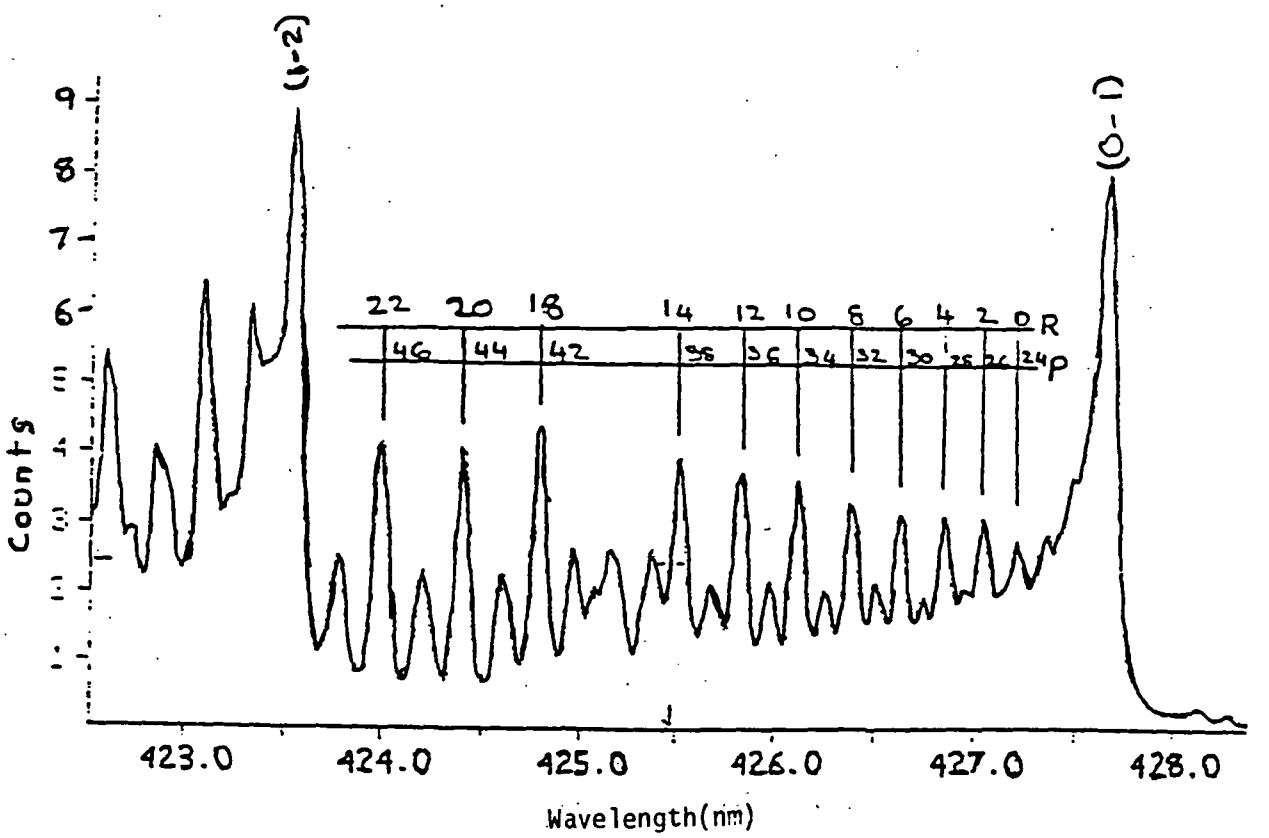


Figure 3.- 0-1 band of N_2^+ B-X system -low resolution

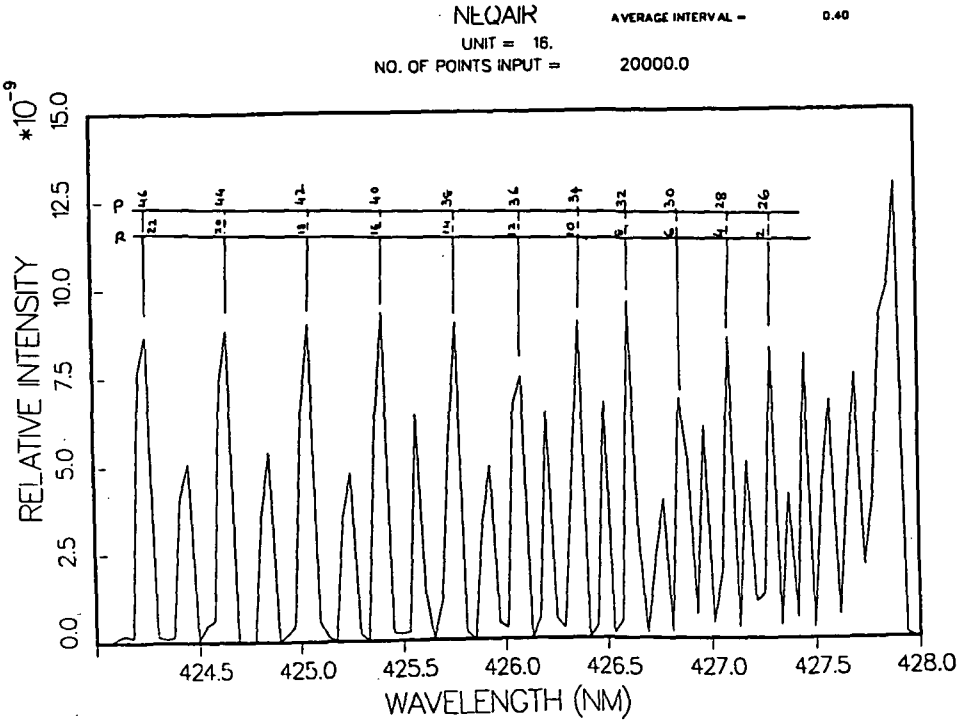


Figure 4.-Simulated low resolution spectrum of 0-1 band of N_2^+ simulated spectra,8000K

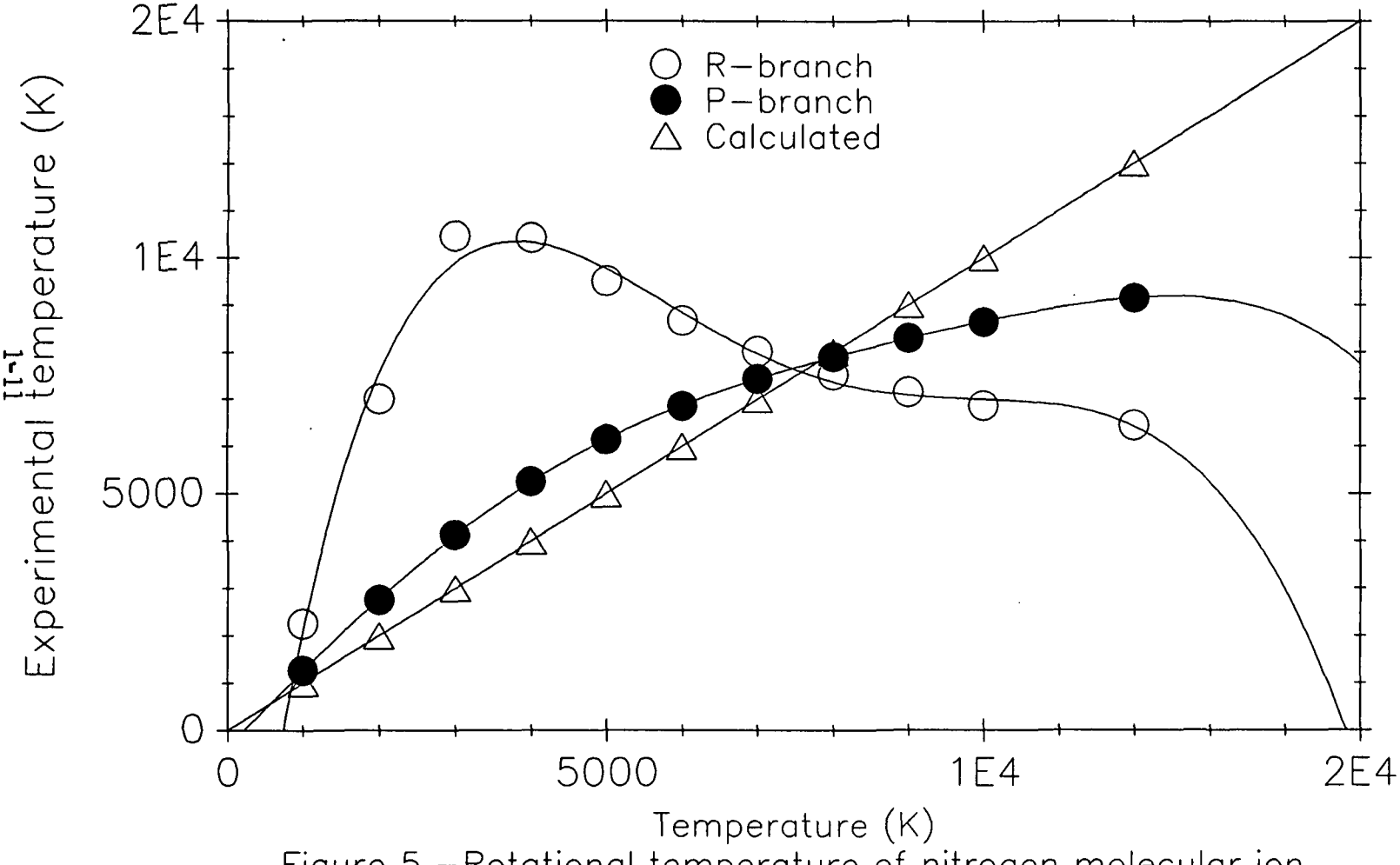


Figure 5.—Rotational temperature of nitrogen molecular ion —simulated spectra, (8000K)

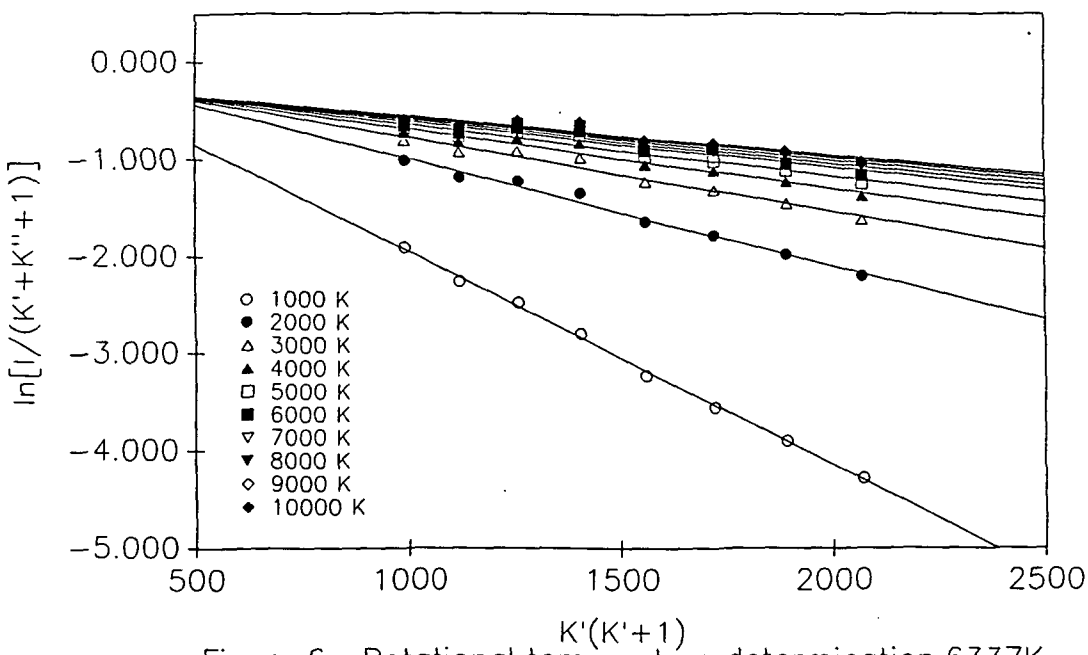


Figure 6.—Rotational temperature determination,6337K

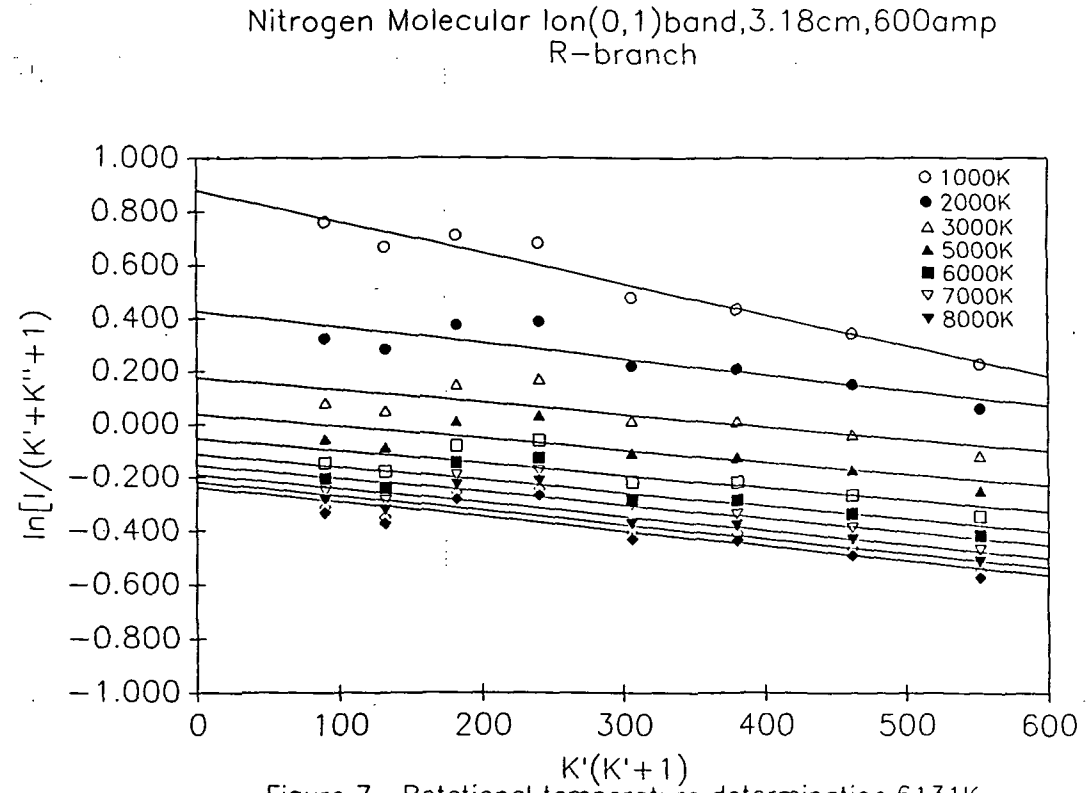


Figure 7.—Rotational temperature determination,6131K

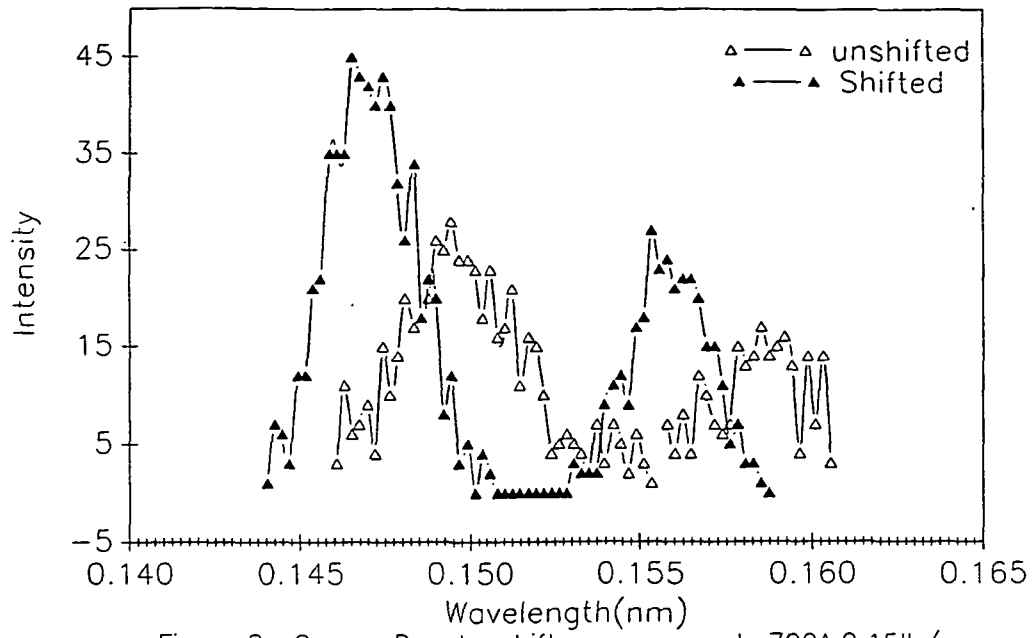


Figure 8.−Copper Doppler shift measurements,700A,0.15lb/s

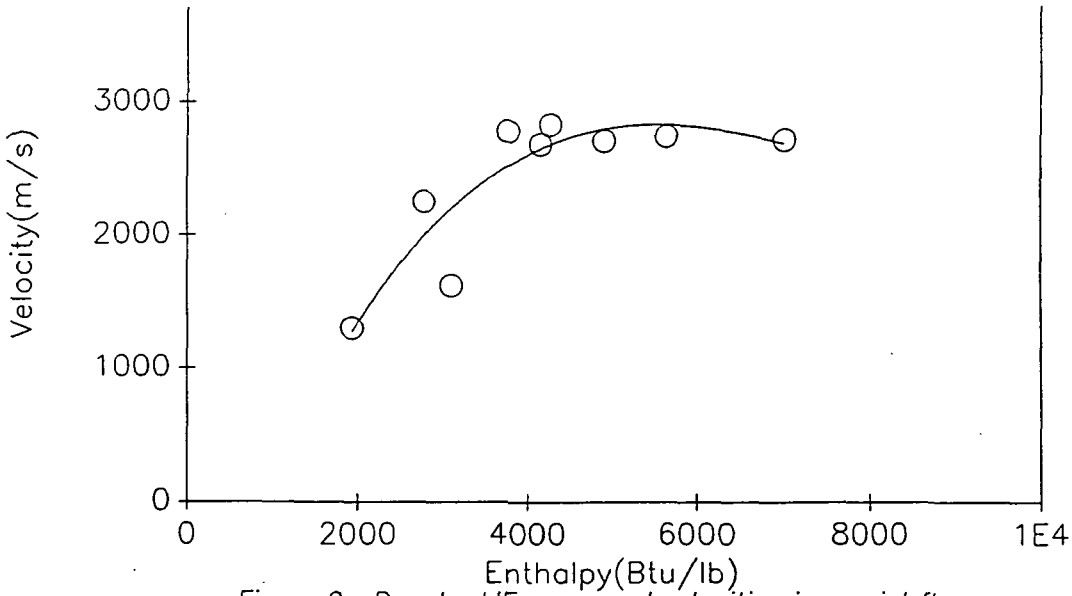


Figure 9.−Doppler LIF measured velocities in arc jet flows

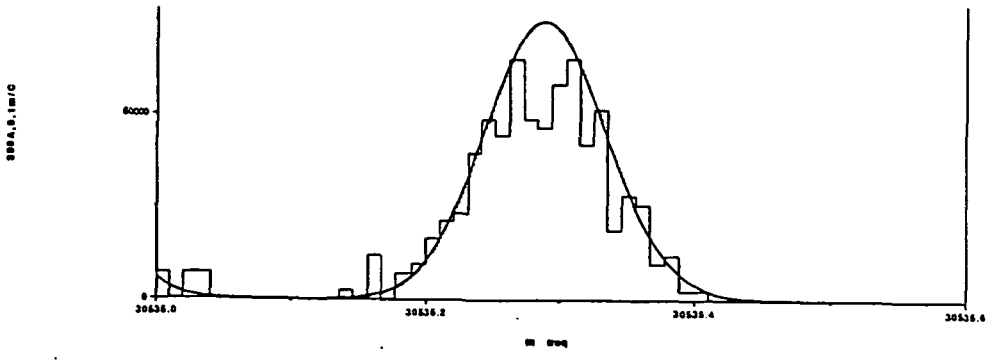


Figure 10.-Copper LIF spectrum and curve fit,300A,0.011b/s,1100K

TABLE 1.- SIMULATION SPECTRA - COMPARISON

T (K)	T_P (K)	% Error	T_R (K)	% Error
8000	7906	1.1	7553	5.8
6000	6234	3.9	6388	6.5
3000	3125	4.2	3217	7.2

TABLE 2.- ROTATIONAL TEMPERATURES OF SHOCK LAYER

Position of Shock layer (cm)	T_P (K)	% Error	T_R (K)	% Error
1.91	3991	0.4	3982	0.2
2.54	2954	1.5	3045	1.5
3.18	6337	5.0	6131	2.2

TABLE 3.- COPPER DOPPLER SHIFT DATA

Current (Amps)	Flow Rate (lb/s)	Doppler Shift 10^{-3}nm	Velocity (m/s)	Temp (K) Curve fit	Temp (K) least Square fit	Enthalpy (Btu/lb)
700	0.05	1.486	2721	--	--	7000
	0.10	1.500	2747	--	2022	5640
	0.15	1.480	2711	1200	1148	4900
	0.20	1.468	2688	1500	1486	4150
	0.25	1.519	2782	--	1076	3760
500	0.10	1.547	2834	1400	1457	4260
	0.20	0.885	1621	1300	1478	3100
300	0.10	1.232	2256	1100	1117	2780
	0.20	0.709	1299	600	498	1925

ROTATIONAL TEMPERATURE EQUATIONS

$$I(P - Branch) = A - 2Je^{-\frac{B_0 J(J-1)hc}{kT}} \quad (1)$$

$$I(R - branch) = A - 2(J+1)e^{-\frac{B_0 (J+1)(J+2)hc}{kT}} \quad (2)$$

Rearranging equations (1) and (2)

$$\ln\left[\frac{I_p}{2J}\right] = \ln A - \frac{B'_0 J(J-1)hc}{kT}$$

$$\ln\left[\frac{I_p}{2J}\right] = A' - \frac{B'_0 J(J-1)hc}{kT} \quad (3)$$

Similarly for R Branch

$$\ln\left[\frac{I_R}{2(J+1)}\right] = A' - \frac{B'_o (J+1)(J+2)hc}{kT} \quad (4)$$

Equation (3) and (4) can general be written for either case as

$$\ln\left[\frac{I}{k' + k'' + 1}\right] = A - \frac{B' K'(k'+1)hc}{kT}$$

where k' - upper rotational state quantum number
k" - lower rotational state quantum number

CORRECTED INTENSITIES

$$I_{EXP} = I_P + I_R \tag{5}$$

$$\frac{I_{EXP}}{I_R} = \frac{I_P}{I_R} + 1 \tag{6}$$

Let $\frac{I_P}{I_R} = B$ where I_p and I_R are the intensities calculated for any temperature using equations (1) and (2)

$$I_{Rc} = I_{EXP}\left[\frac{1}{1+B}\right] \tag{7}$$

$$I_{Pc} = I_{EXP}\left[\frac{1}{1+\frac{1}{B}}\right] \tag{8}$$

I_{Rc} and I_{Pc} are the intensity contributions of P and R branches in the observed spectra

COMPOUND ESTIMATION PROCEDURES IN RELIABILITY

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1990

Johnson Space Center

Prepared By:	Ron Barnes , Ph.D.
Academic Rank:	Associate Professor
University & Department:	University of Houston-Downtown Department of Applied Mathematical Sciences One Main Street Houston, Texas 77002
NASA/JSC	
Directorate:	Safety, Reliability and Quality Assurance
Division:	Reliability and Maintainability
JSC Colleague:	Richard Heydorn, Ph.D.
Date Submitted:	August 16, 1990
Contract Number:	NGT-44-005-803

ABSTRACT

At NASA, components and subsystems of components in the Shuttle and Space Station generally go through a number of redesign stages. While data on failures for various design stages are sometimes available, the classical procedures for evaluating reliability only utilize the failure data on the present design stage of the component or subsystem. Often, few or no failures have been recorded on the present design stage.

Last Summer, in the NASA Faculty Fellow Program , Bayesian estimators for the reliability of a single component, conditioned on the failure data for the present design, were developed. These new estimators permit NASA to evaluate the reliability, even when few or indeed no failures have been recorded. Point estimates for the later evaluation were not possible with the "classical" procedures.

Since different design stages of a component (or subsystem) generally have a good deal in common, the development of new statistical procedures for evaluating the reliability, which consider the entire failure record for all design stages, has great intuitive appeal.

A typical subsystem consists of a number of different components and each component has evolved through a number of redesign stages.

The investigations this Summer considered compound estimation procedures and related models. Such models permit the statistical consideration of all design stages of each component and thus incorporate all the available failure data to obtain estimates for the reliability of the present version of the component (or subsystem).

A number of models were considered to estimate the reliability of a component conditioned on its **total** failure history from **two** design stages.

It was determined that reliability estimators for the present design stage , conditioned on the complete failure history for two design stages have **lower risk** than the corresponding estimators conditioned only on the most recent design failure data.

Several models were explored and preliminary models involving the bivariate Poisson distribution and the Consael Process (a bivariate Poisson process) have been developed. Possible shortcomings of the models are noted. An example is given to illustrate the procedures.

These investigations are ongoing with the aim of developing estimators that will extend to components (and subsystems) with three or more design stages.

INTRODUCTION

Components in the NASA Shuttle and indeed in many other complex systems often go through a number of redesign stages. Classical reliability estimators rely only on the failure data for the present design. Since previous design stages often have a good deal in common with the present design, statistical procedures for estimating reliability may be improved by also taking into account the failure data on the earlier design versions as well.

Last year in the NASA Summer Faculty Fellow program (SFF), Bayesian estimators for the expected value of the reliability of a single component, conditioned on its failure history (for the present design stage) were developed for the cases of (1) a constant failure rate - the exponential model and (2) a variable failure rate - Weibull model [1],[2].

For the constant failure rate model it was shown that:

$$E[\, R(t) \mid N(T) = n \,] = 1/(1 + t/T)^{n+1} \qquad (1)$$

where

R(t) = the reliability or probability that the component will successfully function up to time t in the future,

N(T) = the number of failures up to time T (in the past failure history), and

n = the number of failures recorded up to time T for this component.

These new estimators enable NASA to evaluate reliabilities when few or even no failures have been recorded. Evaluation in the latter case was not possible with the previous "classical " estimators.

In March 1990, these Bayesian estimators were employed along with the classical estimators in a NASA report on the reliability of Orbiter APU hydraulic hoses [3]. Both sets of estimators predicted a very high reliability for success of the hoses on the next mission (.999..).

This report will focus on investigations to extend the constant failure rate model to utilize the total failure history on a component with 2 or more design stages. The investigations considered compound estimation procedures in order to utilize this failure data history. The report incorporates the Bayesian estimators developed last year and noted earlier in this report.

Motivation for the Study

Since redesign versions of a component would appear to have some commonality, the idea of reliability estimators which incorporate the total failure history of all design stages of a component seems worth considering. When data occurs, for example, on a number of identical valves which have been through 2 design stages with the number of failures on the j-th valve with the first design denoted by N_{1j} and the number on the i-th valve with the new design denoted by N_{2i}, one can plot N_1 versus N_2 The data often suggest some correlation between the two failure counts. In general, the number of failures N_1 and N_2 are not independent. In the discussion to follow, $N_1(T_1)$ and $N_2(T_2)$ the number of failures up to times T_1 and T_2 recorded for the old and new designs respectively are each assumed to have a Poisson distribution with possibly different failure rates λ_1 and λ_2 which are unknown. Since $N_1(T_1)$ and $N_2(T_2)$ are not assumed to be independent, the problem of obtaining Bayesian reliability estimators conditioned on the failure data N_1 and N_2 requires some form of a joint or compound probability distribution for N_1 and N_2 .

The impact of such estimators that utilize the total failure history from 2 or more design stages of a component is indicated by the following result.

Proposition

$$E[R(t)-E(R(t)|N_1(T_1),\ N_2(T_2))]^2 \leqq E[R(t)-E(R(t)|\ N_2(T_2))]^2$$

i.e., the estimator conditioned on two stages of failure history has a **lower risk** than the corresponding risk for the estimator conditioned only on the most recent design failure history. Hence as additional failure data on earlier design stages are included, the corresponding risk decreases.

Proof

$$E[R(t)-E(R(t)|N_1(T_1),\ N_2(T_2))]^2 =$$

$$E[R(t)-E(R(t)|N_2(T_2))-\ (E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2)))]^2 =$$

$$E[R(t)-E(R(t)|N_2(T_2))]^2-\ 2E[(R(t)-E(R(t)|N_2(T_2))) *$$
$$(E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2)))] + \qquad (2)$$
$$E[E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2))]^2 =$$

But since

$$E[(R(t)-E(R(t)|N_2(T_2)))(E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2)))] =$$
$$E\{E[(R(t)-E(R(t)|N_2(T_2)))(E(R(t)|N_1(T_1),N_2(T_2)) -$$
$$E(R(t)|N_2(T_2)))|N_1(T_1),N_2(T_2)]\ \} =$$
$$E[E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2))]^2$$

then expression (2) becomes

$$E[R(t)-E(R(t)|N_1(T_1),\ N_2(T_2))]^2 =$$
$$E[R(t)-E(R(t)|N_2(T_2))]^2 - E[E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2))]^2$$

Since the last term is non-negative the result is established. Note that this result implies that, if the last term is positive, a strict inequality holds.

DISCUSSION

The present investigations have considered two preliminary models for the distribution of N_1 and N_2 which lead to reliability estimators.

Model A : Bivariate Poisson Model

The first of these employs the bivariate Poisson distribution.. The joint probability function for a 2 dimensional Poisson process $(N_1(t),N_2(t))$ with cov($N_1(t),N_2(t)$) = υt is given in [4] by:

$$P[\, N_1(t) = n_1,\, N_2(t) = n_2] = f(n_1,\, n_2,\, t) \qquad (3)$$

where

$$f(n_1,\, n_2,\, t) = \exp(-\lambda_1 t - \lambda_2 t + \upsilon t) \sum_{j=0}^{\min(n_1,\, n_2)} ((\lambda_1 - \upsilon)^{n_1 - j} (\lambda_2 - \upsilon)^{n_2 - j} \cdot$$

$$t^{n_1 + n_2 - 2j} (\upsilon t)^j) \,/\, [(n_1 - j)!\, (n_2 - j)!\, j!] \qquad (4).$$

Note that in the discussion to follow, in place of t in equation (3) one could use $T = T_1 + T_2$ which is the total elapsed time for failures for both designs. Utilizing expression (4) and Baye's formula,where $X_2(t) = 1$ denotes that the component in its second design stage is still operating up to time t, one can show that

$$E[R(t) | N_1(T)=n_1, N_2(T)=n_2] = P[X_2(t)=1 \mid N_1(T)=n_1,\, N_2(T))=n_2] =$$

$$P[N_1(T)=n_1, N_2(T)=n_2 \mid X_2(t)=1]\; P[X_2(t)=1] / P[\, N_1(T) = n_1,\, N_2(T) = n_2] \qquad (5)$$

Assuming a joint Bayesian prior distribution for λ_1,λ_2, this expression becomes:

$$\frac{\int\int P[N_1(T)=n_1,N_2(T)=n_2 \mid X2(t)=1,\lambda_1,\lambda_2]\ P[X2(t)=1\mid\lambda_1,\lambda_2]\ g(\lambda_1,\lambda_2)\ d\lambda_1,d\lambda_2}{\int\int P[N_1(T)=n_1,N_2(T)=n_2 \mid \lambda_1,\lambda_2]\ g(\lambda_1,\lambda_2)\ d\lambda_1,d\lambda_2} \qquad (6)$$

If one assumes the joint prior distribution $g(\lambda_1,\lambda_2)$ to be uniform on finite rectangles with λ_1,λ_2 independent, i.e.

$$g(\lambda_1,\lambda_2)= \begin{cases} (1/\theta_1)\ (1/\theta_2) & \text{for } 0<\lambda_1<\theta_1 \text{ and } 0<\lambda_2<\theta_2 \\ 0 & \text{elsewhere} \end{cases} \quad \text{then,}$$

one can show that the uniform priors, with the limits extended to the entire first quadrant, give

$$E[R(t)|N_1(T)=n_1,N_2(T)=n_2] =$$

$$\frac{e^{-\upsilon t}\sum_{j=0}^{\min(n_1,\, n_2)} [1/(1+ t/T)^{n2-j+1}]\,(\upsilon t)^j/j!}{\sum_{j=0}^{\min(n_1,\, n_2)} (\upsilon t)^j/j!} \qquad (7)$$

Note that in the case when $n_1=0$ failures one has

$$E[R(t)|N_1(T)=0,N_2(T)=n_2] = e^{-\upsilon t}(1/(1+ t/T)^{n2+1}) \qquad (8)$$

where T is the total elapsed time from the start of the first design. Notice the similarity of expression (8) to formula (1) developed previously for the reliability estimator given just the failure data for the second design.

Also observe that if N_1 and N_2 are completely independent then

$$E[R(t)|N_1(T)=n_1,N_2(T)=n_2] = E[R(t)|N_2(T)=n_2] = 1/(1+ t/T)^{n2+1}.$$

Thus this model would, in some sense, appear to generalize the earlier model(1) which considered only one design stage.

Example

As an application of this reliability estimator consider the following example.

Arbous and Kerrich [5] recorded the number of accidents of 122 individuals (shunters) in two consecutive time periods. For each individual, the number of accidents in the first 6 year period was recorded and then after new insurance and safety procedures were implemented, the number of accidents for the same individual was recorded in the next 5 years. The authors estimated $\upsilon T \approx .257$ for the $T = 11$ years.

One can use expression (7) to evaluate the "reliability" for the next year of an individual,selected at random, given his/her total past accident (failure) history. Note that in this case the reliability is just the probability that a randomly selected individual with the given accident history will not have an accident in the next t years.

For a randomly selected individual with accident history given by

$$n_1=3 \ , \ n_2 = 1 \ , \ t=1 \ , \ \upsilon T \approx .257 \ , \ T_1 = 6, \ T_2 = 5, T = T_1 + T_2=11$$

one finds by using expression (7) that

$$E[R(t)|N_1(11)=3 \ ,N_2(11)=1 \ ] = .6067624$$

Thus the reliability estimator obtained from this model suggests that for a randomly selected individual who experienced 3 accidents in the first 6 years and 1 in the next 5 years, the probability of no accidents in the next year is approximately 61 percent. Such information may be used in the setting of insurance premiums for the next year for various classes of individuals based on their past accident histories.

Note that in terms of NASA component reliability:

1. The individuals correspond to different copies of a single component with two design stages.

2. Each copy is located in a somewhat similar environment(e.g. the copies may be located on each of the three space shuttles).

3. Data on failures of each copy were recorded in the first design stage (i.e. the first 6 years of accident data).

4. A second improved design replaced the first and failures were recorded.

Then $E[R(t)|N_1(T_1)=n_1,N_2(T_2)=n_2]$ gives the probability that the second design stage with a given failure history in this location will not fail in the next t time units. Note that the NASA description assumes that when the component fails it is replaced or repaired so that it is equivalent to the original system before the failure.

It should be pointed out that in Model A some assumptions about the relationships between λ_1, λ_2, and υ were made. In particular, $\upsilon < \lambda_1, \lambda_2$ suggests that υ is related to the priors. With this in mind, a second preliminary model has been developed.

Model B : Consael Model

As was noted earlier, N_1 and N_2 are probably not independent in general. The Consael process [6] defines a bivariate compound Poisson process by

$$P[N_1(T)=n_1,N_2(T)=n_2] =$$

$$\int\int e^{-\lambda_1 T}\,[(\lambda_1 T)^{n1}]/n1!\;\; e^{-\lambda_2 T}[(\lambda_2 T)^{n2}]/n2! * g(\lambda_1,\lambda_2)d\lambda_1 d\lambda_2 \quad (9)$$

In the Consael process for fixed values of λ_1 and λ_2, N_1 and N_2 correspond to independent Poisson processes while λ_1 and λ_2, have a joint density function $g(\lambda_1,\lambda_2)$.

If one assumes the triangular density function ($\lambda_2 \leq \lambda_1$ since one assumes that the newer design is an improvement) one has:

$$g(\lambda_1,\lambda_2) = \begin{cases} 2/\theta_1^{\,2} & \text{for } 0 < \lambda_1 < \theta_1 \text{ and } 0 < \lambda_2 < \lambda_1 \\ 0 & \text{elsewhere} \end{cases} \qquad (10)$$

Utilizing expressions (9) and (10) ,and Baye's formula in expression (6) (and taking limits on the prior distribution) one can again obtain an estimate for the reliability:

$$E[R(t)|N_1(T)=n_1,N_2(T)=n_2]$$

At the present time investigations are continuing with this estimator which has a rather complex , highly combinatorial closed form. It is anticipated that further consideration of this model will indicate approaches to the development of the general model which can incorporate failure data from any number of redesign stages .

SUMMARY AND CONCLUSIONS

Various NASA failure data suggests support for the development of compound/mixture models to estimate the reliability of components that have failure data recorded on more than one design stage. Bayesian estimators that can utilize all "relevant" failure data, even from an earlier design of the component, were investigated.

A Bayesian estimator based on the bivariate Poisson distribution was developed and an example illustrating the technique applied to insurance data was given. The similarity of this example to NASA reliability problems was also noted.

An additional estimator based on the Consael process was developed. Investigations into this model are continuing. It is anticipated that these investigations will lead to a general model to utilize all relevant failure data on a component (or subsystem of components) that has experienced more that two design stages.

REFERENCES

1. Heydorn, Richard, unpublished paper "A Bayesian Approach to Reliability and Confidence," NASA/JSC Reliability and Maintainability Division, Spring 1989.

2. Barnes, Ron ,"A Bayesian Approach to Reliability and Confidence", NASA Summer Faculty Fellow Final Report, August 1989.

3. "A Reliability Assessment of the Orbiter APU Hydraulic Hose", prepared by NASA/JSC Reliability and Maintatinability Division, April 1990.

4. Haight, Frank A., Handbook of the Poisson Distribution, John Wiley and Sons Inc., New York, 1967.

5. Arbous, A. and Kerrich, J. ,"Accident Statistics and the Concept of Accident-Proneness," Biometrics, 7, 340-432, December 1951.

6. Everitt, B.S., and Hand, D.J., Finite Mixture Distributions, Chapman and Hall, London, 1981.

BIBLIOGRAPHY

1. Efron,B. and Morris, C. N., "Stein's Estimation Rule and Its Competitors - An Empirical Bayes Approach," J. Amer. Statist. Assoc. 68, 117-130, 1973.

2. _____, "Stein's Paradox in Statistics," Scientific American 236, (5) 119-127, 1977.

3. Haight, Frank A., "Index to the Distributiions of Mathematical Statistics," Journ. Res. of NBS -B.Math 65B, 1, Jan-Mar. 1961.

4. Heydorn, Richard, and Basu, Rehka, unpublished report "Estimating Parameters in a Finite Mixture of Probability Densities"NASA/JSC 1989.

5. Holgate, P. , "Estimates for the Bivariate Poisson Distribution," Biometrika 51, 1&2 , 741, 1964.

6. Martz, Harry F. and Waller, Ray A. , Bayesian Reliability Analysis, John Wiley and Sons Inc., New York, 1982.

7. Stein, Gillian and Juritz, June, "Bivariate Compound Poisson Distributions," Commun. Statist. - Theory Meth. 16(12), 3591-3607, 1987.

8 Stigler, Stephen , "The 1988 Neyman Memorial Lecture: A Galtonian Perspective on Shrinkage," Statistical Science, Vol.5, No. 1, 147-155, Feb. 1990.

9. Teicher, Henry "On the Multivariate Poisson Distribution," Skand. Aktuar. Tidskar. 37, 1-9.

ACKNOWLEDGEMENTS

The author would like to thank the members of the Reliability and Maintainability Division of Chief Duane Duston, especially Tom Edwards and Malcolm Himel. Special thanks also to Mary Gibson and the staff for their helpful support and assistance.

Finally, thanks to my NASA sponsor, Dr. Richard Heydorn for making my SFF experiences both stimulating and productive.

TECHNIQUES FOR DETERMINING TOTAL BODY WATER USING DEUTERIUM OXIDE

Final Report

NASA/ASEE Summer Faculty Fellowship program--1990

Johnson Space Center

Prepared by:	Phillip A. Bishop, Ed.D.
Academic Rank:	Assistant Professor
University & Department:	University of Alabama Health and Human Performance Studies Tuscaloosa, AL 35487
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Biomedical Operations and Research
JSC Colleague:	Helen Lane, Ph. D.
Date Submitted:	August 27, 1990
Contract Number:	NGT-44-005-803

ABSTRACT

The measurement of total body water (TBW) is fundamental to the study of body fluid changes consequent to microgravity exposure or treatment with microgravity countermeasures. Often, the use of radioactive isotopes is prohibited for safety or other reasons. It was desired that a safe method of total body water measurement be selected and implemented for use by some Johnson Space Center (JSC) laboratories, which permitted serial measurements over a 14 day period which was accurate enough to serve as a criterion method for validating new techniques. These requirements resulted in the selection of deuterium oxide dilution as the method of choice for TBW measurement. This report reviews the development of this technique at JSC. The recommended dosage, body fluid sampling techniques, and deuterium assay options are described.

INTRODUCTION

The measurement of total body water (TBW) is necessary for studying the response of body fluids to microgravity and microgravity countermeasures. The measurement of TBW with deuterium oxide (D2O), has been well studied. This method is safe, non-radioactive, and potentially very accurate, and is the method of measuring body water turnover in doubly-labeled water studies of long-term metabolic rate (Schoeller et al., 1980; Wong et al., 1988; Schoeller, 1990). By using low doses and measuring the baseline D2O level, repeated TBW measurements can be made over several days without compromising subject safety.

DOSAGE

A wide range of D2O dosages have been used in TBW determinations. The chief limiting factor in determining minimal dosage is the precision of determination of D2O diluted in a body fluid sample. Consequently, minimal dosage is a function of the volume of TBW and the precision of the D2O assay. Minimal dosage reported in the literature for adults was 1 g (Schoeller et al., 1980). Such a dosage results in a D2O concentration increase of 29 ppm in a 60 kg female at 20% fat with assumed water fraction of 73% of the fat free mass or 24 ppm in a 70 kg male with 15% fat. The largest D2O dosage reported in the literature was 107 g for a 70 kg body weight male (Nielsen et al., 1971). The natural abundance of D2O in tap water varies, but is approximately 140-150 ppm (Thomson, 1963; Davis et al., 1987). The body tends to concentrate D2O slightly, yielding saliva baseline concentrations of about 0.02-0.03 ppm above local levels of ingested water (Halliday and Miller, 1977). This suggests that a 1 g dose of 99.9% D2O would raise the baseline value only about 16% for males.

Because many NASA applications require repeated TBW measurements over 10-13 days, the baseline concentration of D2O will increase based upon an elimination half-life of about 10 days (Schloerb, et al., 1950; Schoeller et. al., 1980; Schoeller and Webb, 1984; Wong, et al., 1988; Schoeller, 1990). For maximum accuracy, it is necessary to minimize baseline levels and maximize D2O concentration differences between baseline and post dose. This is accomplished by keeping each dose as small as can be accurately

measured after dilution in body fluids. For the purposes of a 13 day bedrest with requirements to determine TBW before and after a lower body negative pressure and fluid-loading countermeasure, the following doses were administered on the days shown:

		Estimated D2O Conc. (ppm)*			
		Pre	Dose	Post	% Increase
Day	Dose (grams)				
-1	4	150	114	264	76
5	4	218	114	332	52
6	6	317	171	488	54
10	6	349	171	520	49
11	7	438	228	638	46

* Estimated concentrations for 60 kg subject with 20% fat and 73% of fat free mass as water.

Total doses=5; total dosage= 27 g.

In cooperation with Dr. Suzanne Fortney of the JSC Cardiovascular Research Laboratory, this scheduled dosage was administered to one bed rest subject and six controls. This work is currently in progress.

Risks

Toxicity of D2O to humans has not been precisely determined, but it has been used in human research for over 30 years without reported ill effects. The TD_{50} required for reproductive effects in animals is 840 g/kg weight (Registry of Toxic Effects, 1986). Animal studies show that D2O concentration must reach 30-35% of total body water to be lethal. Fusch and Moeller (1988) suggest that D2O concentrations for short term studies be less than 10 g/ kg of body water. Schloerb et al. (1950) found no effects in healty or ill subjects receiving D2O doses of 100g. Doses in this range far exceed that normally required for TBW determinations using sensitive deuterium assay methods.

D2O has been reported to produce nystagmus (Money and Miles, 1974). In recent studies at Johnson Space Center, a dose of approximately 200 g for a 70 kg subject was required to produce symptoms of vestibular impairment in . Therefore, the maximal total exposure of 28 g is extremely safe.

FLUID SAMPLING

One of the advantages of D2O measurement of TBW, is that serum, urine, respiratory water,and saliva have all been used successfully in the technique. For convenience, safety, and requirements for non-invasive methods for space flight, saliva sampling is reviewed as a recommended fluid sampling technique. The D2O method requires an initial fluid sample to determine baseline D2O levels. This can be collected by having the subject expectorate into a small vial. Only approximately 2 ml of saliva is needed, but a capped 5-10 ml vial such as the Sarstedt saliva collection kit used at JSC is convenient to use. It is important that the subject refrain from eating or drinking for 2 hours prior to saliva sampling as the potential exists for obtaining saliva diluted with ingested fluid. Food or drink ingested immediately prior to or following D2O dosing may retard equilibration of D2O with body water. D2O doses should be weighed to the nearest milligram or better in order to maintain overall precision in the ppm range (Schoeller et al., 1980). Because the doses are small, care must be taken to ensure the subject does not lose any of the deuterium. This is accomplished by rinsing the dose vial with at least half its volume of deionized water and having the subject ingest the rinse water. This rinsing is done twice (Thomson, 1963). Equilibration has been found to take about three hours in resting subjects (Schloerb, et al., 1950; Schoeller et al., 1980; Lukaski and Johnson, 1985; Wong et al., 1988). During the equilibration period, no food or drink should be taken. Since any D2O lost must be accounted for in TBW determinations, urine voids during this time must be sampled for D2O concentration and the volume recorded to adjust for this loss. Halliday and Miller (1977) have reported that fractionation results in unequal distributions of D2O in the baseline samples with urine, plasma, serum, and saliva, all showing concentrations higher than local drinking water and respiratory water vapor showing lower baseline concentrations. Fractionation occurs because the heat of vaporization is 3%, and the heat of fusion is 5.5%, higher in D2O compared to H2O (Thomson, 1963). Hence, baseline and equilibration samples should be drawn from the same type of fluid.

DEUTERIUM ASSAY

Methods for deuterium assay include, infrared spectrophotometry, gas chromatography, and radio-isotope ratio mass spectroscopy (Graystone, et al., 1967; Thomson, 1963). Other methods include freezing point depression,(Reaser and Burch, 1958), near-infrared spectrophotometry, refractometry, and falling-drop, (Thomson, 1963); however, these methods are not sufficiently accurate for the dilutions used in our applications and will not be discussed.

Infrared spectrophotometry

Infared spectrophotometry has been used extensively in assaying D2O. It requires careful and laborious sample preparation. Lukaski and Johnson (1985) tested the recovery of plasma and urine using several treatment protocols. Chemical precipitation of proteins with copper sulfate, cadmium chloride, feric chloride, mercuric chloride, and stannous chloride and treatment with activated charcoal all yielded either turbid supernatants, or supernatants could not be recovered. Only vacuum sublimation or distillation yielded high recoveries. Turner et al. (1960) and Fusch and Moeller (1988) used vacuum distillation in which liquid samples were processed at pressures of 0.05 to 1 mmHg with the vapor subsequently refrozen in liquid nitrogen or dry ice and isopropanol. Stansell and Mojica, (1968) used distillation (under atmospheric pressure) with condensate collected in a water-cooled tube.

Davis et al. (1987) used diffusion in Conway dishes to obtain clean samples. The diffusion technique is simpler than most other sample cleansing methods, but it may be inaccurate due to fractionation between D2O and H2O (Wong and Klein, 1986). Also accuracy diminishes because the actual concentrations measured are halved in this method. Volumetric errors in either the sample or the H2O are possible, and to be exact, the saliva water volume (rather than saliva volume) should be matched to the H2O volume. Additionally, the diffusion method would be very laborious for a large number of samples.

For potential use in JSC laboratories, a vacuum distillation system was constructed and tested. Approximately 2 ml of unpurified sample was placed in a stoppered sample tube connected with a small plastic three-way valve to a 7 ml vacutainer with two-inch pieces of tygon tubing. Both the sample tube and the vacutainer were initially heated to approximately 45°C in a water bath. Vacuum was then applied to both tubes via the three-way valve with a Dayton (1/4hp) vacuum pump. The stopcock was then turned to seal off the vacuum side and maintain a vacuum in the two-tube system. The sample tube and vacutainer were then allowed to cool to room temperature before the vacutainer was placed in an acetone-dry ice mixture. Failure to allow the sample to cool before chilling the vacutainer lowered the system pressure too low, resulting in boil-over of the sample which contaminated the vacutainer. After a few minutes, water vapor evolving from the sample raised the pressure sufficiently to allow returning the sample to the water bath. The sample then distilled over several hours until the sample was reduced to dryness. It is important that the sample be dryed because D2O fractionates at a higher temperature than H2O resulting in an artificially low D2O concentration if vaporization is incomplete (Thomson, 1963).

Spectroscopic analysis of the O-D vibrational band at 2500 cm^{-1} (3.98 nm) is conducted with either a single beam fixed-filter (Lukaski and Johnson, 1985; Stansell and Mojica, 1968) or a dual beam (Turner et al., 1960) spectrophotometer in a calcium flouride cell Thermostated at 15 or 20 °C (Lukaski and Johnson, 1985) or 30 and 48 °C (Stansell and Mojica, 1968) for sample and reference cells, respectively. For single beam machines, deionized water is used as the zero reference. Between samples, some investigators cleaned the cell with dry nitrogen, (Lukaski and Johnson, 1985) whereas others did not (Stansell and Mojica, 1968). Davis et al. (1987) were able to measure down to 30 ppm. Reported precision of the single beam method is 2.5% (Lukaski and Johnson, 1985).

Gas chromatography

The gas chromatography (GC) analysis for D2O requires minimal sample preparation. A 50 microliter sample is injected into an evacuated column containing calcium hydride (Arnet and

Duggleby, 1963; Wong and Klein, 1986). Protium and deuterium gas evolves and is injected through a 7 or 8 port valve, into a chromatograph equipped with a 5-20 ml sample loop. The gas is then passed through a 1 meter long activated charcoal column held at room temperature, where the hydrogen is cleaned further. From the column, the gas passes to the detector where the difference in thermal conductivity between the deuterium enriched gas from the sample, and the hydrogen carrier gas is measured with a thermal conductivity cell thermostated at 100°C (Arnet and Duggleby, 1963). The thermal detector cell voltage is output to a recorder. Any deuterium in the carrier gas is zeroed out during set-up. The peak height is measured or better, the area under the curve is integrated to measure concentration relative to known standards.

Mendez et al., (1970) examined the methodology of GC analysis and found it to be as accurate as infrared analysis at a concentration of 1085 ppm (0.12% w/v). They also examined vacuum sublimation and found it did not significantly improve the accuracy or precision of the analysis. In fact, both untreated and vacuum distilled saliva produced exactly the same mean and standard deviation for 11 samples at a concentration of 1356 ppm (0.15%). Duplicate samples showed very high test-retest means, differing by 1 ppm at 1175 ppm (0.13%). Mean recovery in urine samples was 99.3%.

Ratio-isotope mass spectroscopy

Ratio-isotope mass spectroscopy is the most common method of measuring deuterium concentrations in studies which use doubly-labeled water to measure metabolic rate. D2O is assayed because the doubly labelled water technique requires that O18 respiratory turnover be adjusted for water losses of O18. In this method, the fluid sample is treated with a hot reactive metal such as uranium or zinc, thereby liberating gaseous hydrogen. The deuterium and protium evolved are then measured on a mass spectrometer configured for deuterium/protium analysis. This method can measure very low concentrations of deuterium. The chief disadvantages of this technique is the very intricate labware required for sample preparation and the expense of the mass spectrometer. Currently the Planetary Sciences Division of JSC is setting up the procedure to use this method and this would be a desirable technique for D2O determination when completed.

CALCULATIONS

The dilution volume determined by the D2O procedure actually represents hydrogen volume rather than water volume. Evidence from animal studies suggests that hydrogen space over-estimates water space by 2-6 percent (Wong et al., 1988; Schoeller, 1990). Therefore, hydrogen space should be adjusted by about 04% to obtain the most accurate TBW value. This is most readily accomplished by dividing the D2 dilution space values by 1.04 (Schoeller, 1990).

In general, the equation used is:

$$\text{D2 dilution space} = \frac{\text{(volume D2O ingested- volume excreted)}}{\text{final D2O equilibrium concentration}}$$

When isotope ratio mass spectrometry is used, the asssay results are expressed relative to a standard, usually a seawater or precipitation standard such as standard mean ocean water (SMOW) or standard light arctic precipitation (SLAP). The formula utilized for calculating TBW is:

$$\text{D2 dilution space} = (d/MW)*(APE/100)*(18.02/(Rstd*\text{delta D2O}))$$

Where:

d=dose of D2O
MW=molecular weight
APE- atom percent excess of deuterium
Rstd= ratio of deuterium to hydrogen in the standard
Delta= increase between baseline and post-ingestion samples (Dempsey et al., 1984).

The formula provided by Halliday and Miller, (1977) is one of the most comprehensive:

$$\text{D2 dilution space (kg)} = (D_f d * D_e)/[(p-x)*1000]$$

Where:

$$D_e = (A-C(y-x))/D_f$$

Where:

D_f=D2O concentration of administered D2O
x=ppm D2O of baseline fluid sample
p=ppm D2O of equilibrated fluid sample
A=weight of D2O dose administered
C=weight of deuterium passed in urine
y=ppm in excreted urine

Weight-volume doses can be converted to volume/volume by correcting for the density of D2O which is 1.105 at 25°C (Thomson, 1963).

SUMMARY

D2O dilution techniques offer a safe, well-tested method of determining TBW. Saliva, which may be easily and safely collected, serves as a suitable body fluid sample. Numerous options are available for deuterium assay. JSC laboratories have potential access to GC and isotope-ratio mass spectroscopy, but neither method is currently operational.

REFERENCES

Arnet, E. M. and Duggleby, P.M.: A rapid and simple method of deuterium determination. Anal. Chem. 35, 1420-1424, 1963.

Davis, J.M; Lamb, D.R.; Burgess, W.A.; and Bartoli, W.P.: Accumulation of deuterium oxide in body fluids after ingestion of D2O-labeled beverages. J. Appl. Physiol. 63(5):2060-2066, 1987.

Dempsey, D.T.; Crosby, L.O. Lusk, E.; Oberlander, J.L.; Pertschuk, M.J.; and Mullen, J.L.: Total body water and total body potassium in anorexia nervosa. Am. J. Clin. Nutr. 40:260-269, 1984.

Fusch, Ch.; and Moeller, H.: Measurement of D2O concentrations at tracer levels in small samples obtained from paediatric patients. J. Clin. Chem. Clin. Biochem. 26:715-721, 1988.

Graystone, J., Seitchik, J., Milch, R., Shulman, G. P., and Cheek, D.B.: Measurement of D2O in plasma water by freezing point elevation, falling drop, infra-red absorption, and gas chromatography. J. Lab Clin. Med. 69:885-892, 1967.

Halliday, D., and Miller, A.G.: Precise measurement of total body water using trace quantities of deuterium oxide. Biomed. Mass Spect. 4(2); 82-87, 1977.

Lukaski, H.C.; and Johnson, P.E.: A simple inexpensive method of determining total body water using a tracer dose of D2O and infrared absorption of biological fluids. Am. J. Clin. nutr. 41:363-370, 1985.

Mendez, J.; Prokop, E.; Picon-Reategui, E.; Akers, R.; and Buskirk, E.R.: Total body water by D2O dilution using saliva samples and gas chromatography. J. Appl. Physiol. 28(3):354-357, 1970.

Money, K.C., and Miles, W.E.: Heavy water nystagmus and effects of alcohol, Nature, 247, 404-405, 1974.

Moore, F.D.: Determination of total body water and solids with isotopes. Science, 104: 157-160, 1946.

Nielsen, W.C.; Kryzywicki, H.J.; Johnson, H.L.; and Consolazio, C.F.: Use and evaluation of gas chromatography for determination of deuterium in body fluids, J. Appl. Physiol. 31(6):957-961, 1971.

Registry of Toxic Effects of Chemical, 1985-1986, Vol. 5, D.V. Sweet, ed., DHHS, Washington,1987, p 5108.

Reaser, P.B., and Burch, G.E.: Determiantion of deuterium oxide in water by measurement of freezing point. Science, 128:415-416, 1958.

Schoeller, D.A., van Santen, E., Peterson, D.W., Dietz, W., Jaspan, J., Klein, P.D.: Total body water meassurement in humans with O18 and H2 labeled water. Am. J. Clin. Nutr. 33:2686, 1980.

Schoeller, D.A., and Webb, P.: Five day comparison of the doubly labeled water method with respiratory gas exchange Am. J. Clin. Nutr., 40: 153-158, 1984.

Schoeller, D.A.: Isotope dilution methods, in Obesity, Bjornthorp ed. Lippincott, In Press,1990.

Schloerb, P.R., Friis-Hansen, B.J., Edeleman, I.S., Solomon, A.K., and Moore, F.D.: The measurement of total body water in the human subject by deuterium oxide dilution J. Clin. Lab. Invest. 29:1296-1310, 1950.

Stansell, M.J.; and Mojica, L.: Determination of body water content using trace levels of deuterium oxide and infrared spectrophotometry. Clin. Chem. 14(11):1112-1124.

Thomson, J.F.: Biological effects of deuterium. Pergamon Press, N.Y., 1963.

Turner, M.D.; Neeley, W.A.; and Hardy, J.D.: Rapid determination of deuterium in biological fluids. J. Appl. Physiol. 15:309-310, 1960.

Wong, W.W., and Klein, P.D.: A review of techniques for preparation of biological samples for mass spectrometry measurement of hydrogen2/ hydrogen1 and oxygen18/ oxygen16 isotope ratios. Mass Spec. Reviews, 5: 313-342, 1986.

Wong, W.W., Cochran, W.J., Klish, W.J. Smith, E.O.B., Lee, L.S., Klein, P.D.: In vivo isotope-fractionation factors and measurement of deuterium- and oxygen18- dilution spaces from plasma, urine, saliva, respiratory water vapor, and carbon dioxide. Am. J. Clin. Nutr. 47:1, 1988.

MATHEMATICAL MODELING OF THE FLOW FIELD AND PARTICLE MOTION IN A ROTATING BIOREACTOR AT UNIT GRAVITY AND MICROGRAVITY

Final Report

NASA/ASEE Summer Faculty Fellowship Program 1990

Johnson Space Center

Prepared by:	Ernest J. Boyd, Ph.D.
Academic Rank:	Associate Professor
University & Department:	Mankato State University Department of Mathematics, Astronomy & Statistics Mankato, Minnesota 56002-8400

NASA/JSC

Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Biomedical Operations & Research
JSC Colleague:	Steve Gonda, Ph.D.
Date Submitted:	August 23, 1990
Contract Number:	NGT-44-005-803

ABSTRACT

The biotechnology group at NASA Johnson Space Center is developing systems for culturing mammalian cells that simulate some aspects of microgravity and provide a low shear environment for microgravity-based studies on suspension and anchorage-dependent cells. The design of these vessels for culturing cells is based on the need to suspend cells and aggregates of cells and microcarrier beads continually in the culturing medium. The design must also provide sufficient circulation for adequate mass transfer of nutrients to the cells and minimize the total force on the cells. Forces, resulting from sources such as hydrodynamic fluid shear and collisions of cells and walls of the vessels, may damage delicate cells and degrade the formation of three-dimensional structures. This study examines one particular design in both unit gravity and microgravity based on two concentric cylinders rotating in the same direction at different speeds to create a Couette flow between them. The study presents a numerical simulation for the flow field and the trajectories of particles in the vessel. The flow field for the circulation of the culturing medium is modelled by the Navier-Stokes equations. The forces on a particle are assumed to be drag from the fluid's circulation, buoyancy from the gravitational force and centrifugal force from the rotation of the vessel. The problem requires first solving the system of partial differential equations for the fluid flow by a finite difference method and then solving the system of ordinary differential equations for the trajectories by Gear's stiff method. Results of the study indicate that the trajectories in unit gravity and microgravity are very similar except for small spatial deviations on the fast time scale in unit gravity. The total force per unit cross-sectional area on a particle in microgravity, however, is significantly smaller than the corresponding value in unit gravity, which is also smaller than anticipated. Hence, this study indicates that this design for a bioreactor with optimal rates of rotation can provide a good environment for culturing cells in microgravity with adequate circulation and minimal force on the cells.

The biotechnology group at NASA Johnson Space Center is developing rotating cylindrical bioreactors for culturing anchorage-dependent mammalian cells on microcarrier beads. The design of the vessels is motivated by the need to keep the aggregates of cells and microcarriers suspended in the culturing medium, to provide sufficient circulation for adequate mass transfer of nutrients to the cells and to minimize the forces on the cells. These forces, resulting from sources such as hydrodynamic fluid shear and collisions of cells and walls of the vessels, may damage delicate cells and degrade the formation of three-dimensional structures. Several studies, such as the research reported by Cherry and Papoutsakis (refs. 1 and 2) and by Croughan, Hamel and Wang (refs. 3 and 4), have examined ways to quantitate the hydrodynamic effects damaging the cells. However, these reports do not consider the complete dynamics of a particle's motion relative to the fluid. The purposes of this study are to develop a mathematical simulation of a particle's motion and to calculate the total force on a particle suspended in a rotating cylindrical bioreactor.

To begin this study a simple geometry for the vessel was chosen in order to provide a uniform flow field throughout the vessel. The vessel consists of two concentric cylinders both 11 cm long with the outer cylinder having a radius of $r_o = 4.0$ cm and rotating at ω_o rpm's, while the inner cylinder has a radius of $r_i = 2.86$ cm and rotating in the same direction at ω_i rpm's. This narrow gap of 1.14 cm between the cylinders is completely filled with culturing medium into which particles as aggregates of cells and microcarrier beads are introduced. Figure 1 provides a schematic representation of the vessel. Cylindrical coordinates (r, θ, z) will be used to indicate positions within the vessel where r is the radial component outward relative to the cylinders with $r_i \leq r \leq r_o$, θ is the angular component measured positively in the direction of rotation of the two cylinders, and z is the axial direction oriented horizontally in a gravitational field with $0 \leq z \leq 11$. Mathematical simulations of more complex geometries, including perfusion of culturing medium, are being planned.

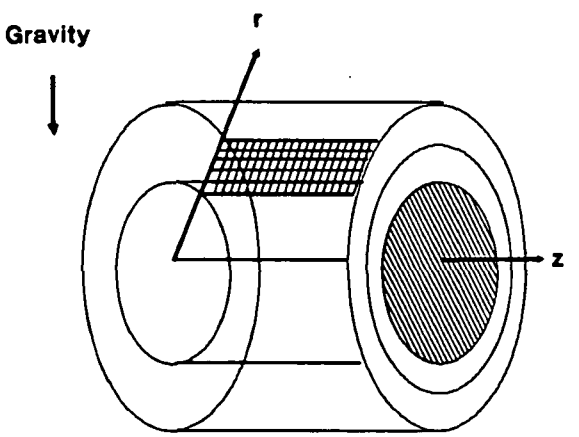


Figure 1.- Schematic representation of the rotating bioreactor.

Assumptions for this mathematical simulation are:

1. The culturing medium is a Newtonian fluid with constant density ρ_s measured in gr/cm^3 and constant viscosity μ measured in gr/cm sec;

2. Particles are spherical in shape with diameter d measured in cm and density ρ_p measured in gr/cm^3, do not interact with one another, and do not affect the flow of the culturing medium;
3. The flow of the culturing medium caused by the rotation of the concentric cylinders is a laminar, axially symmetric Couette flow and is modeled by the Navier-Stokes equations; and
4. The forces acting on a particle are drag from the fluid's circulation, buoyancy from the gravitational force relative to the difference between the densities of a particle and the fluid, and centrifugal force from the rotation of the vessel.

If (u, v, w) represent the (radial, circumferential, axial) components of the velocity of the flow field measured in cm/sec, then the equations for the flow field are:

(1) $$\frac{1}{r}\frac{\partial(ru)}{\partial r} + \frac{\partial w}{\partial z} = 0$$

(2) $$\frac{Du}{Dt} - \frac{v^2}{r} = \frac{\mu}{\rho_s}\left(\nabla^2 u - \frac{u}{r^2}\right)$$

(3) $$\frac{Dv}{Dt} - \frac{uv}{r} = \frac{\mu}{\rho_s}\left(\nabla^2 v - \frac{v}{r^2}\right)$$

and

(4) $$\frac{Dw}{Dt} = -\frac{1}{\rho_s}\frac{\partial P}{\partial z} + \frac{\mu}{\rho_s}\nabla^2 w$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial r} + w\frac{\partial}{\partial z}$$

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{\partial^2}{\partial z^2}$$

and P is the fluid pressure measured in gr/cm sec^2. Note that axial symmetry implies that all derivatives with respect to θ in the Navier-Stokes equations are zero. Boundary conditions for the flow field are the standard constraints of no slip and no penetration.

The equations for a trajectory of a particle are:

(5) $$m\frac{d^2 r}{dt^2} = -\alpha\left(\frac{dr}{dt} - u\right) - \beta g\cos\theta + \beta r\left(\frac{d\theta}{dt}\right)^2$$

(6) $$mr\frac{d^2\theta}{dt^2} = -\alpha\left(r\frac{d\theta}{dt} - v\right) - \beta g\sin\theta$$

and

(7) $$m\frac{d^2z}{dt^2} = -\alpha\left(\frac{dz}{dt} - w\right)$$

where $m = \rho_p \pi d^3/6$, $\alpha = 3\pi d\mu$, $\beta = (\rho_p - \rho_s)\pi d^3/6$ and g is the acceleration due to gravity measured in cm/sec^2. In this study initial conditions for a particle at $t = 0$ are $r = 3$, $\theta = 0$, $z = 1, 2.5, 4, 7, 8.5$ or 10, and the derivatives are all zero.

The values of the parameters for this study are $d = 0.0175$ cm, $\rho_s = 1.02$ gr/cm^3, $\rho_p = 1.04$ gr/cm^3, and $\mu = 0.0097$ gr/cm sec. These values are approximately correct for the type of medium and microcarrier beads being used by the biotechnology group at NASA Johnson Space Center. Preliminary studies consist of four trials using different rotating speeds ω_i and ω_o for the inner and outer cylinders. The speeds are listed in Table 1 below. These values provide numerical results from which optimal rotating speeds hopefully may be estimated.

TABLE 1.- SPEED OF ROTATION IN RPM'S.

Trial	Inner Cylinder, ω_i	Outer Cylinder, ω_o	Differential Rate
1	24	18	6
2	30	18	12
3	36	18	18
4	36	12	24

Dr. Yowmin David Tsao at NASA Johnson Space Center used a semi-implicit finite difference algorithm employing a hybrid scheme developed by Spalding (ref. 5) to compute the values of (u, v, w) from the system of partial differential equations in Equations 1 - 4 on a discrete grid. The grid has 61 subdivisions in the z-direction and ten subdivisions in the r-direction. Since the flow field is axially symmetric, it can be projected into this rectangular region of the (z, r)-plane. Note that v is closely approximated by the standard formula for Couette flow (ref. 6):

(8) $$v = \frac{1}{r_o^2 - r_i^2}\left[r\left(\omega_o r_o^2 - \omega_i r_i^2\right) - \frac{r_o^2 r_i^2}{r}(\omega_o - \omega_i)\right]$$

with ω_i and ω_o measured in radians. The finite difference program was advanced in time t until a steady-state flow was assumed to be reached. This data for the numerical values of the flow field on the rectangular grid was provided to the author, who used linear interpolation to produce an approximation to the complete, steady-state, axially symmetric flow field for (u, v, w) as functions of (r, z).

Figures 2 - 5 show segments of the secondary flow fields for the four trials. Printed along each streamline is the circulation time in seconds computed for the segment plotted in the (z, r)-plane. Circulation is a two-compartmental flow with counterclockwise circulation in the left half and clockwise circulation in the right half. A Runge-Kutta-Fehlberg numerical method (ref. 7) was used to generate the streamlines by solving $dr/dt = u(r, z)$ and $dz/dt = w(r, z)$. However, the zero boundary conditions along the walls of the vessel and the limitations in numerical accuracy due to the linear interpolation of the data for the flow field prevent the complete circulation from being plotted along streamlines in the neighborhoods of the end walls of the vessel. In these narrow regions circulation will be

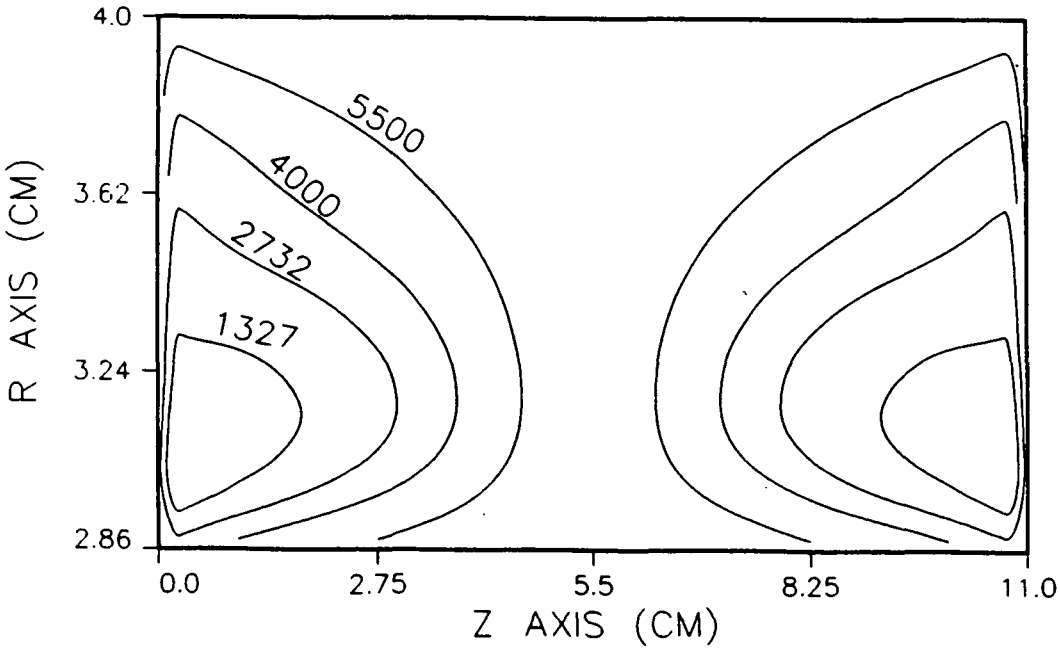


Figure 2.- Streamlines in the *zr* - plane for Trial 1
with the outer cylinder rotating at 18 rpm and the inner cylinder rotating at 24 rpm.

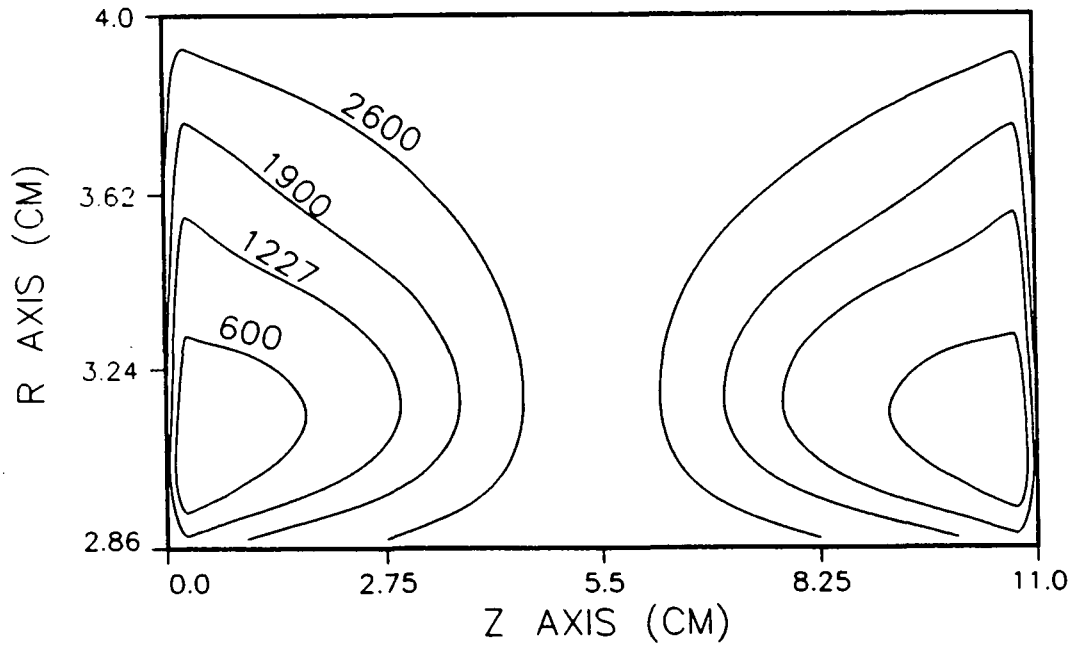


Figure 3.- Streamlines in the *zr* - plane for Trial 2
with the outer cylinder rotating at 18 rpm and the inner cylinder rotating at 30 rpm.

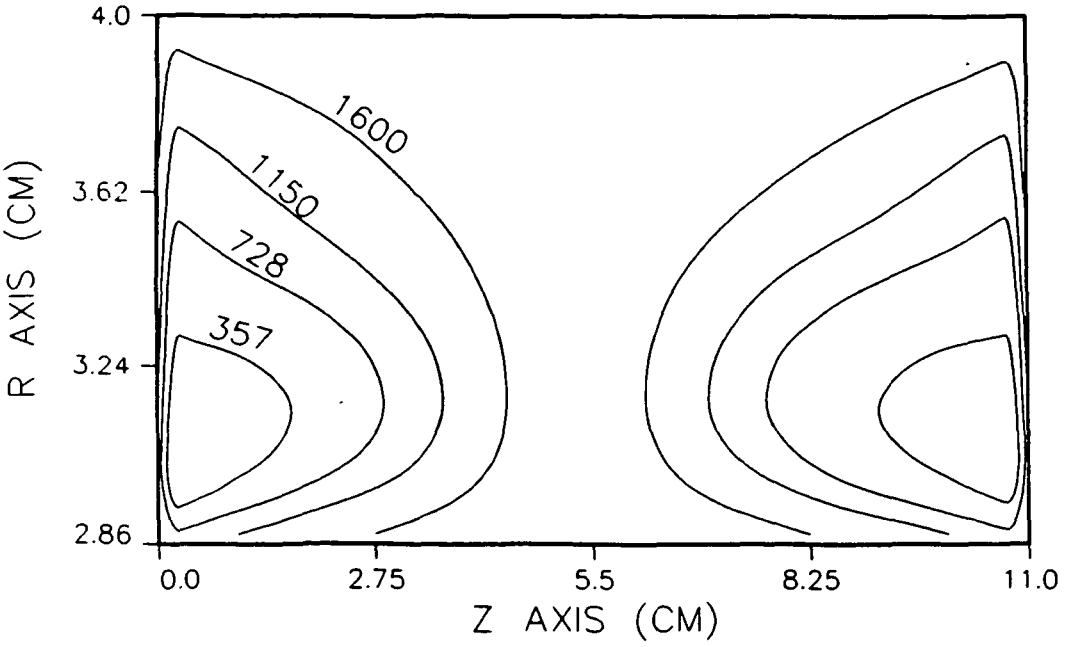


Figure 4.- Streamlines in the *zr* - plane for Trial 3
with the outer cylinder rotating at 18 rpm and the inner cylinder rotating at 36 rpm.

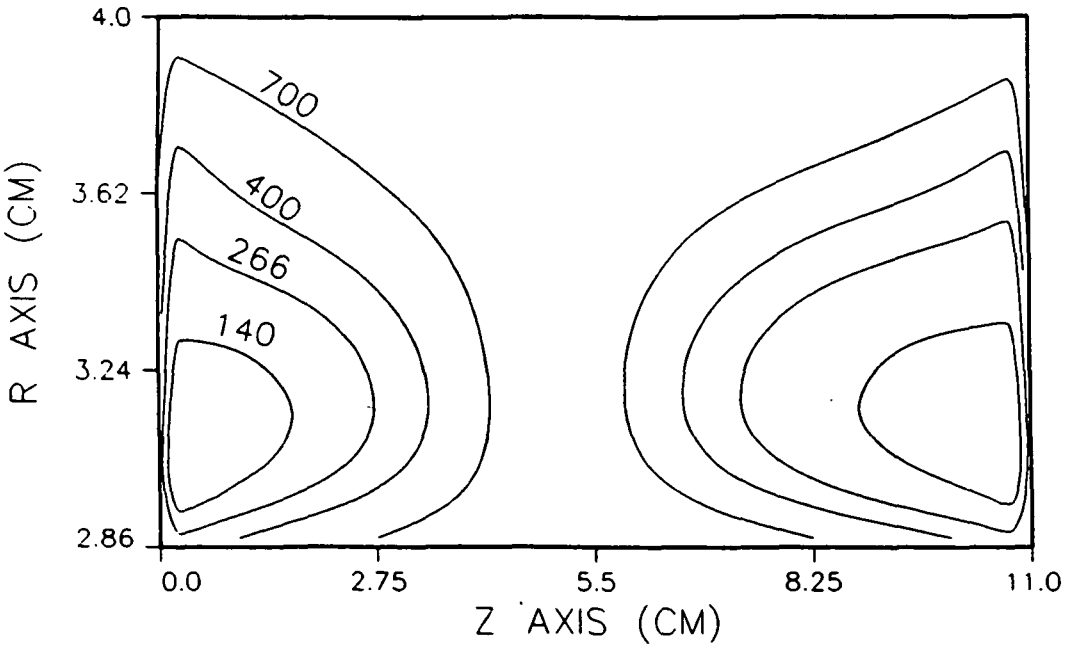


Figure 5.- Streamlines in the *zr* - plane for Trial 4
with the outer cylinder rotating at 12 rpm and the inner cylinder rotating at 36 rpm.

very slow since u and w are close to zero. These figures indicate that there may be some concern about poor circulation in the middle of the vessel. For instance, Trial 1 shows circulation time approximately equal to two hours, while increasing rpm's in Trial 4 shows circulation time decreases to approximately fifteen minutes. Higher differential rates of rotation will create a stronger secondary flow, but more studies are needed to determine what rate allows adequate time for mixing of the nutrient supply relative to the cells' metabolic requirements.

The streamlines are important for studies of mass transfer of nutrients and cellular waste products, but they do not represent the trajectories of particles as aggregates of cells and microcarrier beads. Equations 5 - 7 must be solved for the trajectories given the steady-state values of (u, v, w). This initial value problem is a two time-scale, three-dimensional, second order system of ordinary differential equations. The two time scales in this problem come from the large difference between the rate of rotation and the rate of circulation in the secondary flow. The fast time scale for the rotation of the vessel at approximately 20 rpm's is on the order of seconds, while the slow time scale for the circulation of the flow is on the order of minutes to hours as indicated above. Being a two time-scale problem, Equations 5 - 7 give a stiff system of ordinary differential equations with a 6×6 Jacobian matrix having at least two large eigenvalues and at least two small eigenvalues. Therefore, Gear's stiff method (ref. 8) was used to calculate trajectories.

In each of the four trials six trajectories are plotted in Figures 6 - 9 at unit gravity (g = 980 cm/sec^2) and in Figures 10 - 13 at microgravity (g = .0980 cm/sec^2). The initial conditions for these six trajectories are $r = 3$, $\theta = 0$, $z = 1, 2.5, 4, 7, 8.5$ and 10, respectively, with all first derivatives initially equal to zero. These figures show that a particle does not follow the streamlines precisely, but instead migrates across streamlines. If a particle is within a region of strong circulation, then it does complete its own cycle and remain suspended in the fluid. However, in most regions of the vessel a particle will not remain suspended and will hit either the wall of the outer cylinder or an end wall. Computation of a trajectory was terminated at a time when either it completed one cycle or it hit a wall of the vessel. The total lengths of time for trajectories are listed in Tables 2 - 5. In Trial 1 particles were taking more than twenty minutes to migrate to a wall of the vessel, while in Trial 4 this time frame was reduced to less than nine minutes. There is no appreciable difference in the time frames between unit gravity and microgravity. Thus higher differential rates of rotation do provide stronger circulation, but also appear to increase the frequency of interactions between the particles and the walls of the vessel. It is hypothesized that these interactions cause significant damage to the cells. Further analysis is required to quantitate the amount of damage incurred from a collision with a wall and to determine the response in a particle's trajectory afterwards.

The most noticable difference between trajectories in unit gravity and microgravity are the small, rapid oscillations in unit gravity shown in Figures 6 - 9, but lacking in Figures 10 - 13. These oscillations, which give the particles the appearance of tumbling, occur once every revolution of the flow field within the vessel and have amplitudes on the order of 0.02 cm. Figure 14 shows one example of these small deviations in the r-direction for sixty seconds. The amplitudes of these oscillations increase as r increases, since the outer cylinder is rotating slower than the inner cylinder, but are uniform in the z-direction. Intuitively, one can say that these oscillations represent a deviation in the trajectory at unit gravity from the trajectory at microgravity due to sedimentation downward in a stronger gravitational field.

The major advantage in using dynamic modeling in this study is the ability to calculate the total force on a particle as a function of time. Traditional studies have looked at fluid

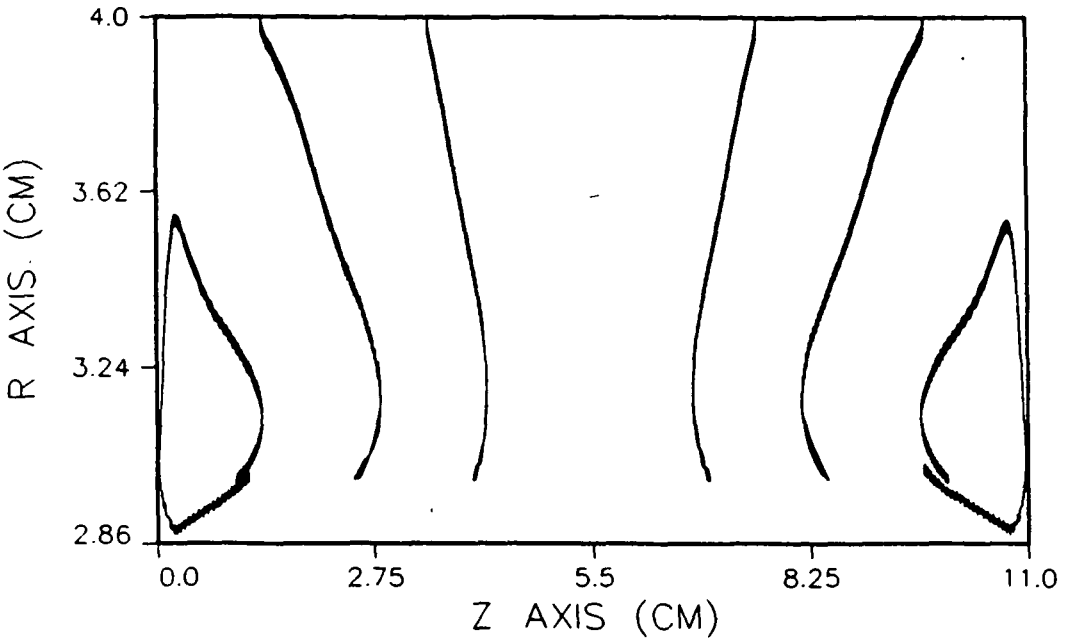


Figure 6.- Trajectories in the *zr* - plane for Trial 1
with a particle measuring 175 μm in diameter at unit gravity.

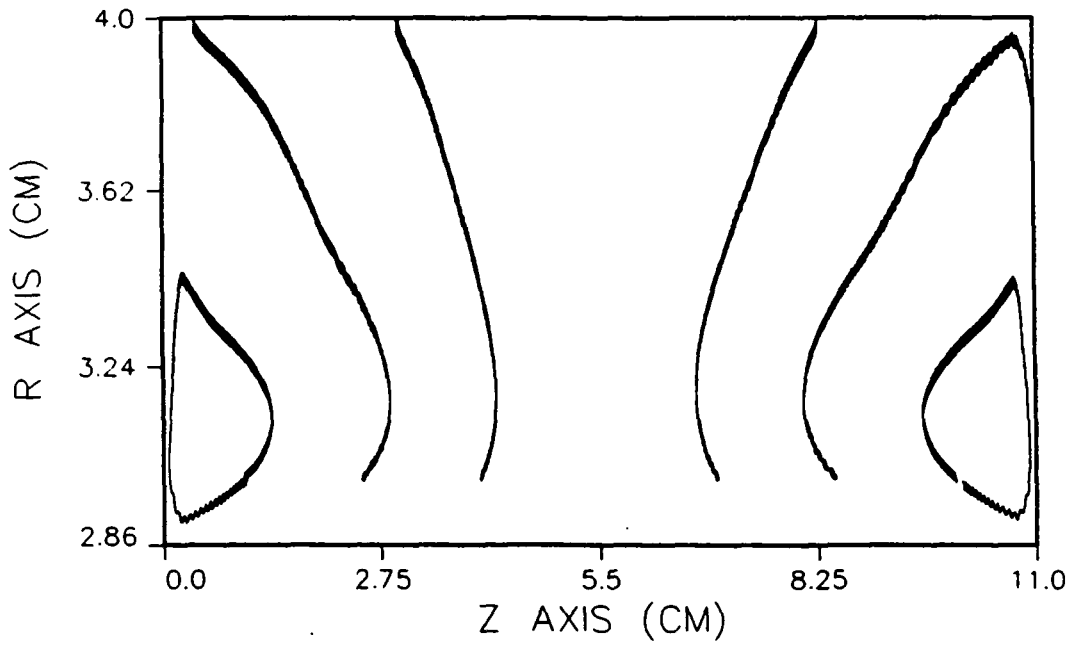


Figure 7.- Trajectories in the *zr* - plane for Trial 2
with a particle measuring 175 μm in diameter at unit gravity.

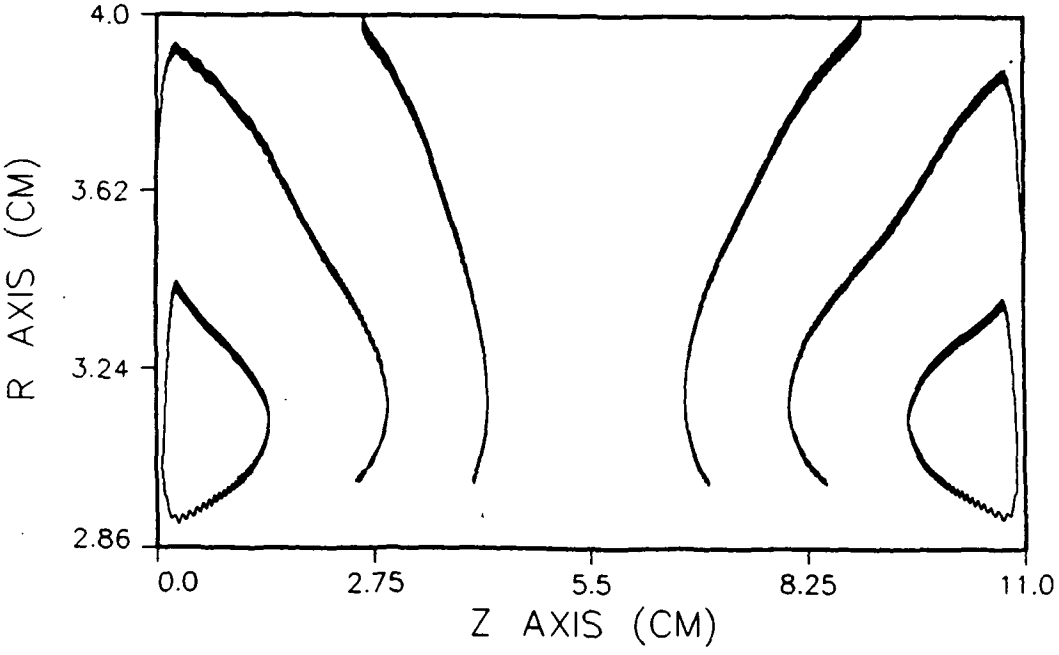


Figure 8.- Trajectories in the *zr* - plane for Trial 3
with a particle measuring 175 μm in diameter at unit gravity.

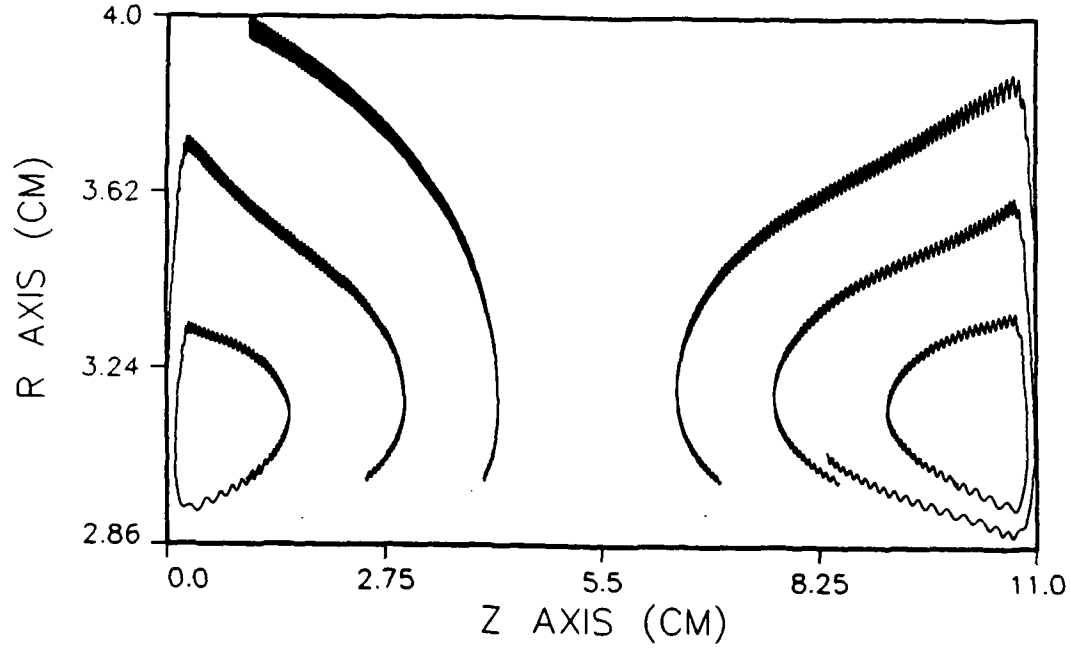


Figure 9.- Trajectories in the *zr* - plane for Trial 4
with a particle measuring 175 μm in diameter at unit gravity.

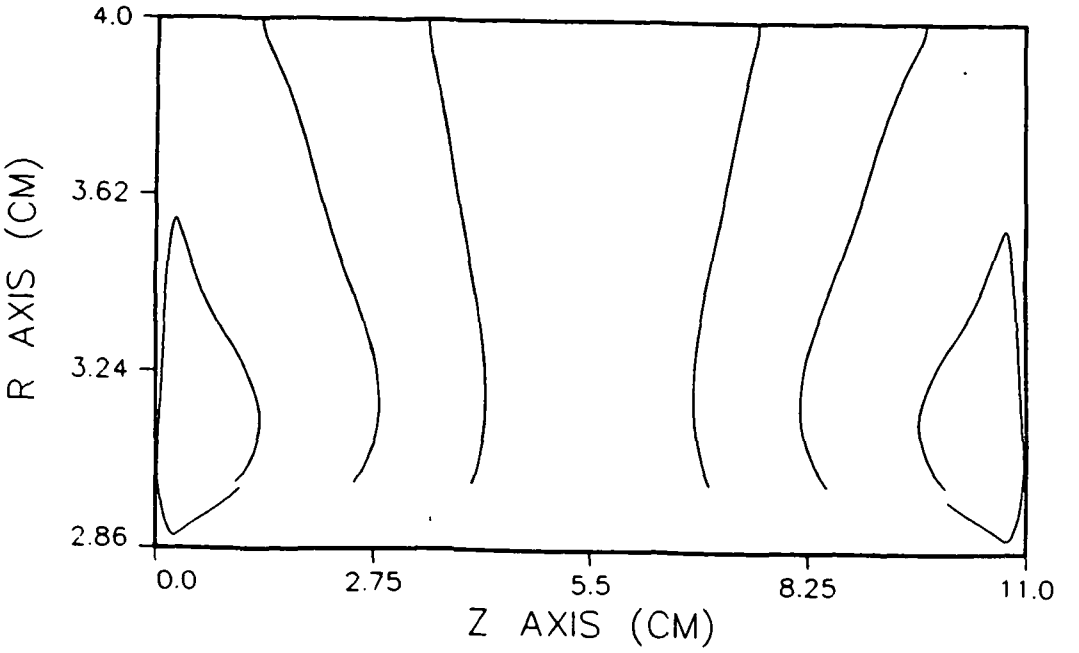


Figure 10.- Trajectories in the *zr* - plane for Trial 1
with a particle measuring 175 μm in diameter at microgravity.

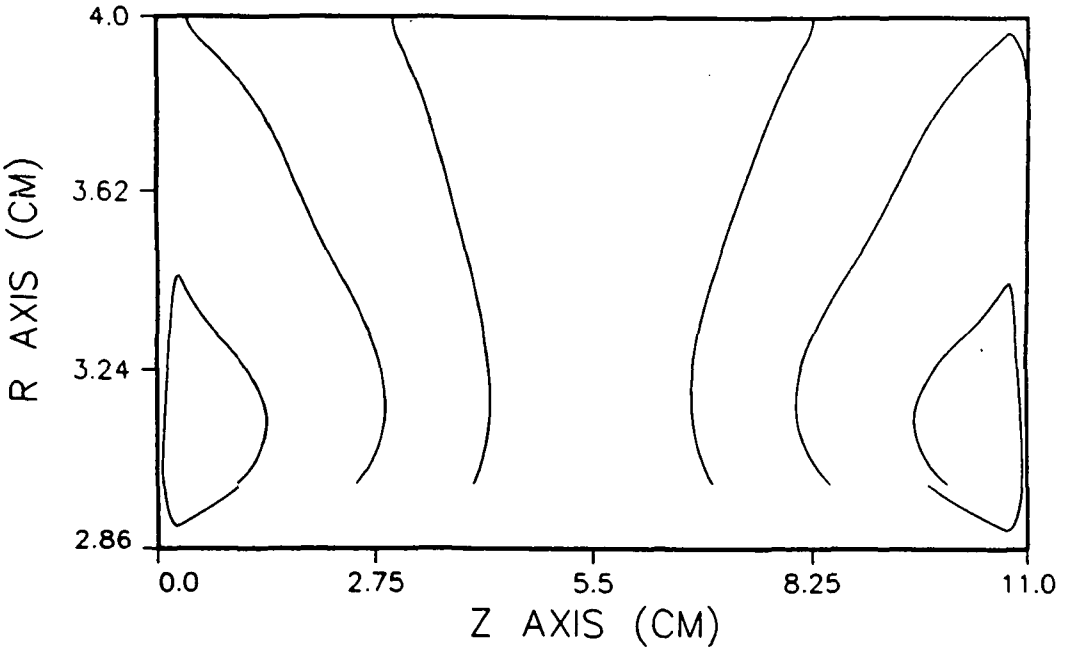


Figure 11.- Trajectories in the *zr* - plane for Trial 2
with a particle measuring 175 μm in diameter at microgravity.

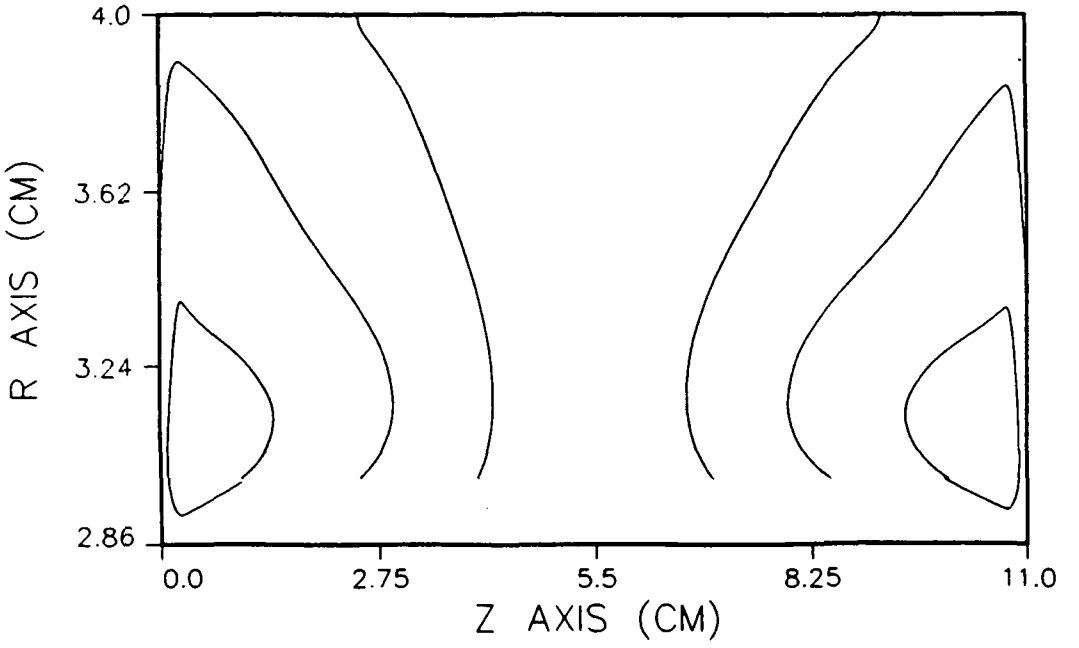


Figure 12.- Trajectories in the *zr* - plane for Trial 3
with a particle measuring 175 μm in diameter at microgravity.

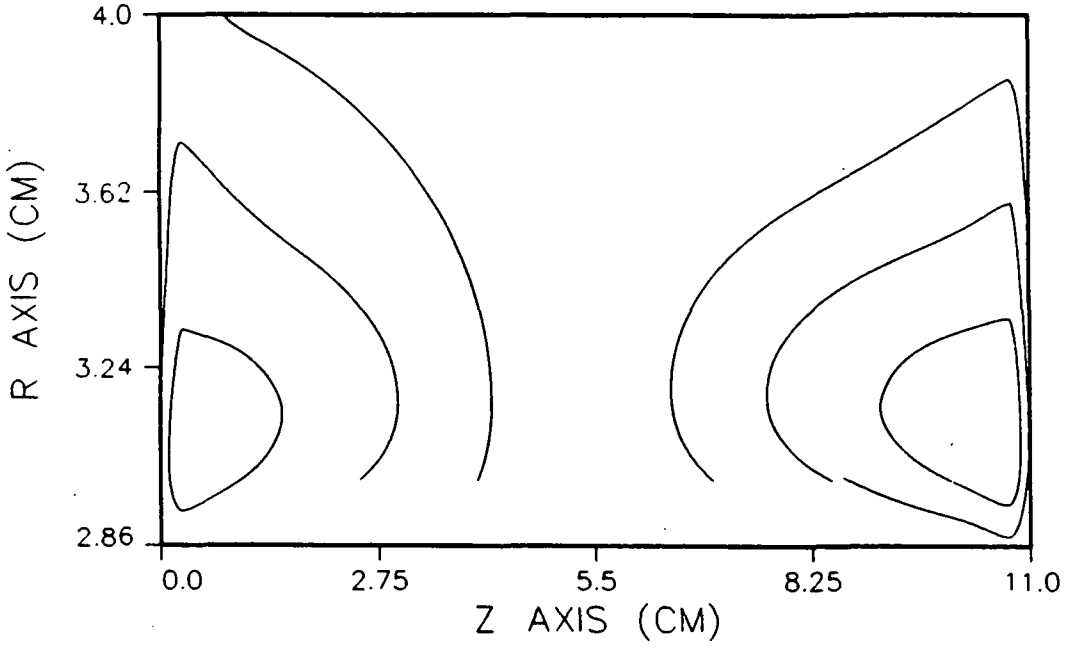


Figure 13.- Trajectories in the *zr* - plane for Trial 4
with a particle measuring 175 μm in diameter at microgravity.

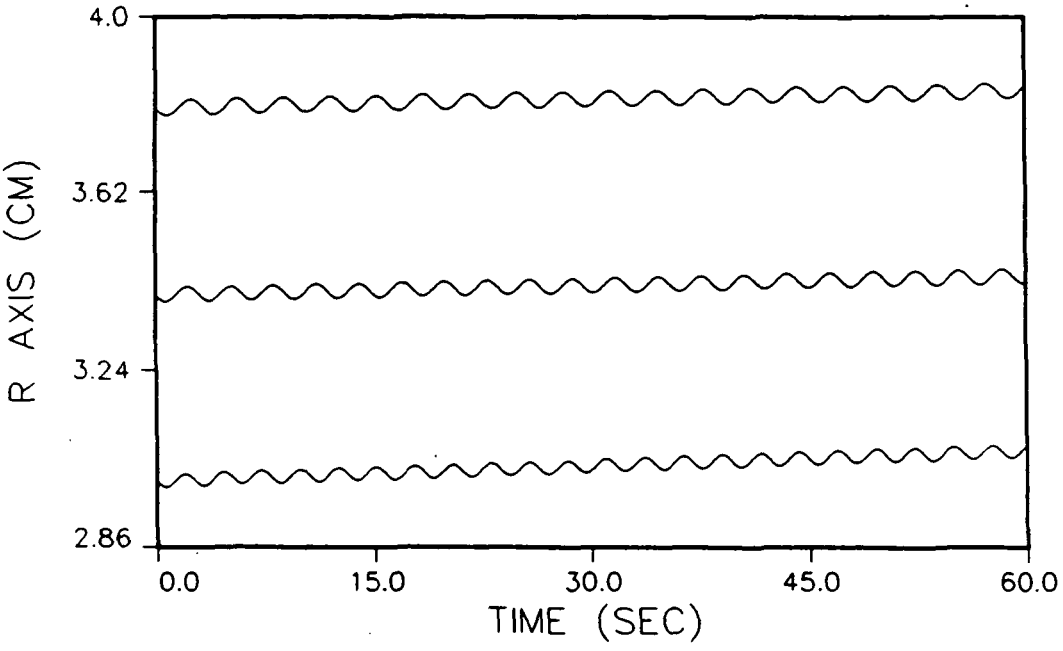


Figure 14.- Three trajectories from Trial 1
showing the r - component for sixty seconds with inital values of $\theta = 0$ and $z = 1$.

shear and drag induced by sedimentation in terms that do not reflect the dynamical response by a particle to the sum of all forces. Traditionally, from Formula 8 fluid shear relative to the cross-sectional area of a particle measured in dyne/cm^2 is given by

$$\tau = 4\mu \left| r \frac{\partial}{\partial r}\left(\frac{v}{r}\right) \right| = 8\mu r_o^2 \left| \frac{\omega_o - \omega_i}{r_o^2 - r_i^2} \right|$$

which is between 0.08 and 0.4 for the trials in this study. Drag as a force per unit cross-sectional area induced by sedimentation is

$$C_D \left[\pi \left(\frac{d}{2}\right)^2 \right] \left[\frac{1}{2} \rho_s v_\infty^2 \right]$$

where v_∞ is the terminal velocity of a particle and the drag coefficient C_D depends on the Reynolds number (ref. 9). For the trials at unit gravity in this study, drag is between 0.5 and 1.0 dyne/cm^2. The question to be answered is whether or not these values indicate the magnitude of the force on a particle in real time. In this study the total force per unit cross-sectional area on a particle can be computed in real time by calculating the vector norm of the sum of the forces in Equations 5 - 7 divided by the cross-sectional area, $\pi d^2/4$, of a particle. The maximum values and the average values over time for the total force are listed in Tables 2 - 5 for each of the six trajectories in each of the four trials at both unit gravity

and microgravity. The average force ranges from 0.0005 to 0.002 dyne/cm^2 in unit gravity and from 0.00003 to 0.0007 dyne/cm^2 in microgravity. The maximum force ranges from 0.002 to 0.013 dyne/cm^2 in unit gravity and from 0.0001 to 0.01 dyne/cm^2 in microgravity. In unit gravity drag and buoyancy seem to balance each other at approximately 0.2 - 0.4 dyne/cm^2, while centrifugal force is much smaller at approximately 0.004 - 0.007 dyne/cm^2. In microgravity the same effect is observed, but here drag and centrifugal force seem to balance each other at approximately 0.004 - 0.007 dyne/cm^2, while buoyancy is much smaller at approximately 0.00002 - 0.00004 dyne/cm^2. In general, the force on a particle in microgravity is an order of magnitude smaller than the force on a particle in unit gravity. Figures 15 and 16 show profiles in time of the force per unit cross-sectional area for sample trajectories. All of the trajectories in this study show similar profiles. Note in Figure 16 the significant increase in force experienced by the particle during the time when it is close to the end wall indicating that any damage done to the cells would most likely occur when the particle is in the neighborhood of a wall of the vessel.

TABLE 2.- RESULTS ON TRAJECTORIES IN TRIAL 1 AT 6 RPM DIFFERENTIAL.

	Unit gravity			Microgravity		
Initial Z-coord (cm)	Time (sec)	Average Force (dyne/cm^2)	Maximum Force (dyne/cm^2)	Time (sec)	Average Force (dyne/cm^2)	Maximum Force (dyne/cm^2)
1.0	1000	.0006	.0029	1000	.00005	.0003
2.5	1335	.0005	.0026	1368	.00004	.0001
4.0	1356	.0005	.0016	1389	.00004	.0003
7.0	1356	.0005	.0018	1387	.00003	.0002
8.5	1339	.0005	.0016	1371	.00004	.0003
10.0	1000	.0006	.0029	950	.00006	.0007

TABLE 3.- RESULTS ON TRAJECTORIES IN TRIAL 2 AT 12 RPM DIFFERENTIAL.

	Unit gravity			Microgravity		
Initial Z-coord (cm)	Time (sec)	Average Force (dyne/cm^2)	Maximum Force (dyne/cm^2)	Time (sec)	Average Force (dyne/cm^2)	Maximum Force (dyne/cm^2)
1.0	430	.0009	.0046	430	.00013	.0014
2.5	930	.0007	.0020	960	.00005	.0002
4.0	980	.0007	.0026	1013	.00005	.0002
7.0	977	.0007	.0029	1010	.00005	.0003
8.5	1000	.0007	.0022	1160	.00006	.0003
10.0	400	.0010	.0029	430	.00013	.0014

TABLE 4.- RESULTS ON TRAJECTORIES IN TRIAL 3 AT 18 RPM DIFFERENTIAL.

	Unit gravity			Microgravity		
Initial Z-coord (cm)	Time (sec)	Average Force (dyne/cm^2)	Maximum Force (dyne/cm^2)	Time (sec)	Average Force (dyne/cm^2)	Maximum Force (dyne/cm^2)
1.0	270	.0014	.0048	275	.00028	.0034
2.5	905	.0010	.0026	894	.00010	.0004
4.0	765	.0010	.0032	845	.00009	.0002
7.0	751	.0010	.0024	832	.00009	.0003
8.5	684	.0011	.0083	736	.00015	.0018
10.0	250	.0014	.0048	9265	.00028	.0034

TABLE 5.- RESULTS ON TRAJECTORIES IN TRIAL 4 AT 24 RPM DIFFERENTIAL.

	Unit gravity			Microgravity		
Initial Z-coord (cm)	Time (sec)	Average Force (dyne/cm^2)	Maximum Force (dyne/cm^2)	Time (sec)	Average Force (dyne/cm^2)	Maximum Force (dyne/cm^2)
1.0	125	.0021	.0107	125	.00068	.0097
2.5	337	.0016	.0053	345	.00033	.0034
4.0	541	.0014	.0036	604	.00015	.0005
7.0	432	.0016	.0063	600	.00031	.0021
8.5	265	.0019	.0132	265	.00055	.0115
10.0	125	.0021	.0125	125	.00066	.0099

In conclusion, the rotating bioreactors being developed by the biotechnology group at NASA Johnson Space Center do partially simulate microgravity with respect to the trajectories and time frame of a particle's motion. However, there is still a distinct advantage to having the bioreactor in microgravity since the total force on a particle would be significantly reduced. The total force on a particle in a rotating bioreactor is less than one might anticipate from the values reported by other researchers using different geometries to model hydrodynamic effects on cell growth. In either unit gravity or microgravity it appears that the greatest potential for damage to the cells comes when the particle is near to a wall of the vessel or actually collides with a wall. Hence, further study needs to be conducted to determine an optimal design and optimal rates of rotation to minimize the migration of particles toward the outside cylinder and still provide adequate circulation for sufficient mass transfer of nutrients and cellular waste products.

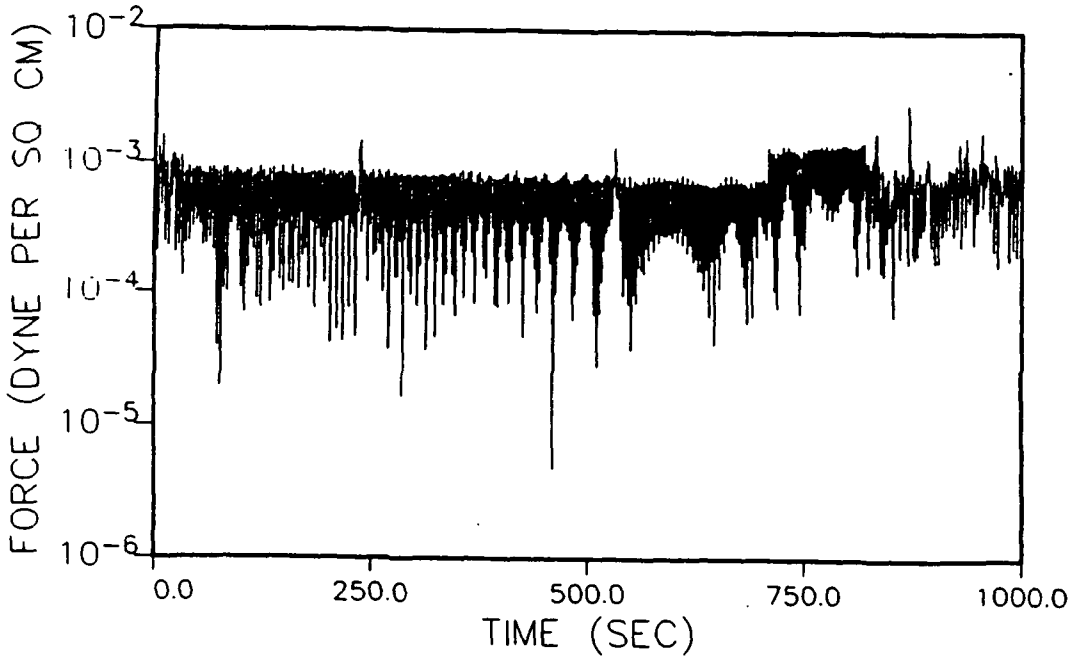


Figure 15.- Force per unit area on a particle
for the trajectory starting at $r = 3$, $\theta = 0$ and $z = 1$ in Trial 1 at unit gravity.

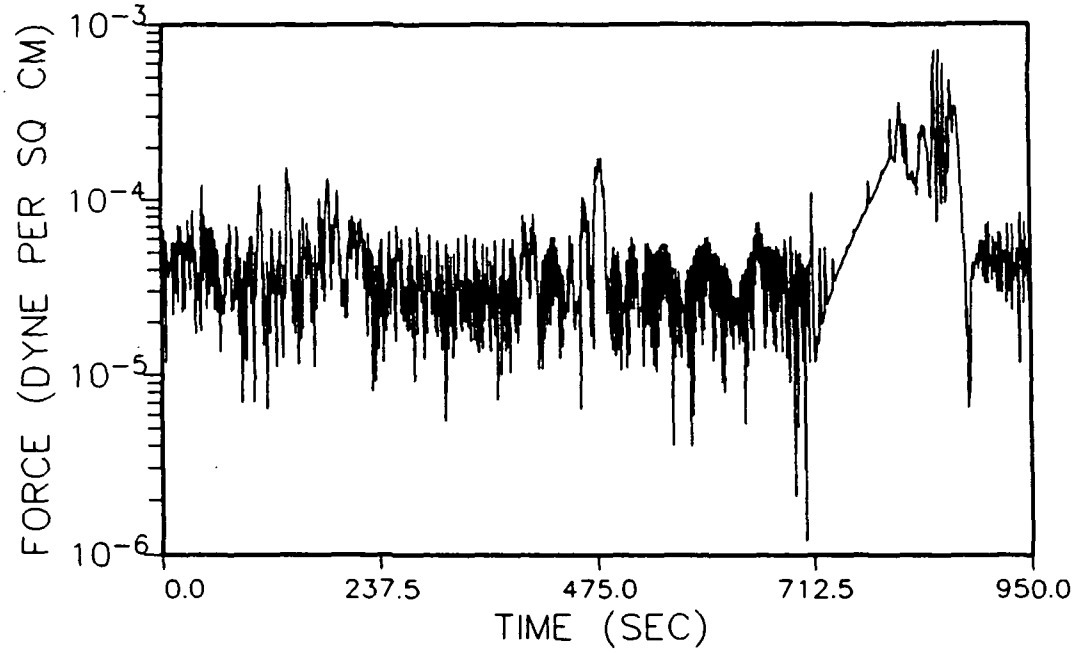


Figure 16.- Force per unit area on a particle
for the trajectory starting at $r = 3$, $\theta = 0$ and $z = 10$ in Trial 1 at microgravity.

REFERENCES

1. Cherry, R. S.; and Papoutsakis, E. T.: Hydrodynamic Effects on Cells in Agitated Tissue Culture Reactors. *Bioprocess Engineering*, vol. 1, 1986, pp. 29 - 41.

2. Cherry, R. S.; and Papoutsakis, E. T.: Physical Mechanisms of Cell Damage in Microcarrier Cell Culture Bioreactors. *Biotechnology and Bioengineering*, vol. 32, 1988, pp. 1001 - 1014.

3. Croughan, M. S.; Hamel, J. F.; and Wang, D. I. C.: Hydrodynamic Effects on Animal Cells Grown in Microcarrier Cultures. *Biotechnology and Bioengineering*, vol. 29, 1987, pp. 130 - 141.

4. Croughan, M. S.; and Wang, D. I. C.: Growth and Death in Overagitated Microcarrier Cell Cultures. *Biotechnology and Bioengineering*, vol. 33, 1989, pp. 731 - 744.

5. Patankar, S. V.: *Numericla Heat Transfer and Fluid Flow*. McGraw-Hill, New York, 1980, p. 197.

6. Schlichting, Hermann: *Boundary-Layer Theory*. McGraw-Hill, New York, 1979, p. 87.

7. Burden, R. L.; and Faires, J. D.: *Numerical Analysis*. PWS-Kent, Boston, 1989, pp. 251 - 255.

8. Burden, R. L.; and Faires, J. D.: *Numerical Analysis*. PWS-Kent, Boston, 1989, pp. 302 - 308.

9. Bird, R. B.; Stewart, W. E.; and Lightfoot, E. N.: *Transport Phenomena*, John Wiley and Sons, New York, 1960, pp. 190 - 194.

AN AIR REVITALIZATION MODEL (ARM) FOR REGENERATIVE LIFE SUPPORT SYSTEMS (RLSS)

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1990

Johnson Space Center

Prepared By:	Maxwell M. Hart
Academic Rank:	Associate Professor
University:	University of Mary Hardin-Baylor
Department:	Dept. of Mathematics and Physics Belton, Texas 76513

NASA/JSC

Directorate:	Space and Life Sciences
Division:	Solar System Exploration
Office:	Space Resources Utilization
JSC Colleague:	Donald L. Henninger
Date Submitted:	August 17, 1990
Contract Number:	NGT-44-005-803

ABSTRACT

The primary objective of the Air Revitalization Model (ARM) is to determine the minimum buffer capacities that would be necessary for long duration space missions. Several observations are supported by the current configuration of ARM. There are at least two factors affecting the buffer sizes: the baseline values for each gas, and the day-to-day or month-to-month fluctuations that are allowed. The baseline values depend on the minimum safety tolerances and the quantities of life support consumables necessary to survive the worst-case scenarios within those tolerances. Most, if not all, of these quantities can easily be determined by ARM once these tolerances are set. The day-to-day fluctuations will also require a command decision. It is already apparent from the current configuration of ARM that the tighter these fluctuations are controlled, the more energy will be used, the more nonregenerable hydrazine will be consumed, and the larger will be the required capacities for the various gas generators. All of these relationships could clearly be quantified by one operational ARM.

INTRODUCTION

Earth is, without a doubt, the most successful known example of a regenerative life support system. Because past and current space missions continue to be of relatively short duration, these life support systems all rely heavily upon expendable materials. Life support systems of many future space missions may try to imitate Earth's own system. One of the major problems encountered with this reproduction is that the amount of buffered air volume on the earth is a great many times larger, per person, than that available on a space mission.

The volume of air available to the life support system can be substantially increased by the use of stored air buffers of each important gas. However, strict mass and volume constraints for extended space missions require that these buffers be kept to a minimum. It will therefore be necessary to use automated physicochemical systems to modify the atmosphere and keep the relative abundance of the various gas constituents in balance.

METHODOLOGY

This investigation involves the development of a computer simulation model that emulates operation of the air revitalization component of an actual regenerative life support system. The added benefit of this procedure is that it also reveals the amount of replenishment from outside sources (e.g. from Earth or locally produced) that would be necessary under various configurations of the system. The particular system that is being modeled is an initially unmanned test bed facility that is now in the earliest stages of development at JSC. This JSC growth chamber is being built in support of NASA's long-duration missions to the moon and Mars. JSC's long-range plans also include design, construction, and operation of a man-rated test bed facility for verification of integrated regenerative life support systems, operations development, and crew-training. The Air Revitalization Model (ARM) is being designed around a built-in facility for evolution to allow it to keep pace with test bed upgrade.

Due to the complexity of this computer model and because of the great danger of undetected errors, it is advisable to approach it's design sequentially. Therefore development of a fully functional ARM has been subdivided into three phases. Of these three phases, the first two have now been essentially completed. Planning is currently under way for the commencement of phase three. It is hoped that phase three can be completed during summer, 1991.

Phase One

In this phase ARM will generate data on a short cyclic interval. The length of each cycle will be set for any value from a fraction of an hour up to a full day or more. ARM will compute the amount of each gas in the storage buffers and then activate the physicochemical systems whenever any controlled amount falls below a baseline, called zero, whose value is unimportant to ARM. In phase one there will be no restriction on the maximum size of any of the buffers. They will be assumed to be extremely large. The controlled atmospheric gases will be nitrogen and oxygen with the possibility of adding a baseline for carbon dioxide at a later date. Also controlled with baseline values will be water and hydrogen with the possibility of adding a baseline for methane at a later date. Available to the system, but with no baseline, will be hydrazine. Hydrazine is nonrenewable and must be resupplied when exhausted. Fortunately, however, hydrazine is often used as fuel and reserve supplies can be transferred to life support systems as needed.

ARM will be provided with a leakage variable. This value will be subtracted from the available amounts in each cycle. The user may choose any rate in either or both of two categories: 1) large leaks consisting of breathable air, or 2) small leaks in which gases are differentiated by molecular size.

Phase Two

This phase consists mainly of internal reorganization and restructuring for greater speed and efficiency. The various baselines will be coordinated and consolidated into a unified structure. This is all necessary as provision for a month-by-month emulation capability for ARM.

All essential documentation will be supplied to ARM. This includes both internal and auxiliary documentation. Many months, even years, of data will be generated by ARM and this data will be thoroughly tested for accuracy, reasonableness, and continuity from month to month.

ARM will also be endowed with a table of months and days so that actual monthly data can be emulated rather than only generic months of 30 days each. This is a necessary step before ARM can actually emulate the JSC growth chamber which will be operating on a real-world basis.

Phase Three

This phase will see the addition of maximum baseline

values. This will allow ARM to emulate a system with limited storage capacities. By adjusting these variables we can discover the reaction of a regenerative life support system to any variety of reserve storage buffer sizes. Only in this way can we finally determine the minimum safe capacities for these buffers. It also seems possible to provide ARM with two additional aspects of limitation. One would be an absolute lower boundary below the baseline. This lower boundary would represent an empty tank. The other additional limitation would be an absolute upper boundary. This, by analogy, would represent a full tank. These two limitations would open up a whole new realm of possibilities, especially in the absolute minimum. For one thing it would give us insight into the use of emergency air rations and improve our understanding of just how much is really needed for different types of emergencies.

New overlays will be added to ARM to allow year-by-year operation. This will provide insight to the long-term effect of various configurations. The entire structure of ARM would also need to be redesigned to reduce computer memory and storage requirements and enable the program to run on conventional hardware.

When the JSC test bed is completed and in operation, the various model parameters will be adjusted to reflect the actual chamber data. As more data becomes available, a database of various plant growth profiles will be incorporated into ARM to identify different reactions of the environment to the presence of a variety of types of plants.

ARM will be programmed to respond to an assortment of unscheduled "events." The importance of surviving unexpected occurrences will surely increase the minimum buffer sizes. This increase can best be calculated by modeling these events. However, once the values of these increases have been determined, they can be added (moved) below the minimum baseline and need not affect the day-to-day operation of the system.

PARAMETERS

System Constants

These values are well known constants and unless an error is discovered, these constants will not be changed. These are the values used by ARM:

A. Atomic weights
(Based on IUPAC Atomic Weights of the Elements 1981)

1.	Hydrogen	= 1.00794
2.	Carbon	= 12.01100
3.	Nitrogen	= 14.00670
4.	Oxygen	= 15.99940

B. Conversion factors
(CRC Handbook 1987)
1. Lbs. to kg = 0.45359237
2. Ft. to meters = 0.3048

C. Other factors
(Basic definitions included for exhaustivity)
1. Percent = 0.01
2. Total = 100%

D. Physicochemical system formulas
1. Sabatier $CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$
2. Oxygen generation system $2H_2O \rightarrow O_2 + 2H_2$
3. Nitrogen supply system $N_2H_4 \rightarrow N_2 + 2H_2$
4. Methane burner $CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$

Controlled Constants

These values can be easily changed as the system may require. The values currently being used by ARM are given here. However, the user should substitute his own constants for the particular system being modeled.

A. Crew Cabin
1. Crew size = 1
2. Cabin volume = 30 cubic meters
3. Air density = 1225 grams/cubic meter
4. Oxygen conversion = 0.83 kilograms/man/day
5. Conversion efficiency = 87%
6. Atmosphere composition
a. Nitrogen = 77.5%
b. Oxygen = 21.0%
c. Water vapor = 1.0%
d. Carbon Dioxide = 0.5%

B. Plant growth chamber
1. Plant growth Plots
a. Length = 1.778 meters
b. Width = 0.762 meters
c. Number = 8
2. Chamber volume = 25 cubic meters
3. Air density = 1225 grams/cubic meter
4. CO_2 conversion = 2.5 grams/square meter/hour
5. Conversion efficiency = 90%
6. Atmosphere composition
a. Nitrogen = 76.0%
b. Oxygen = 18.0%
c. Water vapor = 3.0%
d. Carbon Dioxide = 3.0%

C. Time
1. Time between readings = 8 hours
2. Plant cycle daylight = 16 hours

D. Physicochemical system efficiencies
1. Sabatier = 99%

2. Oxygen generation system = 99%
3. Nitrogen supply system = 99%

E. Leakage rates
1. Whole air type = 0.9% per day
2. Permeable membrane type = 0.1% per day

Input Variables

Some variables will often change from one run to the next. These must be chosen by the user depending on the conditions being investigated. They are:

A. Initial quantities of gases in the storage buffers
1. Nitrogen
2. Oxygen
3. Water
4. Carbon dioxide
5. Hydrogen
6. Hydrazine
7. Methane

B. Leaf Area Index (LAI)
C. Leaf growth rate

DISCUSSION

The graph in figure 1 was generated by ARM. It represents a month, January, in which no plants were growing. Thus the leaf area index has been set to zero. The nitrogen level can be seen to be dropping below an initial, arbitrarily chosen, value of 30 due to an assumed overall leakage rate of 1% per day. The water vapor level is initially decreasing very slowly below 5 because of the low percentage of water vapor contained in the atmosphere. Of course, the gases which are not contained in the atmosphere are not seen to be leaking. See hydrogen, for example, at 15. A peek at the data would reveal that oxygen is leaking by about one third as much as nitrogen. But even the graph shows that the rate at which oxygen is dropping below 10 is greater than the nitrogen decline. This is due to oxygen consumption by the one crew member. This crew member is also responsible for the increase in carbon dioxide. The zero line near the bottom of the graph represents the baseline for each gas in the buffer. It can be seen on the graph that the excess oxygen is depleted on January 10 when the oxygen trace reaches the baseline. At this time the oxygen generation system will begin generating just enough oxygen to keep the oxygen level from dropping below the baseline. The oxygen generation system is also responsible for the hydrogen increase and the greatly accelerated rate of water decline which begins on January 10.

Figure 1.- RLSS/ARM Graph for January

LAI = 0 N2H4 Used = 0.0 Kg 1.0% LEAK

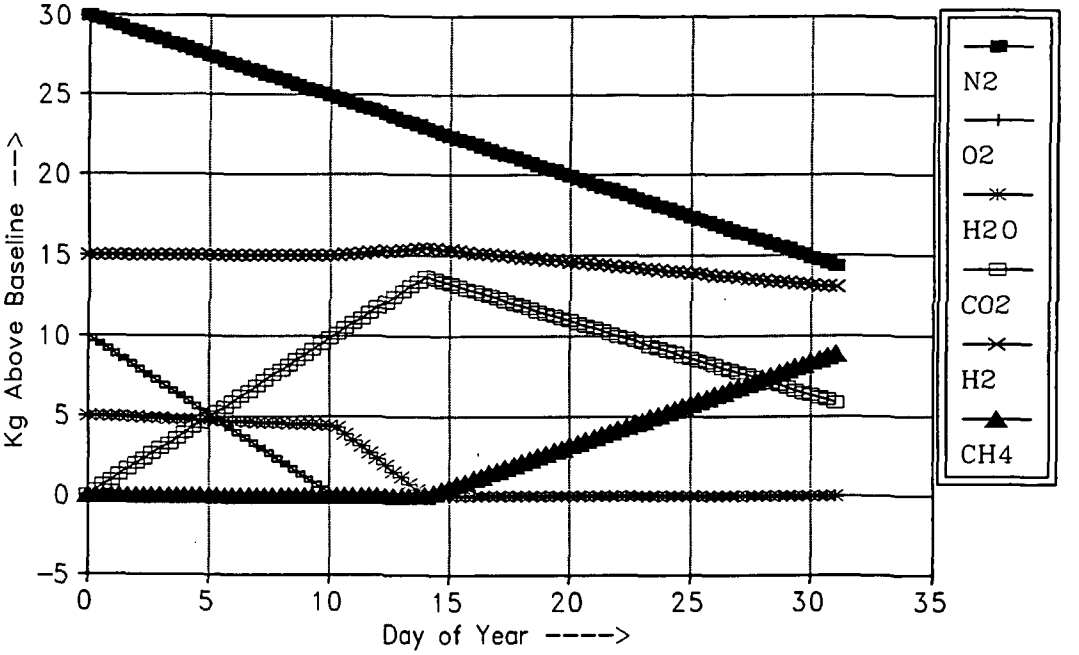
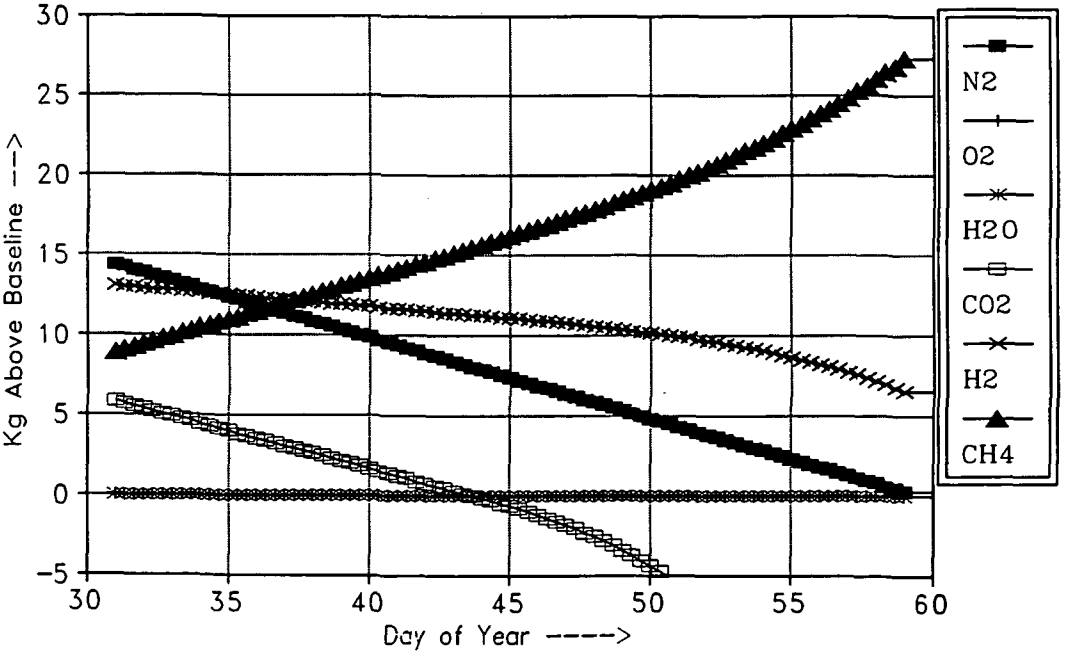


Figure 2.- RLSS/ARM Graph for February

LAI = 0 N2H4 Used = 0.0 Kg 1.0% LEAK



Notice on January 14 that the water level has dropped to zero. At this time the Sabatier process begins operation at exactly the right rate to keep the water level from dropping any further. The Sabatier process causes a decrease in the hydrogen level and the carbon dioxide level on January 14. It is also responsible for the production of methane which is seen to begin on January 14. It should be pointed out that both oxygen and water remain on the baseline for the remainder of January. Also notice, at the top of the graph, the label that indicates that no hydrazine was used during January. All of the traces end on the thirty-first, the last day of January.

The graph in figure 2 was generated by ARM immediately after the previous graph without any user intervention. These are the results for February. It uses the final values at the end of January as the initial values for February. It can be seen that the day numbers at the bottom of the graph are days of the year rather than days of the month, and each day ends at the corresponding mark.

On day 44 the carbon dioxide is depleted. But since there are no plants and since both oxygen and water levels are on the baseline, the carbon dioxide is allowed to go below the baseline rather than use precious oxygen or water to try to stop it.

Nothing noteworthy happens for the remainder of February until February 28, day number 59. On this day the nitrogen is depleted. However, this same situation is duplicated in the next series of graphs and will be illustrated at that time.

There are two very important reasons for studying the operation of ARM with an LAI of 0. Initially, it gives us an opportunity to see this operation without the further complication of input from the biological component of the life support system. This makes it easier to understand and validate ARM. Secondly, this is the mode in which ARM must operate, at least temporarily, if some misfortune destroys all of the plants. This will make it possible to determine the quantities of buffer gases that must be stored for just such contingencies.

The specifics for figure 3 are similar to those for figure 1, except that the initial values are set somewhat differently and the leaf area index is set to 4. The slow leakage of water vapor from the atmosphere is more noticeable in this graph than it was in figure 1. The wavy nature of the oxygen trace as well as that of carbon dioxide, and others to a lesser degree, is due to the day and night growth cycle of the plants. The plants cause the oxygen levels to increase and the carbon dioxide levels to decrease until January 25 when the carbon dioxide is exhausted. At this time the emulated methane burner converts just the right amount of methane to hold the carbon dioxide level at

Figure 3.- RLSS/ARM Graph for January

LAI = 4 N2H4 Used = 0.0 Kg 1.0% LEAK

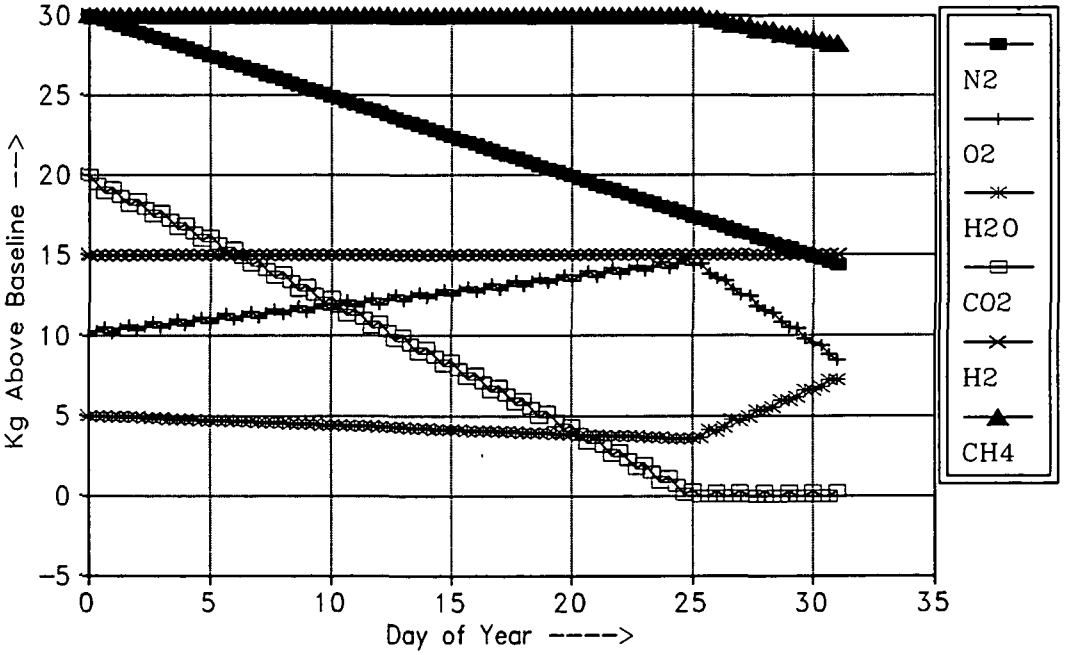
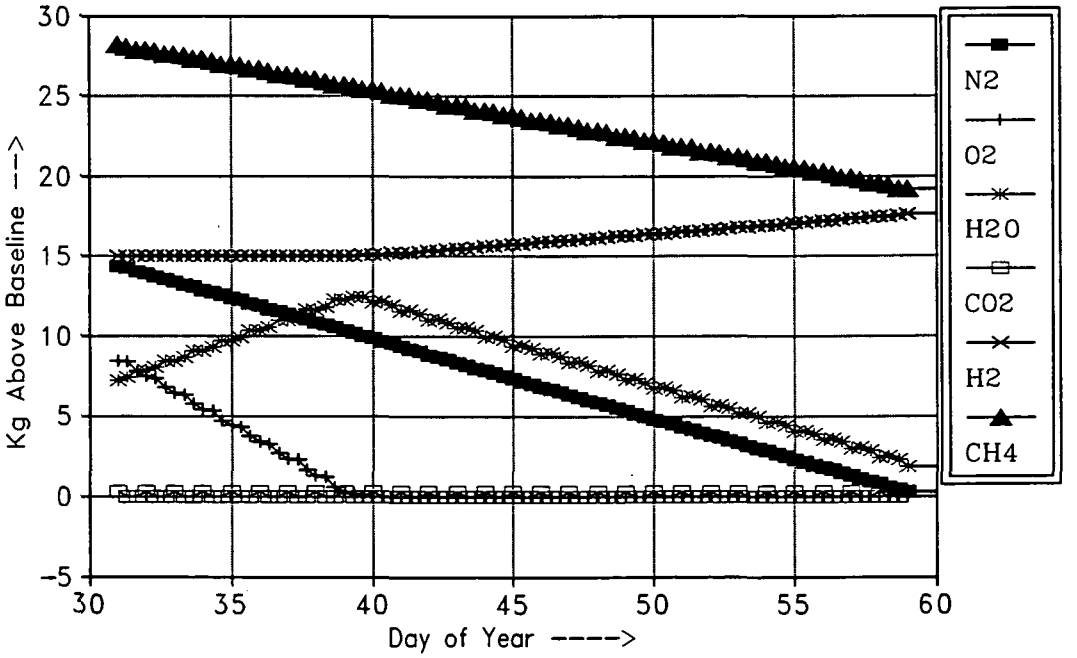


Figure 4.- RLSS/ARM Graph for February

LAI = 4 N2H4 Used = 0.0 Kg 1.0% LEAK



the baseline while the water level begins to rise as a by-product. But, the methane burner also consumes oxygen and this, too, begins to decline on January 25.

In figure 4 we see that the oxygen is exhausted by day number 39. This activates the oxygen supply system which begins using the excess water to produce oxygen and hydrogen. Even though twice as much hydrogen is generated as oxygen, the hydrogen is seen to rise very slowly. This is due to the low mass of hydrogen.

Until now, no hydrazine has been consumed. This is indicated at the top of figures 3 and 4. However, as in figure 2, the nitrogen is depleted on day 59. There are now three gases on the baseline. Since nitrogen reaches the baseline on February 28, we will not see the effects of this until March 1.

In figure 5 we first notice that the Nitrogen Supply System that was activated on February 28 consumes hydrazine and prevents the nitrogen from falling below the baseline. But, of course it has no effect on the water level, which is also falling. Thus, we see the excess water depleted on day 63. At this time we have four gases at the baseline. These four are carbon dioxide, oxygen, nitrogen, and now water. It is clear that these four cannot all be maintained as leakage continues. We have essentially the same problem as seen in figure 2 when carbon dioxide reached the baseline. We also have the same solution. Carbon dioxide is allowed to deteriorate below the baseline. You will notice on the graph of figure 5 that carbon dioxide drops below the baseline just as water is coming onto the baseline on day 63. This leaves water, nitrogen, and oxygen on the baseline. The Nitrogen Supply System continues to generate both nitrogen and hydrogen. But this is not enough to prevent the large drain on the hydrogen supply from using all of the reserve supply of hydrogen. This occurs on day 87. Now the Nitrogen Supply System is called upon to balance the hydrogen drain caused by maintaining the water supply at the baseline. As we saw, the drain on the water supply was caused by a shortage of oxygen. This, in turn, was caused by a lack of carbon dioxide. We will come back to the carbon dioxide in a moment. On day 87, in order for the Nitrogen Supply System to balance the hydrogen, it must also generate huge amounts of nitrogen. This is why the nitrogen level is seen sailing off the top of the graph on day 88.

The system is now in desperate straits. During the month of March the RLSS used 120 kg of hydrazine, mostly in the last few days. During the month of April the system is able to keep hydrogen, oxygen and water at the baseline, but only with a huge consumption of hydrazine as indicated by the label at the top of figure 6, almost 26 metric tons. This indicates that the physicochemical system will do everything it can to maintain the baseline integrity until

Figure 5.- RLSS/ARM Graph for March

LAI = 4 N2H4 Used = 120.2 Kg 1.0% LEAK

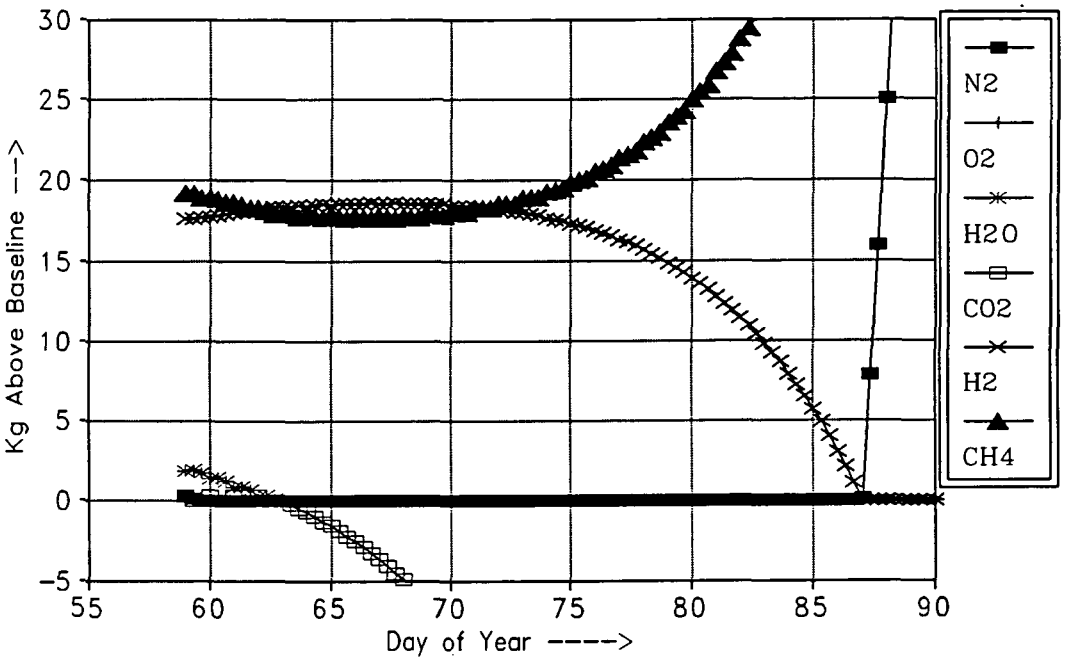
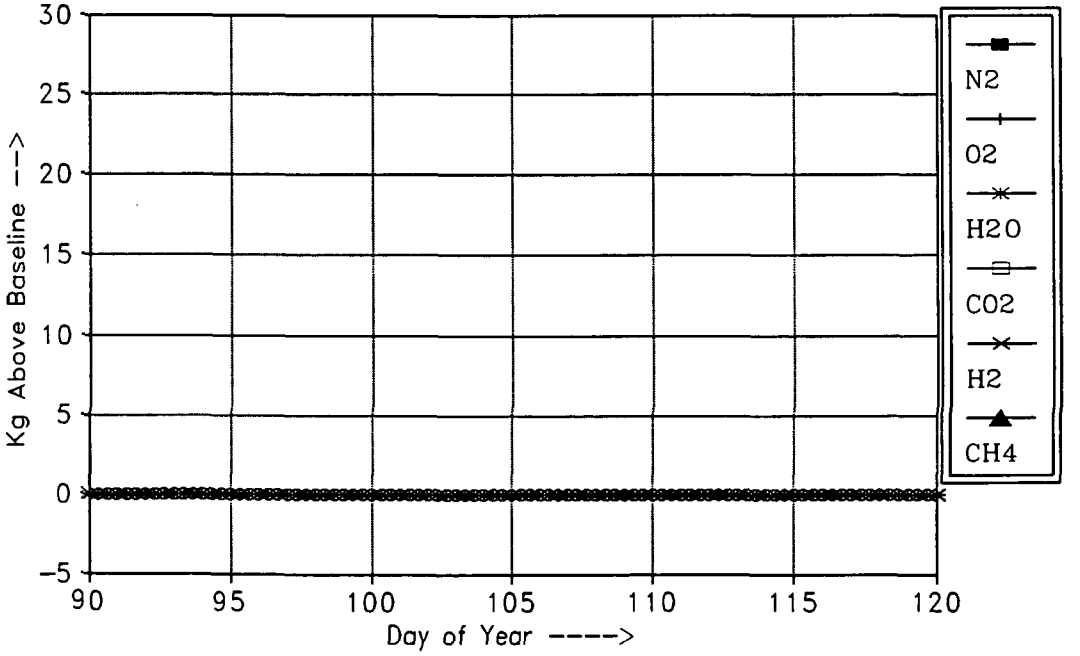


Figure 6.- RLSS/ARM Graph for April

LAI = 4 N2H4 Used = 25959.6 Kg 1.0% LEAK



it totally runs out of resources. We can only hope that never happens.

Now, returning to the carbon dioxide. We have seen that this whole breakdown of the system was first initiated by the shortage of carbon dioxide. There are auxiliary sources of carbon dioxide that can prevent the system from running down once they are utilized. This includes the recycling of waste and the eating and digesting of food which will likely be resupplied from Earth. If these factors are figured in, we will see a totally different end result. It is therefore very likely that a future version of ARM will include these features.

Important Discovery

This procedure is probably known, but it was a challenge for me. In order to keep the RLSS in balance it is quite important to poll the different gas supply systems in an appropriate sequence. So far, I have found only one sequence that works. This is an important understanding, not only for a simulation model, but also for a computer control system. The sequence used by ARM is given in table 1. Ceiling checks will not be done by ARM until it reaches phase three. They are included here for reference only. If the value in the floor check is negative, the corresponding system will be used to bring it up to the baseline. If the value in the ceiling check is above established maximums it will be consumed until brought below the ceiling. It should be understood that these systems can run simultaneously, but in three shifts. Systems 1 through 4 can run together as long as each knows the results of all previous systems and stops polling in sequence. Systems 5 through 7 and then systems 8 and 9 can repeat the process. There should, however, be no time lapse between the three shifts, especially if the leakage rate is high.

It can be seen from this table that the only products which can be generated above its ceiling will be methane, water, and nitrogen. Excess water can be stored in liquid form and excess nitrogen can be used to increase atmospheric volume or pressure to some degree. Excess methane beyond storage capacity could be vented or used as fuel.

The only products which can run short are water, carbon dioxide and hydrazine. Hydrazine is nonrenewable and must be resupplied from sources beyond RLSS domain, anyway. Similarly, resupplied food, to supplement the edible biomass provided by the plants, will help alleviate the shortages of both carbon dioxide and water when the resupplied food is digested by the crew.

TABLE 1.- SEQUENCE OF OPERATION

#	Name*	Floor Check	Ceiling Check	Byproduct	Requisite
1.	NSS	N_2 <0		H_2	N_2H_4
2.	S	H_2O<0	CO_2>Max	CH_4	CO_2 & H_2
3.	MB	CO_2<0	O_2 >Max	H_2O	CH_4 & O_2
4.	OGS	O_2 <0	H_2O>Max	H_2	H_2O
5.	S	H_2O<0	H_2 >Max	CH_4	CO_2 & H_2
6.	MB	CO_2<0 and H_2O>0 or O_2>0	CH_4>Max	H_2O	CH_4 & O_2
7.	OGS	O_2 <0	H_2O>Max	H_2	H_2O
8.	S	CH_4<0	H_2 >Max	H_2O	CO_2 & H_2
9.	NSS	H_2 <0		N_2	N_2H_4

*NSS = Nitrogen Supply System
S = Sabatier System
MB = Methane Burner
OGS = Oxygen Generation System

CONCLUSIONS

The Air Revitalization Model (ARM) seems to indicate, on the basis of minimum baseline constraints only, that Regenerative Life Support Systems (RLSS) can be held on or above all of the minimum baselines for very long periods, months or even years. If this feat can be duplicated for maximum baselines, then only very small buffer storage would be necessary except that which may be required for recovery from catastrophic events. However, there is a price. This model does not track the energy requirements, but it does show a need for a continuous resupply from outside sources, such as from Earth or from local resource extraction. Under the present configuration of ARM, resupply would be required only for carbon dioxide, water and hydrazine. The carbon dioxide and water could, of course be supplied in the form of food stock, and the hydrazine does not differ from hydrazine fuel.

Under this configuration, ARM also indicates that the only large storage buffers beyond those necessary for the baseline amounts would be for storage of excess methane, water and nitrogen. However, these storage requirements are diminished if resupply is reliable.

It remains for future enhancements of ARM to determine the effects of additional constraints on the operation of RLSS. Some future enhancements which may be considered for annexation to ARM include:

1) maximum baseline constraints
2) absolute minimum constraints
3) absolute maximum constraints
4) year-to-year overlays
5) plant profiles
6) unscheduled "event" scenarios
7) resupply schedules
8) waste oxidation
9) liquid water recycling
10) tracking of gas production
11) tracking of gas usage totals
12) tracking of energy usage totals

It would seem that there is still much to be done.

PHOTOGRAPHIC IMAGE ENHANCEMENT

Final Report

NASA/ASEE Summer Faculty Fellowship Program-1990

Johnson Space Center

Prepared by: Gerald E. Hite

Academic Rank: Associate Professor

Universiy & Department: Texas A&M Univ. at Galveston
Dept. of Marine Science
Galveston Texas 77553-1675

NASA/JSC

Directorate: Space and Life Sciences

Division: Solar System Exploration

Branch: Space Shuttle Earth Observation Office

JSC Colleaque: David E. Pitts

Date Submitted: August 10, 1990

Contract Number: NGT-44-005-803

Abstract

Deblurring capabilities would significantly improve the scientific return from Space Shuttle Crew acquired images of the Earth and the safety of Space Shuttle missions. This summer, deblurring techniques were developed and demonstrated on two digitized images that had been blurred in different ways. The first was blurred by a gaussian blurring function analogous to that caused by atmospheric turbulence while the second was blurred by improper focussing. It was demonstrated in both cases that the nature of the blurring i.e. gaussian and Airy, and the appropriate parameters could be obtained from the Fourier transformation of their images. The difficulties posed by the presence of noise necessitated special consideration. It was demonstrated that a modified Wiener frequency filter judiciously constructed to avoid over emphasis of frequency regions dominated by noise resulted in substantially improved images. Even though the deblurred images were similar to the original unblurred images, their Fourier transformed images were similar not as similar, indicating that more refined techniques applied to the Fourier image could result in further improved images. Several important areas of future research were identified. Two areas of particular promise are the extraction of blurring information directly from the spatial images and improved noise abatement from investigations of select spatial regions and the elimination of spike noise.

Introduction

The Space Shuttle Earth Observation Office serves as the repository of over 100,000 photographs form the Manned Space Flight Program. Photographs are taken to monitor both the lift-off and the landing of the Space Shuttle, and to document environmental, geological, meteorological and oceanographic phenomena observed from orbit. Inadvertently a number of these are blurred by improper focussing, relative camera-object motion, atmospheric distortions i.e. shock waves or thermal gradients, etc. Atmospheric distortions are particularly annoying since they restrict NASA's ability to monitor the effects of lift-off on the shuttle and the decent of the orbiter's wheels during landing. To gain access to important engineering and scientific information contained in blurred photographs, it is necessary to apply image processing techniques. The more sophisticated of these involve digitization of the raw image and subsequent computer processing of the degitized images or their Fourier transforms.

The research accomplished as a Summer Faculty Fellow involved the development and implementation of a deblurring procedure to handle a) the digital gaussian blurring which is analogous to that caused by atmospheric turbulence and b) blurring due to improper focussing e.g. actual focal lengths differing from expected values due to thermal effects, etc. Although there is no apparent indication of the source of blurring in the spatial images, there are definite indications of the source of blurring in the Fourier (transformed) images. Thus to improve a spatial image one first diagnoses the Fourier images for signs of the source of blurring and then used that information to construct a frequency "filter" that reduces the effects of blurring in the Fourier image which in turn would produce a sharper spatial image. However, the presence of unknown noise in the original images and increased by the digitalization process, necessitates special consideration. In particular, even though blurring reduces the high frequency components compared to the lower frequency ones, simple enhancement of the higher frequency amplitudes will not necessarily improve the image, since it could increase the overall noise to signal ratio by introducing high frequency amplitudes that are noise dominated. The trick is to judiciously emphasize only those frequency regions where the signal to noise ratio is acceptable. There are many excellent text dealing with digital processing (1,2,3) and discrete Fourier transforms(4,5) that provide the background for the work described in this report.

II Theory and Approach

II.1 Blurring as a Convolution Process: the Point Spread Function

In general a blurred image is the result of light that would have arrived at a particular point being spread over a number of points. If the distribution is the same for all points in the image and if the light is incoherent then in the absence of noise the final intensity at a point (pixel) m,n would be given by

$$\sum_{m',\,n'} \mathrm{P.S.F.}_{m-m',\,n-n'}\; f^{s}_{m',\,n'}$$

where P.S.F. is the point spread function, i.e. the image of an ideal point source in the object plane, $f^{s}_{m,n}$ is the sharp image that would have formed in the absence of blurring and noise.

For the images of interest the digitization process yields an image made up of NxN pixels (N=512). Due to the blurring bringing light rays into the image region that would have otherwise not nave entered it, the summation in this expression should include m"s and n"s outside the range of 0 to N-1. Nevertheless, the sum is usually assumed to be limited to this range to avoid having to deal with an "underdetermined problem" (6).

Noise is always present in the image. It can be the result of electronic cross-talk, round-off errors in the digitization process, random, inhomogeneous chemical processes in the original film, etc. Noise can be divided into two contributions: one independent and the other dependent on the signal, i.e. f^{s}(7). In the latter case it could be incorporated into the point spread function. Since we have little knowledge of the nature of the noise, except that it is usually dominant at high frequencies, we will describe it by an unknown function $n_{m,n}$ and write the expression for the blurred image as

$$f^{b}_{m,\,n} \;=\; \sum_{m',\,n'} \mathrm{P.S.F.}_{m-m',\,n-n'}\; f^{s}_{m',\,n'} \;+\; n_{m,\,n} \qquad (1)$$

The next step is to invert the sum in equation 1 and solve for f^{s}. Since the summation is usually referred to as convolution, the inverse process is referred to as deconvolution. The basic technique is to expand all functions in terms of a complete set of basis functions which will turn the summation into a simple product of the expansion coefficients.

II,2 Fourier Transformations

For simplicity we use the traditional discrete Fourier expansion

$$f_{m,n} = \sum_{k,l} F_{k.l} \exp 2\pi i(km + ln)/N \qquad (2a)$$

where there are N values of k and l.
The inverse Fourier transformation is then given by

$$F_{k,l} = (1/N^2) \sum_{k,l} f_{m,n} \exp -2\pi i(km + ln)/N \qquad (2b)$$

In the following we will use lower case letters , e.g. f, for the spatial images and upper case letters, e.g. F, for the Fourier transformed image. The Fourier transform of the point function,P.S.F., is called the modulation transfer function,M.T.F..

In terms of the Fourier transformed variables equation 1 takes the form

$$F^b{}_{k,l} = F^s{}_{k,l} \cdot M.T.F._{k,l} + N_{k,l} \qquad (3a)$$

Deleting the common subscripts and solving this equation for the Fourier transform of the sharp image function yields

$$F^s = F^b/M.T.F. - N/M.T.F. \qquad (3b)$$

II.3 Noise Considerations

If the modulation transfer function and the noise, N, were both known then equation 3b could be used to find F^s and a subsequent inverse Fourier transformation would yield a sharp image. For the sake of discussion let us assume that we have a reasonably good understanding of the blurring process and have a good approximation to P.S.F. and thus to M.T.F. Most likely the blurring extends over a few spatial pixels and M.T.F. will be small at high frequencies, i.e. large values of k and l. Blurring due to improper focussing, relative motion, atmospheric turbulence are in general of this type. The first two sources of blurring have M.T.F.'s that vanish on contours in the k,l space. Where ever M.T.F. is small the contribution from noise, N/M.T.F., in equation 3b can be significant. Consequently, approximating F^s by F^b/M.T.F.can result in large noise contributions. The normal procedure to avoid this problem is to multiply F^b by a "filter" function which uses information about the noise to approximate 1/M.T.F. where noise is

unimportant and avoids contributions from regions where noise dominates. The Wiener filter is the most common filter discussed in the literature (2,3,4,6):

$$W^{-1} = M.T.F. \cdot (1 + |N/F^{s} \cdot M.T.F.|^{2})$$

As it stands this expression requires knowledge of the amplitude of the function to be calculated and the unknown noise contribution. A simple way to suppress noise is to replace N/F^{s} with a constant (3) or a constant times a gaussian, e.g. C exp $c(k^2+l^2)$. In the former case the constant could be chosen to be the value of M.T.F. at frequency in the range where the maximum value of W is desired. A gaussian would do a better job of modelling N/Fs than a constant but it involves finding appropriate values for two instead of one parameter.

Another approach is to notice that when M.T.F. vanishes F^{b} is equal to N. Consequently, if F^{b}_{o} is the value of F^{b} where M.T.F. = 0, then $(F^{b} - F^{b}_{o})/M.T.F.$ would avoid the introduction of excessive noise, since - F^{b}_{o} removed the noise.

III Applications : A Tale of Two Images

In this section we discuss two applications of the deblurring techniques described in the previous section. The first application involves an image blurred digitally by convolution with a gaussian spread function. This is the digital analogy of atmospheric turbulence for long exposure times(1,8). The second involves an image blurred by defocussing the lens of the digitizer and thus is an example of blurring due to improper focussing.

III.1 Enhancement of an Image Blurred by Gaussian Convolution

Figure 1a shows the original image and figure 1b shows the image blurred by convolution with a point spread function which is gaussian within a 5 by 5 square of pixels. Figure 1c shows an enhanced image obtained by applying a filter to the Fourier transform of the blurred image. It should not be considered an ultimate product but simply a demonstration of the results that could be obtained without expenditure of the time needed to fine tune the enhancement process. Figures 2a,b and c show the Fourier transformed images corresponding to those shown in figure 1. Clearly the Fourier image (figure 2c) reconstructed from the Fourier transform of the blurred image (figure 2b) is an improvement but still differs considerably from the Fourier transformation of the sharp image (figure 2a) especially in the intermediate to

high frequency region. Figure 3 shows the inverse of the deblurring filter used to obtain figure 2c from figure 2b.

There were three important aspects in constructing the deblurring filter:

1. Only in a 5 by 5 spatial pixel region was the P.S.F. non-zero.
2. Inside that region the P.S.F. could be described by a single coefficient in the exponent of the gaussian.
3. The filter would have to deviate sufficiently form 1/M.T.F. in the high frequency region where 1/M.T.F. would be large and noise could be expected to dominate.

The Fourier images shown in figures 2a and 2b were used to obtain the coefficient for the "gaussian" M.T.F. This was done by taking the Fourier transform of their ratios which in the absence of noise would be the P.S.F. Although the resulting function did not vanish outside the 5 by 5 pixel region, presumably due to noise, there was sufficient accuracy within that region to determine the coefficient within a 5% variance.

Knowing that in the continuous case the Fourier transform of a gaussian is again a gaussian, a preliminary filter was constructed using a gaussian M.T.F. The resulting image was unrecognizable, presumably due to high frequency noise. Next a one constant Wiener-like filter (see section II.3) which emphasized the medium to low frequencies ($N/\sqrt{6}$) was used. This gave a reasonable but rather grainy image. The final result shown in figures 1c and 2c was obtained by replacing the gaussian M.T.F. in the Wiener-like filter by the M.T.F. resulting from a P.S.F. having only a gaussian distribution within a 5 by 5 pixel region. The squareness of this pixel region gave extra emphasis to the high frequency components near the axes and consequently removed most of the graininess. Further refinements could be expected from a more judicious choice of the constant in the filter or optimization of a two constant filter (see section II.3).

III.2 Enhancement of an Image Blurred by Improper Focussing

Figures 4a, b and c show the original, blurred and the first order restored image and figures 5a, b and c show their respective Fourier transforms. The first two images were made using the video digitization system. Though it is no evident from the figures as shown, the top 1/16 of both images is missing as a result of the digitization process. Such spatial-domain truncation causes over shooting or ringing in subsequent filtered image (p. 26 ref. 1). A crude attempt was made to fill in these region, which replaced a few strong $l \approx 0$ constant frequency by the scratchy like region of with N/16 centered on the l axis as shown in figure 5b. It is thought that remnants of this spatial domain truncation are the cause of the ringing seen in figure 4c.

Knowing that the blurring process is due to defocussing one can expect to see the zero contours of the appropriate M.T.F. in the Fourier transform of

the blurred image. If the P.S.F. is a uniform disk, the the M.T.F. is the discrete equivalent of the Airy function. A subroutine was then written to generate Airy-like M.T.F.'s for a uniform solid disk of variable radius surrounded by an annulus whose value decreases from that of the disk to zero at a variable outer radius. The halo visible where F^b drops down to the background noise, was used to estimate the radius of the disk. Then using this estimate, Airy-like M.T.F.'s were generated. A disk with a diameter of about five pixels and a fuzzy annulus of two pixels width was selected. Comparing it to the F^b shown in figure 5b, it was seen that the first zero contour was in the halo region and that the second zero contour just enclosed the region where F^b was still significant. Thus the Fourier transform of an image blurred by improper focussing contained sufficient information to identify by visual inspection an appropriate M.T.F. However, there were other M.T.F.'s resulting from varying the disk radius and the annulus width that had similar zero contours. The one chosen gave slightly better results although the fuzzy annulus had little of no effect on the ringing. Having an appropriate M.T.F. , a Wiener-like filter was chosen that gave maximum emphasis in a frequency circle of radius 80 pixels. Clearly the resulting. Fourier image represents an improvement but is not identical to that of the unblurred image. Perhaps the most improvements in the deblurred image are to the letters on the wings of the orbiter and the upper regions of the solid fuel rockets.

IV Future Research

The research described herein revealed several interesting avenues of future work. Two prominent areas of importance in designing the deblurring filters are the identification of the appropriate M.T.F. from Fourier transforms of the blurred images and noise abatement. The latter area of work may lead to preprocessing the blurred images, e.g. with windowing functions. Another important investment of research time is that of progressing systematically through a series of blurred images. These images should be classified as much as possible beforehand into groups having similar blurring processes. It is clear that any strategies developed in working with a few images need to be tested and refined by considering other images. In this section each of these areas of work will be reviewed.

V.1 Construction of M.T.F.'s from Images

Each blurring process can be expected to have its own point spread function (P.S.F.) and an associated modulation transfer function (M.T.F.). A priori knowledge of the blurring process is helpful, but not essential. For example, the circular halo observed in the second but not the first image considered in section III is characteristic of improper focussing. The lack of a halo would suggest another type of blurring, e.g. gaussian. Relative motion would have a line segment for a P.S.F. For the simplest case, that of a straight line segment, i.e. constant velocity, the resulting M.T.F. would have rectangular symmetry. A combination of improper focussing and relative motion has a M.T.F. that is the product of the individual M.T.F.'s for each process. Blurring due to atmospheric turbulence under certain conditions would be gaussian and would not result in a halo pattern. In many cases, the only means of identifying the M.T.F. and the parameters needed to describe it will be from the Fourier image. Visual inspection of a Fourier image for the halo pattern probably is the fastest means of identifying the source of blurring. Having appropriate M.T.F. parameters, e.g. the exponent coefficient of a gaussian or the diameter of the disk in the case of improper focussing. How much of this work could be accomplished by a computer remains to be seen.

In principle, it is possible to extract P.S.F. parameters directly from features of the spatial image. In particular, if the image pixels associated with a single point of the sharp image can be identified then one has the P.S.F. In theory such information is also contained in the profiles of edges and lines in the blurred image. The change of pixel values perpendicular to an edge can be shown to be proportional to the pixel values of the P.S.F. measured along the diameter in the same directions.

IV.2 Noise Abatement

The two examples considered in section III have clearly shown that identification of the M.T.F. is only the first step. Care must be taken to avoid emphasis of noise. The approach used there of a one or two parameter Wiener-like suppression of high frequency contributions was a rather crude method of noise abatement. It would be better if the source or the nature of the noise could be identified and removed by a separate frequency filter. For example, in working with the first image, it was found that setting all amplitudes inside circular arcs of radius N/4 around the four corners of the Fourier image had no visual effect on the unblurred image. In the case of spike noise, where individual pixels are radically different in value from the

surrounding pixels, a subroutine could search and remove such "sparkle". Sparkle is most prominent near the corners of the Fourier image where little if any signal is present. It is in contrast to the more cloud-like structure typical of the low frequency region where signal dominates. It should be possible to obtain information on noise by studying select portions of the spatial image, e.g. if a region is known to be uniform in intensity, then any ripple in its Fourier transform may be assumed to be noise. Identifying regions or individual pixels of the Fourier image as noise dominated and their removal would greatly reduce the importance of the choice of parameters in the deblurring filter.

IV.3 Image Truncation and Windowing

In dealing with the image blurred by improper focussing, ringing supposedly due to spatial truncation was observed. In this particular case, the upper sixteenth of the image was not present due to the digitalization process. Because of the cut-off of the camera, it is barely noticeable in figures 2a, b or c. However, an attempt to fill in the missing area removed strong vertical lines near l=0 and created the scratchy band about 32 pixels wide, i.e. N/16. This is an indication of the importance of spatial truncation on the Fourier image. It would be worthwhile to see what effect the edges of an image have on its Fourier image. There are many different types of window functions(1) that remove the effects of the edges and could help us to see which Fourier features are important and which are spurious.

V Summary and Conclusion

The two examples discussed in section II, demonstrated the significant improvement that can be accomplished in enhancing blurred images. Information needed to construct the frequency filters which manipulate the Fourier images can be extracted from the Fourier as well as the spatial image. For the two examples considered, the nature of the blurring mechanism was evident from the presence or absence of halos in their Fourier transforms. It was also found that once the type of blurring was recognized if was easier to use the Fourier images to fix the appropriate parameters. Supposedly methods could be developed to extract the needed information from features of the spatial images that might complement the work with Fourier images. Similarly, noise reduction could proceed from identification of spike noise etc. in the Fourier image or from studying select regions of the spatial image.

In conclusion, it has been demonstrated that blurred images can be enhanced using information extracted from the images themselves. However, there is plenty of research and art work to be done before JSC has a cookbook for deblurring images.

VI References

1 Moik, J.G.,Digital Processing of Remotely Sensed Images, NASA SP-431, 1980

2 Castleman, K.R., Digital Image Processing, ISBN 0-13-21365-7 Prentice-Hall Inc. 1979

3 Gonzalez, Wintz, Digital Image Processing

4 Press,W.H., B.P.Flannery, S.A. Teukolsky,W.T. Vetterling, Numerical Recipes, ISBN 0-521-30811-9, Camgridge Univ. Press, 1986

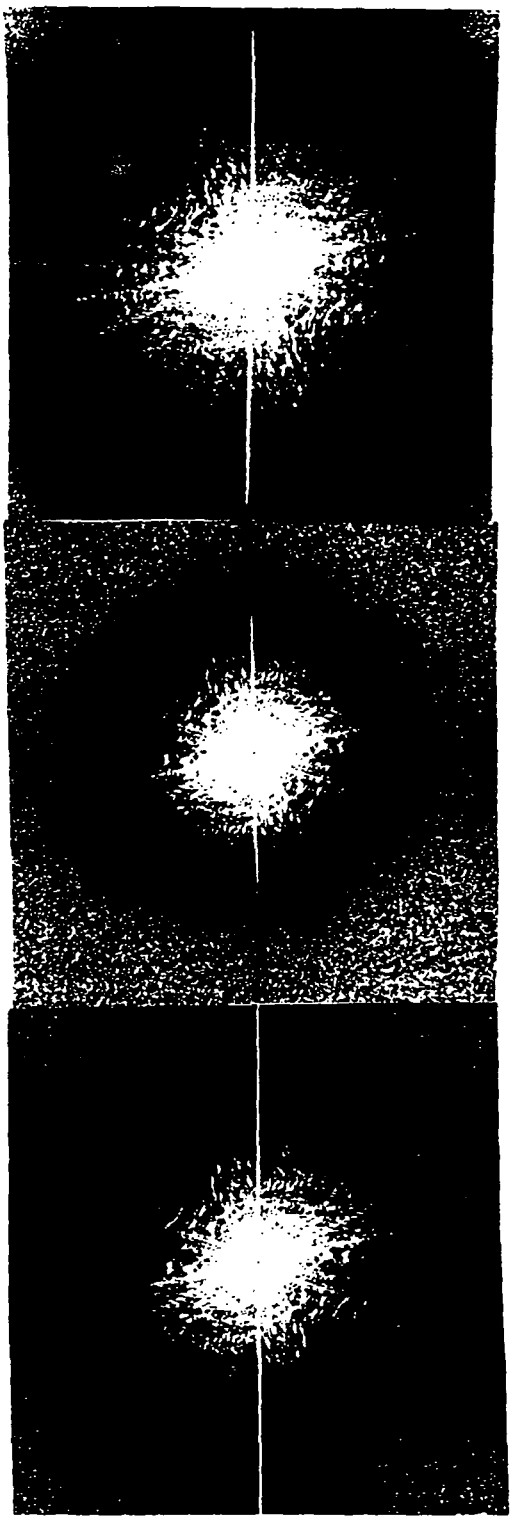
5 Reynolds, G.O., J.B. DeVelis, G.B Parrent Jr.,B.J. THompson, Physical Optics Notebook: Tutorial in Fourier Optics, ISBN 0-8194-0130-7, 1989

6 Menke, W., Geophysical Data Analysis: Discrete Inverse Theory", ISBN 0-12-490921-3 , AP, 1989

7 Ward,R.K., B.E.A. Saleh, Restoration of images Distorted by System of Random Impulse Resonses, J. Opt. Soc. Am.,A2,pp1254-1259,1985

8 Hufnagel, R.E., N.R. Stanley, Modulation Transfer Function Associated with Image Transmission through Turbulent Media, J. Opt. Soc. Am, vol.54, pp 52-61, 1964

Figures 1a, 1b and 1c



Figures 2a, 2b and 2c

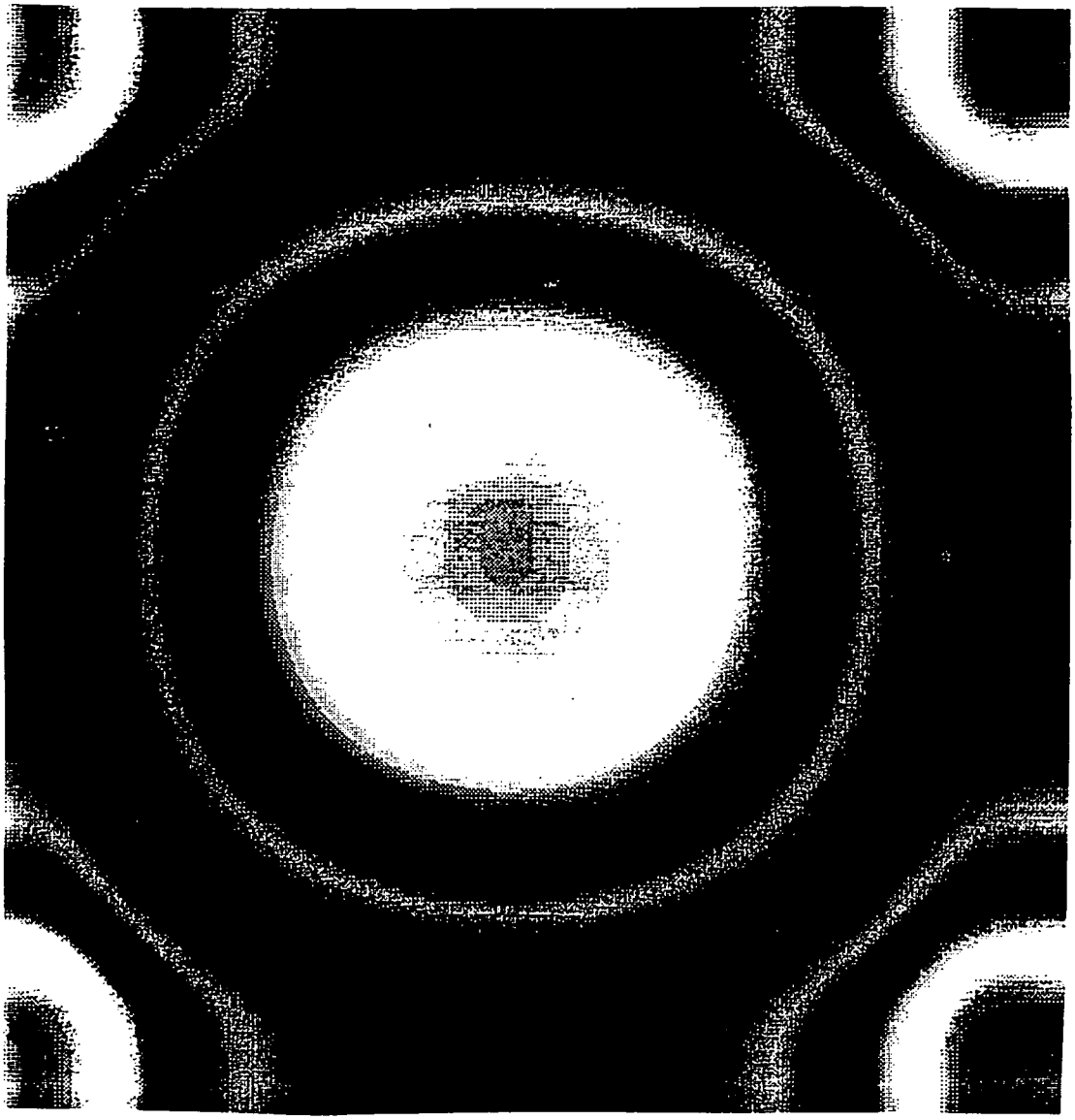
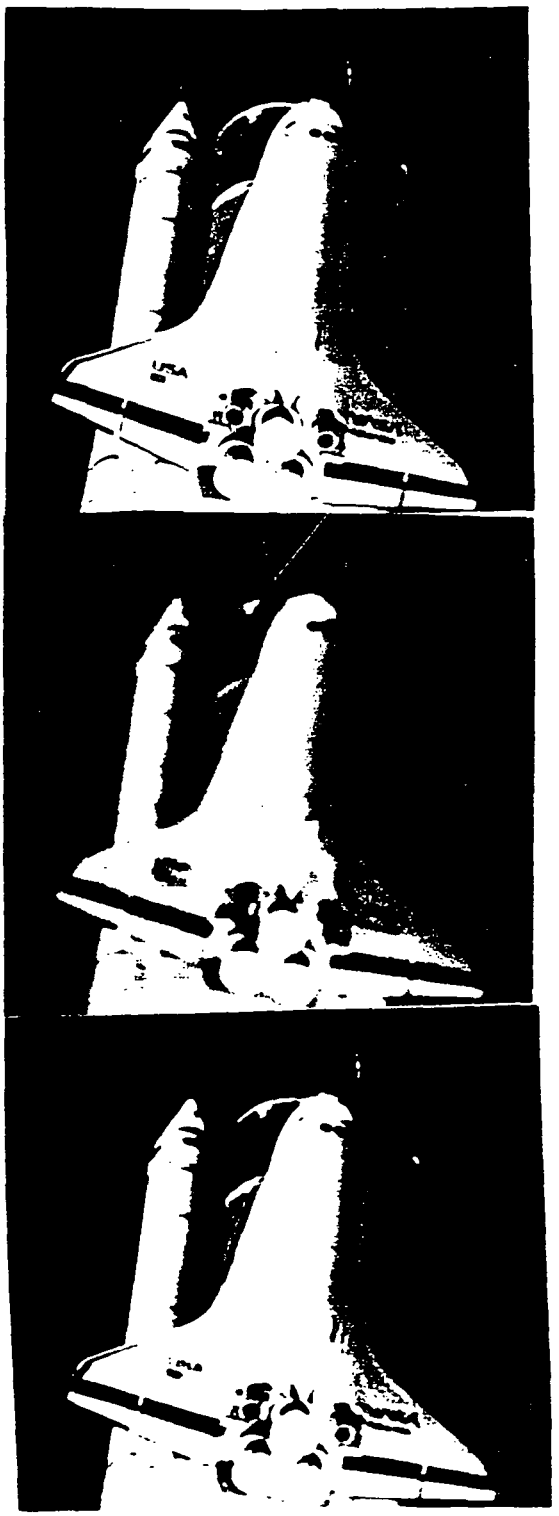
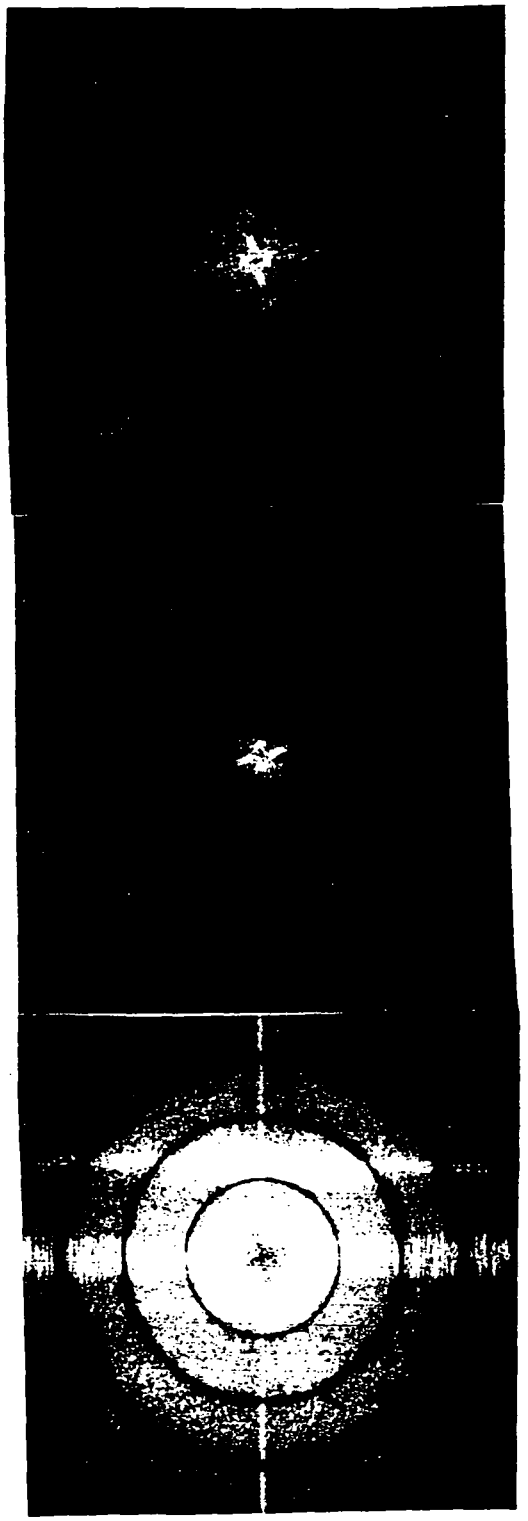


Figure 3



Figures 4a, 4b and 4c



Figures 5a, 5b and 5c

NONLINEAR SYSTEMS DYNAMICS IN CARDIOVASCULAR PHYSIOLOGY: THE HEART RATE DELAY MAP AND LOWER BODY NEGATIVE PRESSURE

Final Report

NASA/ASEE Summer Faculty Fellowship Program 1990

Johnson Space Center

Prepared By:	John C. Hooker, Ph.D.
Academic Rank:	Professor
University & Department:	Houston Baptist University Department of Biology Houston Baptist University
NASA/JSC	
Directorate:	Life Sciences
Division:	Medical Sciences
Branch:	Space Biomedical Research Institute
JSC Colleague:	John B. Charles, Ph.D.
Date Submitted:	August 17, 1990
Contract Number:	NGT-44-005-803

ABSTRACT

This report is a presentation of a preliminary study of the applicability of nonlinear dynamic systems analysis techniques to LBNP studies. In particular, the applicability of the heart rate delay map is investigated.

It is suggested that the heart rate delay map has potential as a supplemental tool in the assessment of subject performance in LBNP test and possible in the determination of susceptibility to cardiovascular deconditioning with spaceflight.

INTRODUCTION

The human cardiovascular system is a very complicated, highly nonlinear, dynamic system. As such, it is a primary candidate for analysis using the techniques of chaos theory which have evolved in the recent past in an attempt to understand the nature of deterministic systems whose outputs appear to be ramdom.

One tool which has been used for discrete systems -- systems with outputs only at discrete times and described, at best, by difference equations, as opposed to systems with continuous time outputs and described by differential equations -- is the delay map. The delay map for discrete systems is analogous to the phase space trajectory for continuous systems. Each of these represents a description of the dynamic behavior of the variable under consideration.

This report discusses some of the findings of a preliminary study into the applicability of the delay map in the analysis of heart rate varation during lower body negative pressure (LBNP) tests. This study was performed in the Cardiovascular Laboratory, Space Biomedical Research Institute, NASA Johnson Space Center, as part of the NASA/ASEE Summer Faculty Fellowship Program 1990.

DATA ACQUISITION AND METHODOLOGY

The heart rate data used in this study was derived from data recorded during pre-bedrest, pre-syncopal LBNP tests. Actual data recorded included ECG (in most cases), blood pressures (systolic, diastolic, and mean), and chamber pressures (LBNP). Cardiac cycle duration (CCD) was calculated from the ECG and the blood pressure. Where available, the CCD from the ECG was used to determine the heart rate as a function of time. If CCD from the ECG was not available the CCD from the blood pressure was used. Heart rate was determined as the inverse of the CCD multiplied by 60 seconds per minute. The time sequence for the heart rate profile was determined by summing the CCDs. The recorded data was digitized and presented in worksheet format on floppy discs. Data on four subjects was available.

The heart rate delay map was constructed by comparing the heart rate after a twelve second delay with the heart rate before the delay, i.e., the present heart rate versus the heart rate twelve seconds ago. A simple program was written in the C language to search back through the heart rate data to determine the heart rate twelve seconds ago.

The heart rate delay map was then constructed by plotting the heart rate after the delay versus the heart rate before the delay.

RESULTS PRESENTATION AND DISCUSSION

The heart rate delay of twelve seconds represents the delay associated with autonomic nervous system control of the cardiovascular system primarily via the sympathetic and parasympathetic inputs to the SA node pacemaker of the heart. Several researchers (1,2) have shown through spectral analyses of heart rate data that there is a low frequency component of the spectrum at approximately 0.08 Hz cooresponding to this delay. Based on these analyses the twelve second delay was used for this study.

The heart rate time profile of Subject A is shown in Figure 1. Superimposed on that profile is the LBNP. Each chamber pressure was maintained for approximately three minutes with the transitions taking only a few seconds. The LBNP test was terminated when the systolic blood pressure dropped below a given threshold and the heart rate ceased to increase or decreased. For Subject A the test was terminated after a few minutes at a LBNP of -50 mmHg.

Figures 2 through 6 present the heart rate delay map segmented according to the steady LBNPs for Subject A and Figure 7 shows the composite heart rate delay for Subject A. The baseline or atmospheric pressure case (Figure 2) indicates that the heart rate delay map is concentrated in a small region of the map. If the -20 mmHg condition (Figure 3) is also considered as a non-stressful situation and only an extension of normal, the heart rate delay map resembles what Goldberger (1), and others in the chaos liturature, refer to as a stranger attractor. This attractor pattern indicates that, although the heart rate is quite variable, after excursion away from the norm, the heart rate returns. Note the "butterfly pattern" of Figure 3.

Figures 4, 5 and 6 indicate that as the LBNP stress increases (lower LBNP) the heart rate increases, becomes more variable, and the excursions increase in magnitude and duration. Of significance is the decrease in variability of the map -- a concentrating of the map -- during the late -50 mmHg case and a decrease in the heart rate -- seen as the map trends down and to the right as the end of the test approaches.

One other subject presented a similar delay map.

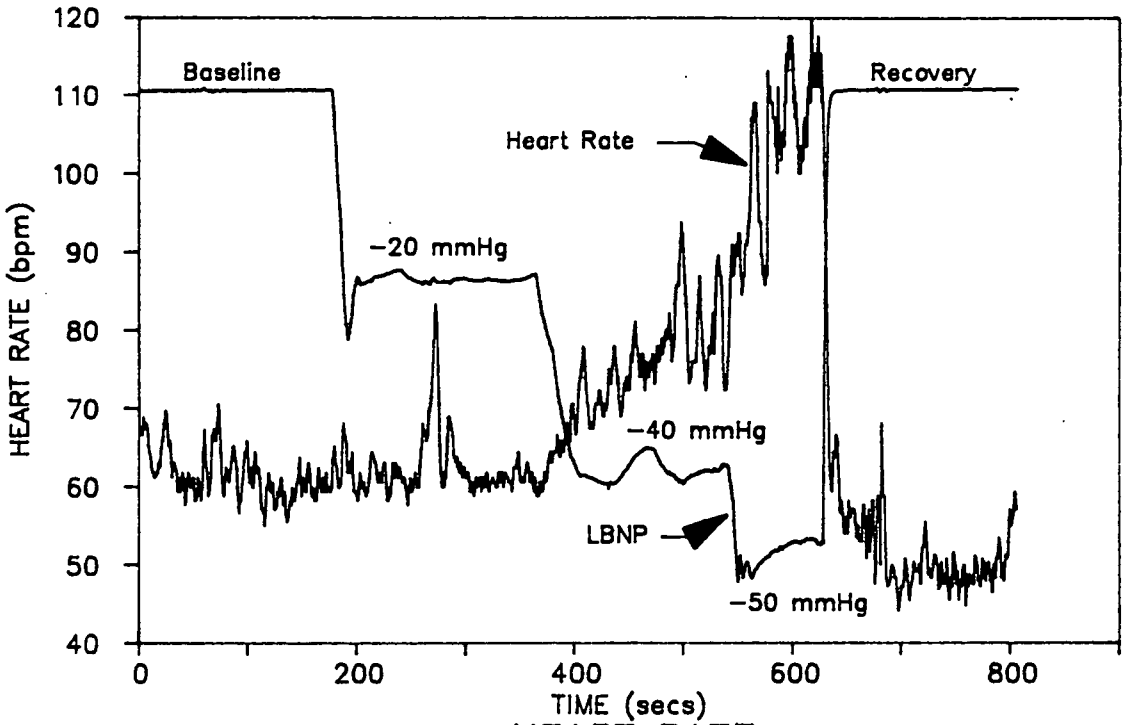


Figure 1.- HEART RATE
SUBJ A, PRE-BEDREST, PRE-SYNCOPAL TEST

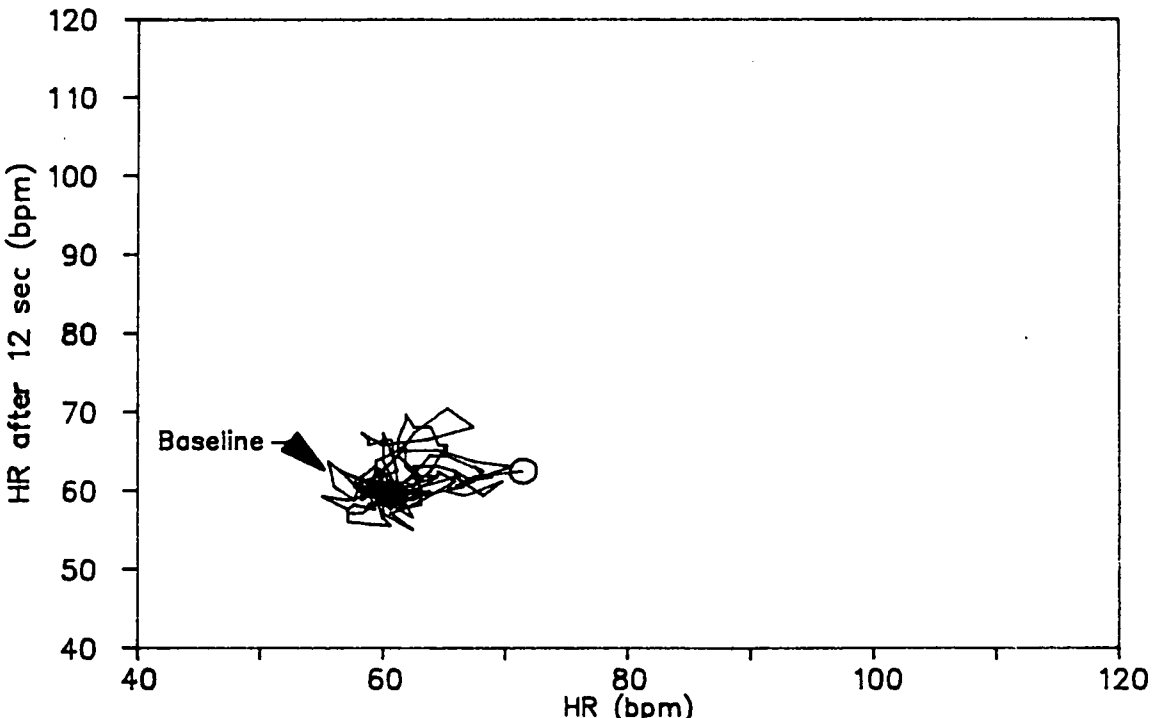


Figure 2.- HEART RATE DELAY MAP
SUBJ A, PRE-BEDREST, PRE-SYNCOPAL TEST

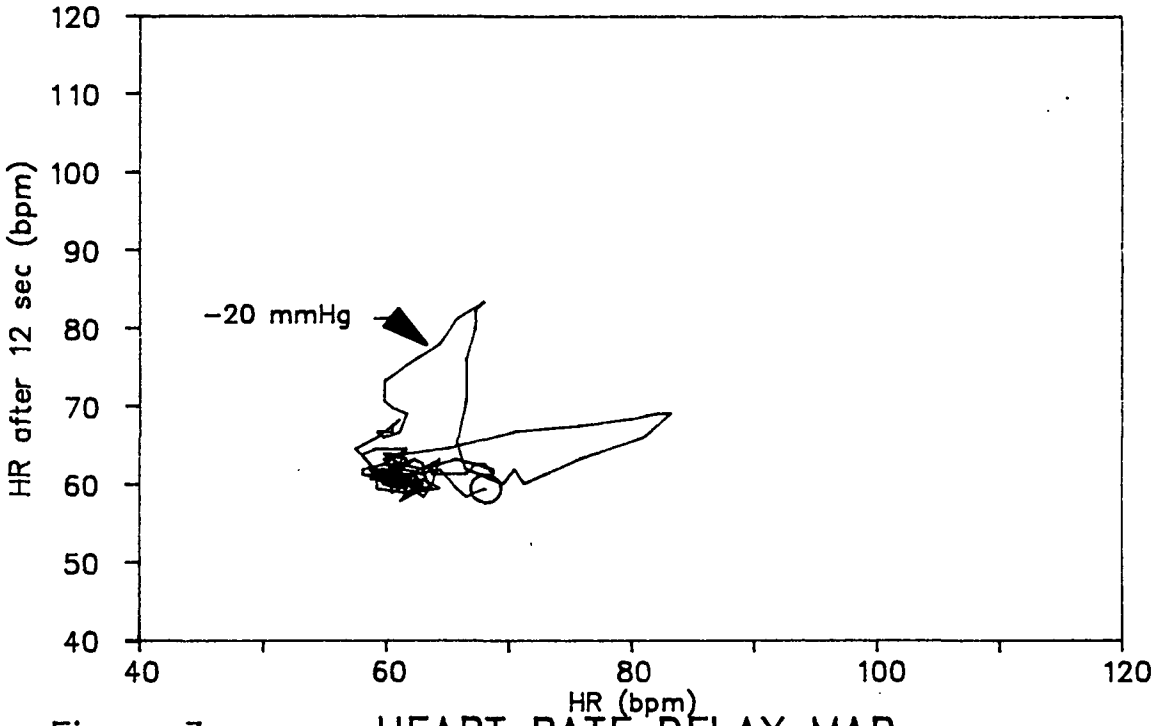


Figure 3.— HEART RATE DELAY MAP
SUBJ A, PRE−BEDREST, PRE−SYNCOPAL TEST

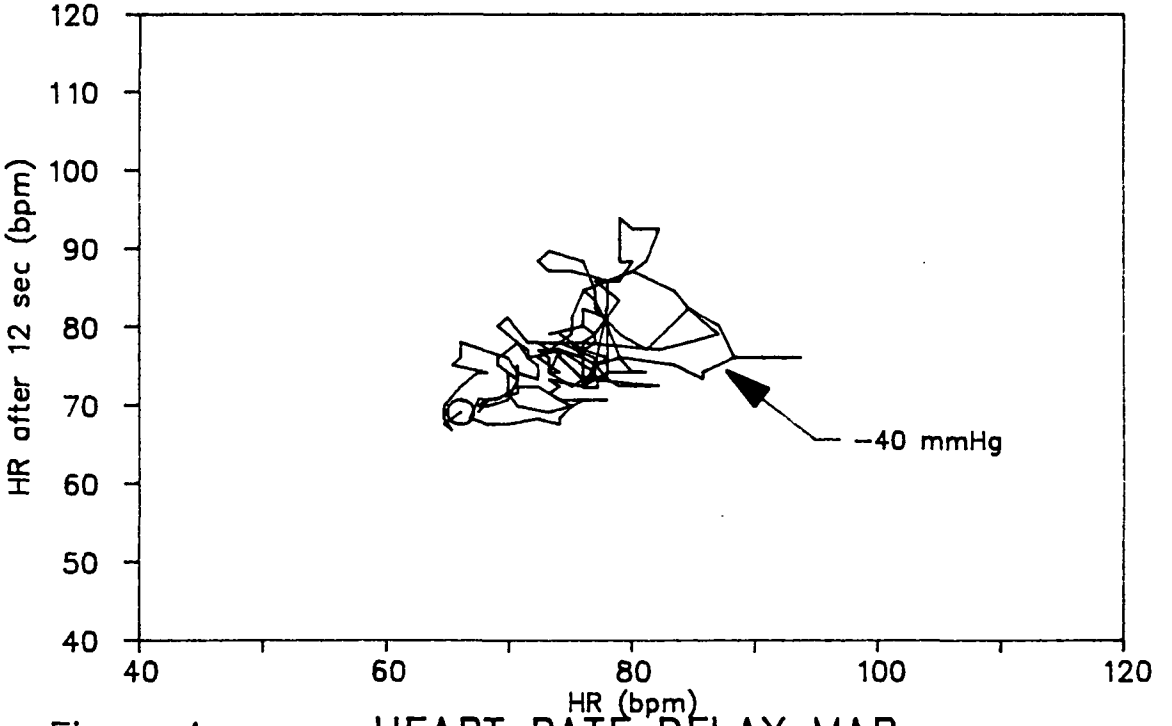


Figure 4.— HEART RATE DELAY MAP
SUBJ A, PRE−BEDREST, PRE−SYNCOPAL TEST

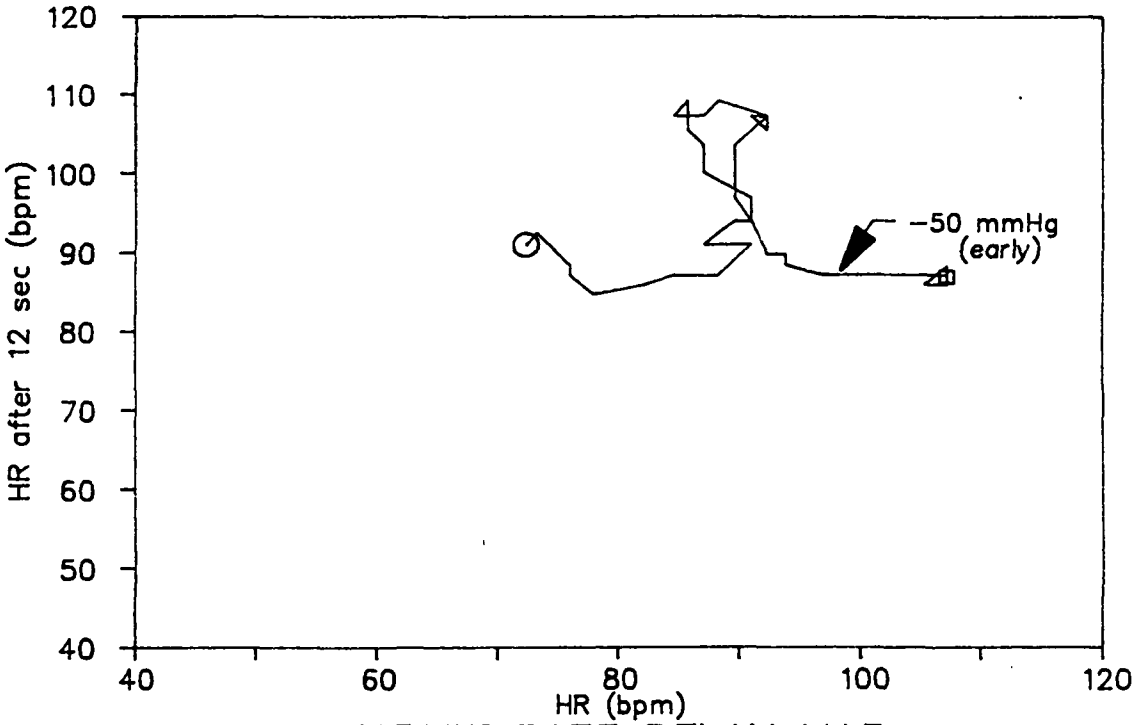


Figure 5.– HEART RATE DELAY MAP
SUBJ A, PRE–BEDREST, PRE–SYNCOPAL TEST

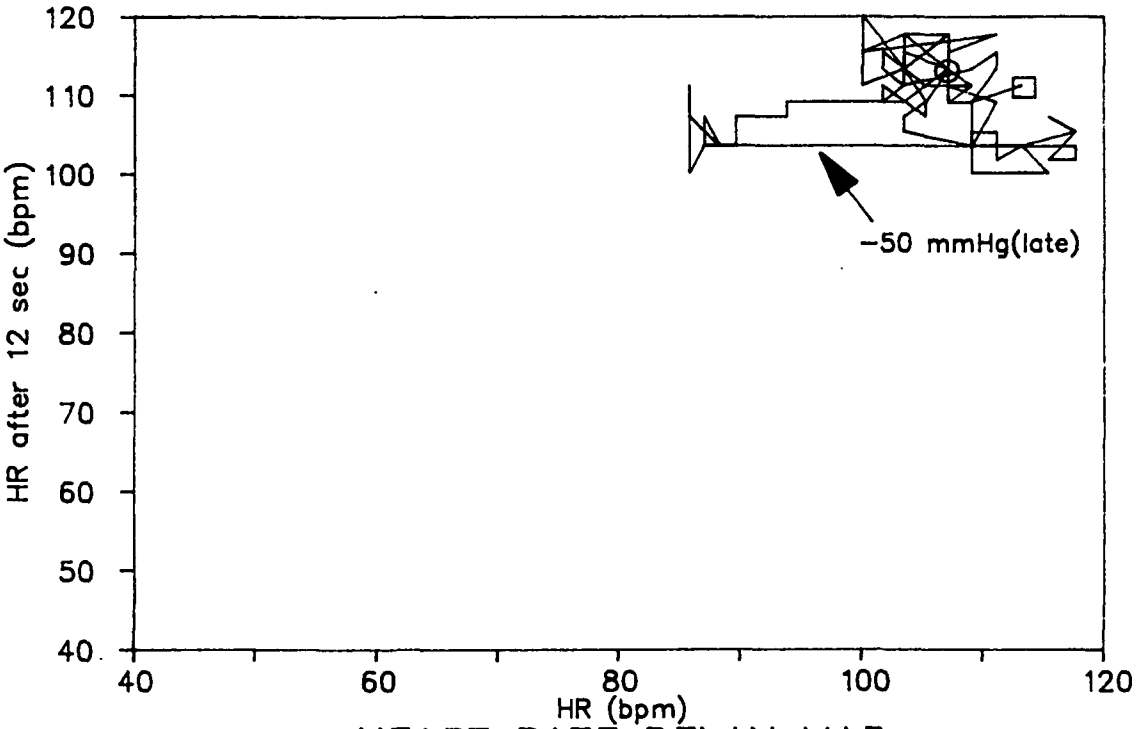


Figure 6.– HEART RATE DELAY MAP
SUBJ A, PRE–BEDREST, PRE–SYNCOPAL TEST

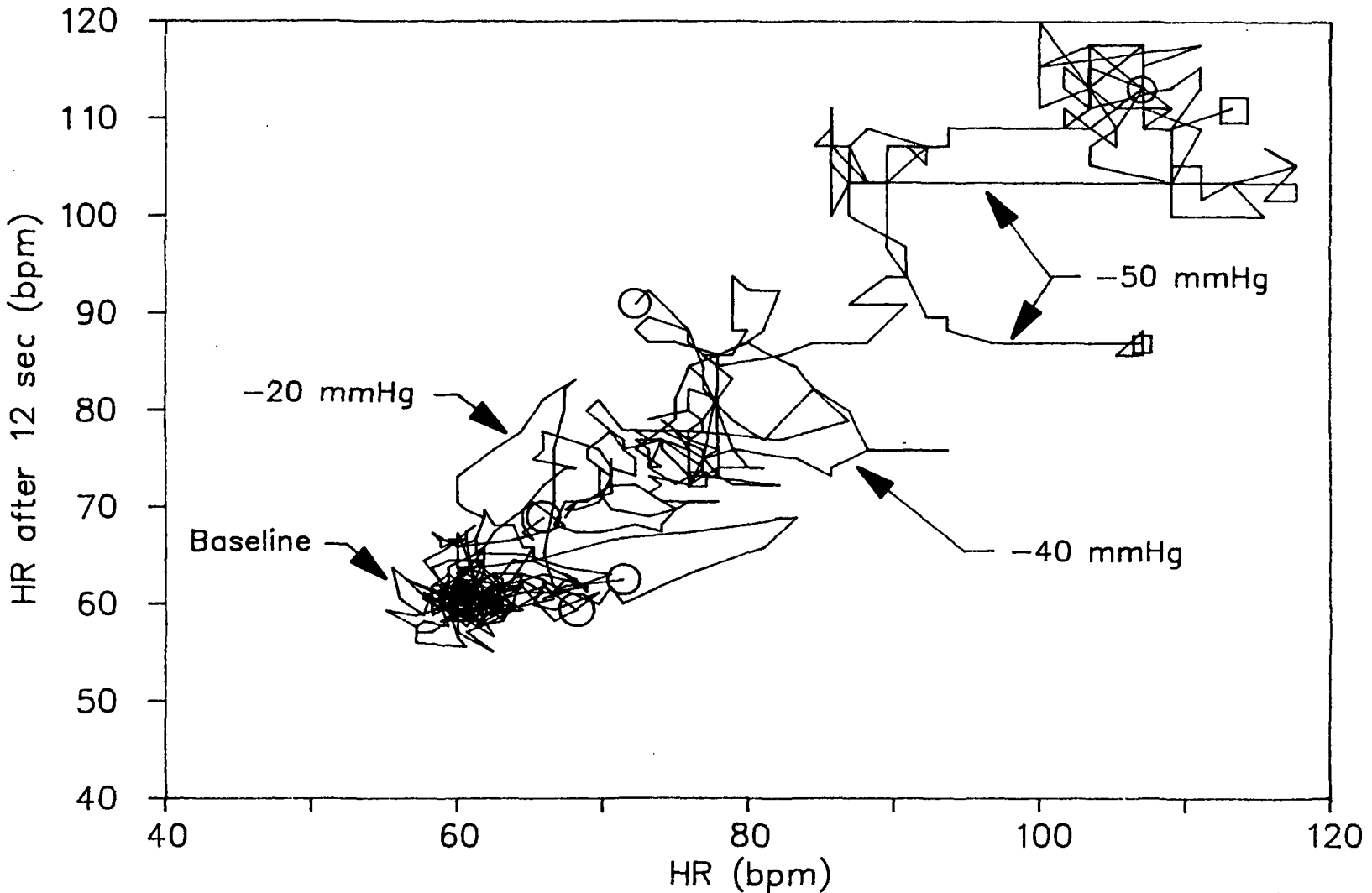


Figure 7.— COMPOSITE HEART RATE DELAY MAP
SUBJ A, PRE−BEDREST, PRE−SYNCOPAL TEST

Subject B (Figures 8 and 9) was able to withstand a greater LBNP stress as shown by LBNPs of -60 to -80 mmHg. Of significance in this study is the contrast to Subject A as seen in the decrease in heart rate variability (Figure 8) and a narrowing of the delay map (Figure 9) with LBNP stress. The geometric, rectangular pattern seen in the upper regions of heart rate (Figure 9) is a function of only two significant figures in the CCD.

Another subject presented a similar heart rate delay map to Subject B. It would appear from these two cases that as the subject is able to move to higher LBNP stress levels the heart rate delay map narrows.

CONCLUSIONS AND RECOMMENDATIONS

Although this was only a prelimimary, cursory and small sample study, it appears that the heart rate delay map has potential as an analytic tool in LBNP studies. It would appear that the delay map has use as a supplemental aid in ascertaining the termination point of an LBNP test, particularly, if an on-line, real-time presentation is implemented. The delay map may also have potential as a predictor of subject susceptibility to cardiovascular deconditioning in spaceflight.

The following recommemdations are made for further studies:

1) Longer time profiles of heart rate at constant LBNP should be analyzed. Such studies should gave better indicators of strange attractors and deviations from them and the signicance of such deviations.

2) The longer time profiles should also be subjected to spectral analyses and correlations done between changes in the heart rate delay map and the frequency spectrum.

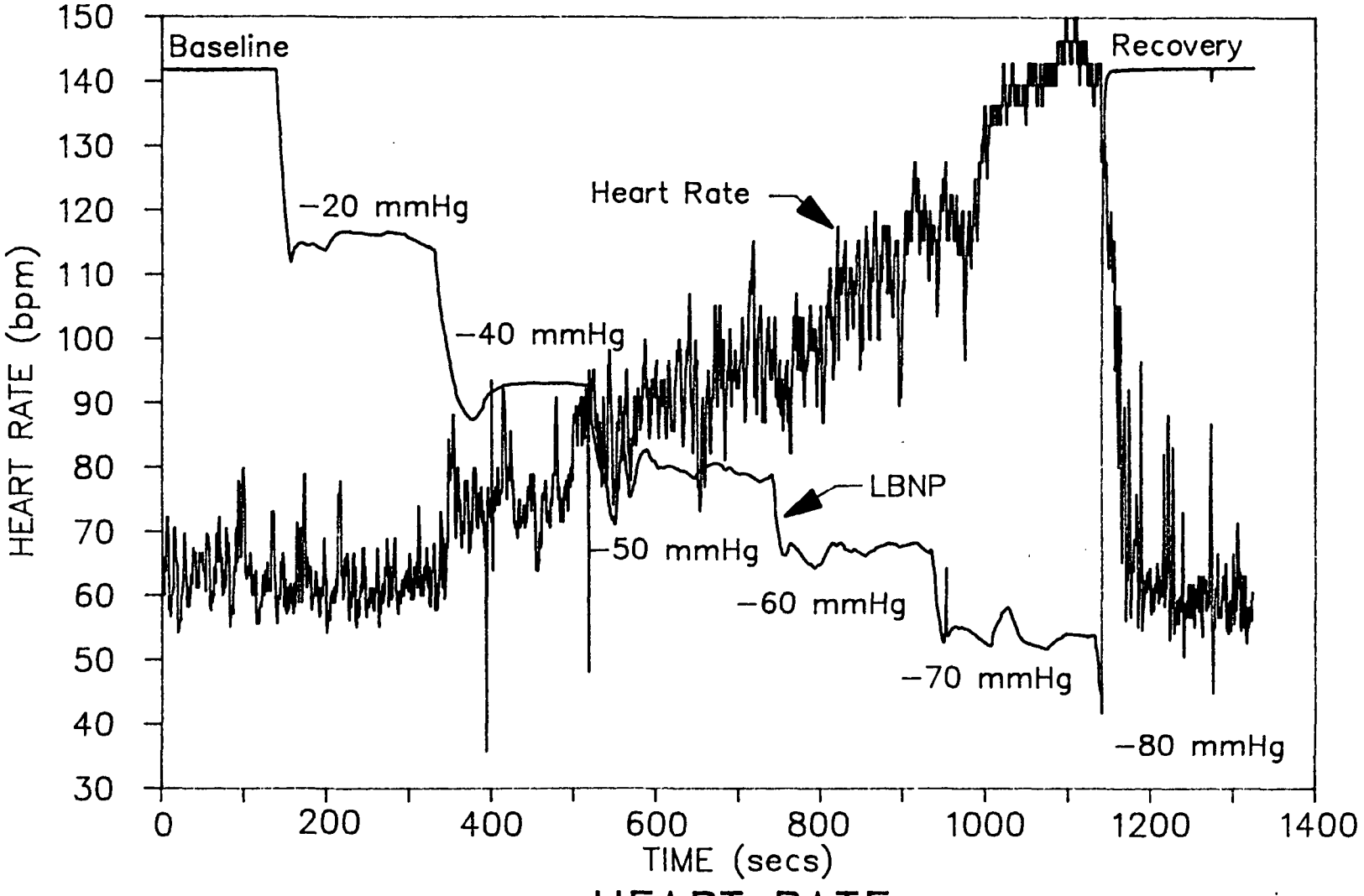


Figure 8.— HEART RATE

SUBJ B, PRE−BEDREST, PRE−SYNCOPAL TEST

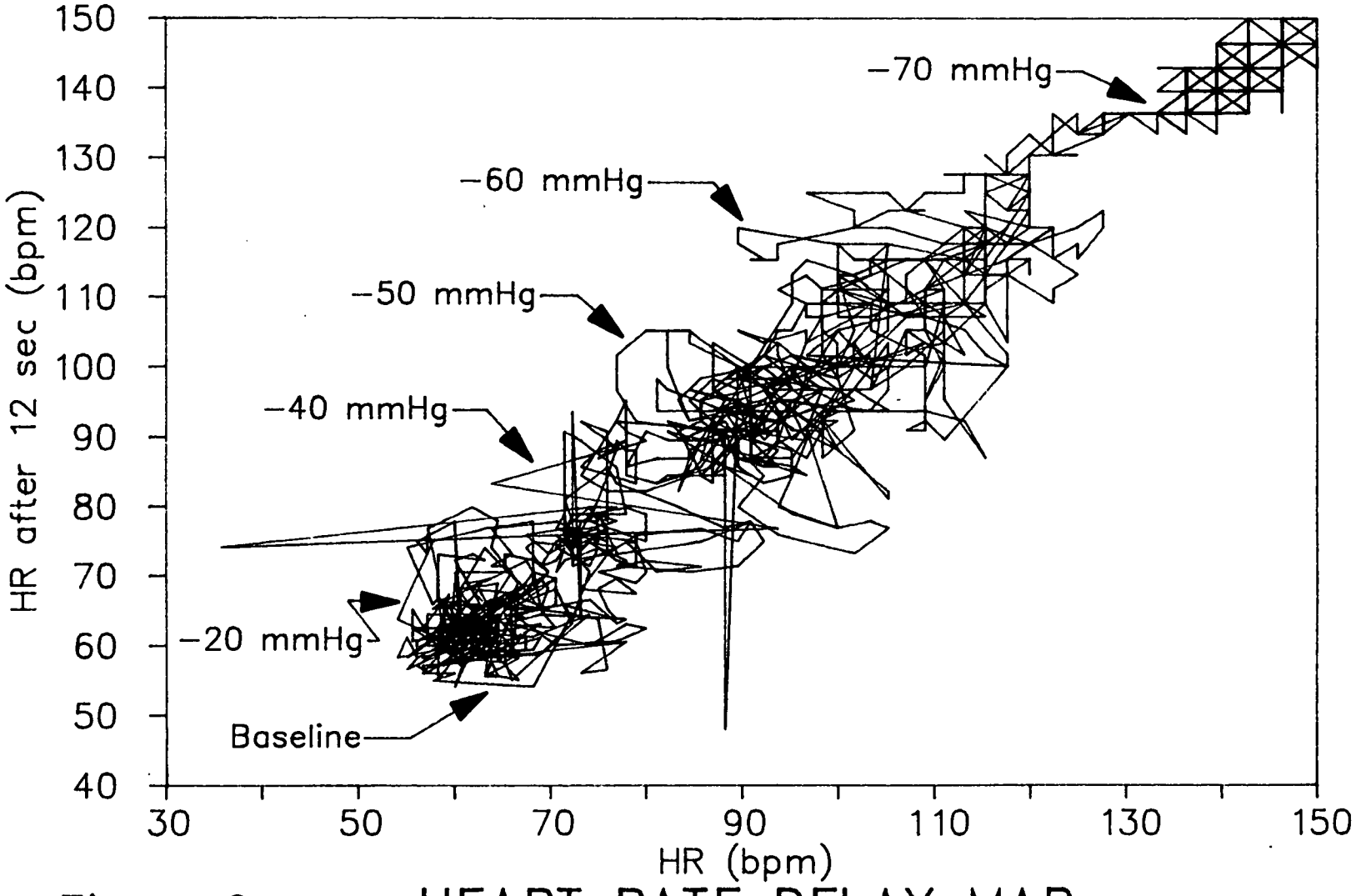


Figure 9.— HEART RATE DELAY MAP
SUBJ B, PRE−BEDREST, PRE−SYNCOPAL TEST

REFERENCES

1. Goldberger, A. L., Rigney,D. R., and West, B.J., Chaos and Fractals in Human Physiology, Scientific American, Feburary, 1990.

2. Akselrod, S., Gordon, D., Ubel, F., Shannon, D., Barger, A., and Cohen, R., Power Spectrum Analysis of Heart Rate Fluctuation: A Quantitative Probe of Beat-to-Beat Cardiovascular Control, Science, vol. 213, July, 1981.

INCORPORATION OF SHUTTLE CCT PARAMETERS IN COMPUTER SIMULATION MODELS

Final Report

NASA/ASEE Summer Faculty Fellowship Program - 1990

Johnson Space Center

Prepared By:	Terry Huntsberger, Ph.D.
Academic Rank:	Associate Professor
University & Department:	University of South Carolina Department of Computer Science Columbia, South Carolina 29208
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Man-Systems
Branch:	PLAID Laboratory
JSC Colleague:	Linda Orr
Date Submitted:	August 10, 1990
Contract Number:	NGT-44-005-803

Abstract

Computer simulations of shuttle missions have become increasingly important during recent years. The complexity of mission planning for satellite launch and repair operations which usually involve EVA has led to the need for accurate visibility and access studies. The PLAID modeling package used in the Man-Systems Division at Johnson currently has the necessary capabilities for such studies. In addition, the modeling package is used for spatial location and orientation of shuttle components for film overlay studies such as the current investigation of the hydrogen leaks found in the shuttle fleet.

However, there are a number of differences between the simulation studies and actual mission viewing. These include image blur caused by the finite resolution of the CCT monitors in the shuttle and signal noise from the video tubes of the cameras. During the course of this investigation the shuttle CCT camera and monitor parameters are incorporated into the existing PLAID framework. These parameters are specific for certain camera/lens combinations and the SNR characteristics of these combinations are included in the noise models. The monitor resolution is incorporated using a Gaussian spread function such as that found in the screen phosphors in the shuttle monitors.

Another difference between the traditional PLAID generated images and actual mission viewing lies in the lack of shadows and reflections of light from surfaces. Ray tracing of the scene explicitly includes the lighting and material characteristics of surfaces. The results of some preliminary studies using ray tracing techniques for the image generation process combined with the camera and monitor effects are also reported.

INTRODUCTION

Geometric properties such as visibility from a given vantage point and clearance for movement are adequately modeled with the PLAID package currently used in the Man-Systems Division at the Johnson Space Center. This modeling package uses engineering drawings and sketches to build a full three-dimensional database of objects based on polygons. This database can then be used to display a filled polygon view from any viewpoint. Clarity of the image is only limited by the spatial resolution of the display device, which is typically 1024 rows by 1024 columns. In an actual mission scenario, the display device resolution is lower and camera effects such as noise are inherent in the imaging process. This study investigates the correction of the computer generated images for degradation during the imaging process.

Surface material properties and lighting characteristics also have an effect on the image generation process. There is a dynamic relationship between the camera iris settings and the relative brightness of surfaces in a scene. Strong variations in lighting can be caused by strong reflective surfaces such as satellite shrouds, as well as shadows due to payload occlusion of light sources. Incorporation of these properties into the PLAID database structure and image generation process is another major component of this study. Preliminary results from some shuttle scenes are also reported.

IMAGING SYSTEM EFFECTS

Finite monitor resolution and camera noise characteristics can be modeled using parameters specified by the manufacturers. The blurring caused by the lower display device resolution can be modeled as a point spread function. In other words, a point in the actual scene is displayed as a blurred circle on the monitor. This process of blurring is illustrated in Figure 1, where Figure 1(a) shows the original intensity of the point in the scene as a narrow peak and Figure 1(b) shows the same point after spreading on the monitor. The classical point spread function derived using the

Fourier transform would be a sinc. This is an expensive function to apply to an entire image and the same visual appearance can be obtained using a Gaussian or triangle point spread function. The original image is resampled using the horizontal and vertical resolution supplied by the manufacturer for the display monitor.

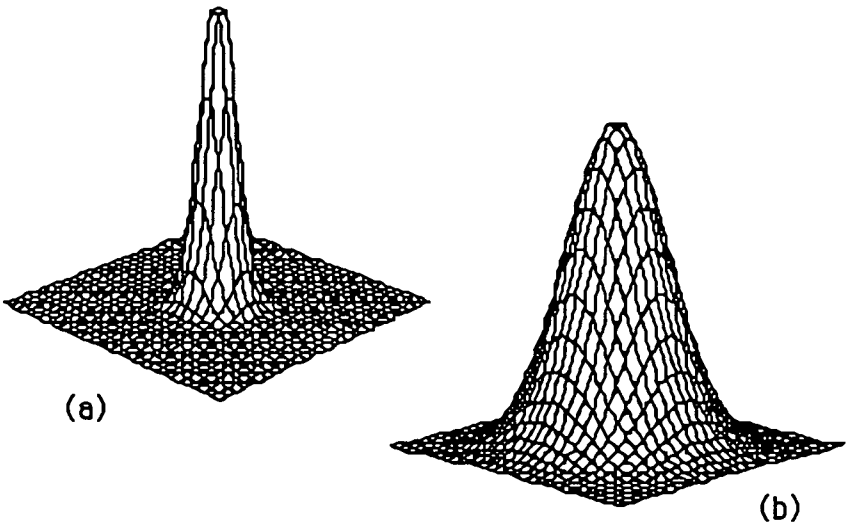


Figure 1. Point spread process caused by monitor phosphors

System noise is the other source of degradation in the imaging process. Noise manifests itself in the camera videotube and in the connecting electronics and wires. Since it is extremely difficult to measure the noise characteristics of the connecting wires, only the videotube noise will be included in the correction process. If the tube is in good calibration, noise will appear randomly throughout the field of view. The amount of noise will depend on the settings of the camera iris, since an iris setting that is narrow in a low light setting will lead to a much greater degradation of the image. The auto-iris setting of the camera causes this to happen most frequently when there are extremely bright objects such as shroud covered satellites in the scene. The iris will close down in this situation with the result that areas in low light are extremely noisy.

Apart from iris effects, a Gaussian white noise process can be used to model the degradation caused by the camera system. The intensity of points in the image as sensed by the camera system are randomly modified by sampling from a uniform Gaussian

distribution, whose width can be derived from the signal to noise ratio (SNR) for the camera system. Once again this is information that is supplied by the manufacturers of the system. This process is equivalent to adding or subtracting a percentage of the intensity of the actual image point for point in a random manner. Generally the noise level of the camera system is plus or minus ten percent. The next section discusses extensions to the monitor and camera degradation processes.

SCENE EFFECTS

There are a number of possible extensions to the present model. First, depth of field effects appear as blurring, dependent on relative distance of objects in the scene from the camera. This blurring must be part of the image generation process and can not be added with post-processing. Second, the diffuse and specular components of surface materials will have a direct impact on the overall appearance of the scene. Finally, the effects of limited lighting conditions or directional light will cause shadows in the scene, which can obscure objects of interest. Two methods can be used to add these extensions to the present model. These are ray tracing and radiosity.

Ray tracing is a technique that recursively follows a light ray from the observers' viewpoint throughout the scene to any light source. This process traces the light ray path as it is reflected off of surfaces or transmitted through transparent surfaces such as glass. A ray is spawned for each picture element or pixel in the final image, which is typically on the order of 1024 rows by 1024 columns. Since there are typically thousands of objects in a scene, this is a computationally intensive task. As such, with the present generation of computer hardware, it is not a real-time image generation process such as that used in the present PLAID system. However, single images can be generated using the present computer hardware in about nine hours.

Radiosity is a technique that is based on thermal heat transfer methods applied to light transmission. This method calculates the light transferred between every pair of surfaces in the scene. As mentioned previously, this is typically on the order of thousands of surfaces. A set of simultaneous linear equations which is equal to

the number of surfaces is set up and solved using any of the common techniques. This process is then followed by an image generation step such as ray tracing. What is gained in the radiosity step before ray tracing is the inclusion of the diffuse component of scattered light from light sources, such as that cast from the Earth into the shuttle bay.

Ray tracing was the method that was decided upon for the preliminary investigations in this report. The shading model that was used for the rendering portion of the ray tracing process was that of Phong. This model includes ambient, diffuse and specular components of the reflection process. It can be expressed as:

$$I = I_a k_a + \frac{I_0}{k+r}[k_d(\vec{N}\cdot\vec{L}) + k_s(\vec{R}\cdot\vec{V})^n] \qquad (1)$$

where I_a is the ambient intensity, I_0 is the intensity of the light source, k_a, k_d and k_s are the ambient, diffuse and specular coefficients, and n is the specular power factor. The various vectors depend on the geometry of the surface, viewer and light source as shown in Figure 2.

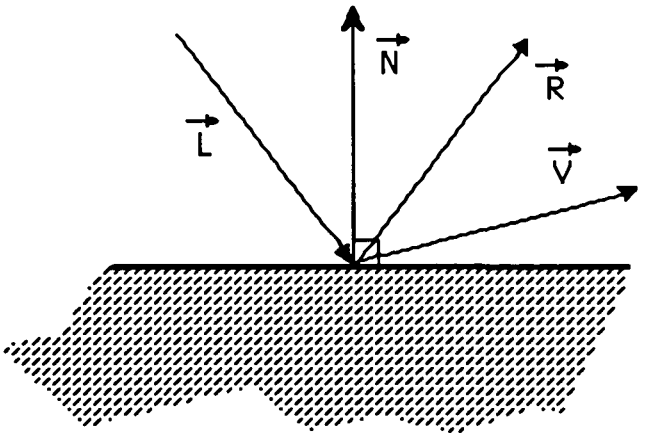


Figure 2. Scene geometry for surface shading.

Specification of surface characteristics relies on knowledge of k_d, k_s and n. These factors can be derived from experimental data collected in the NASA laboratories.

An example of this data is shown in Figure 3, taken from JSC Internal Note 85-SP-1, where orbiter tile reflectance characteristics are plotted vs. degrees off specular angle.

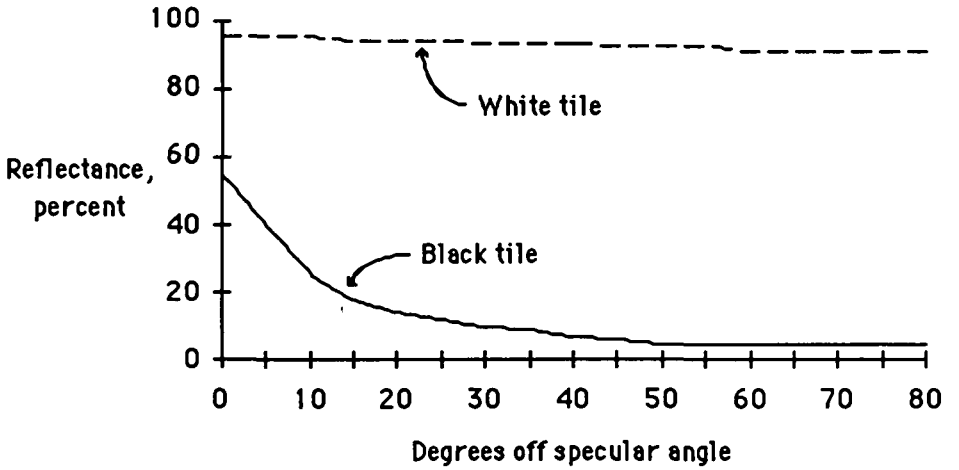


Figure 3. Reflectance vs. specular angle (From C.D. Wheelwright, "Orbiter Solar/Artificial Illumination During STS Flights," JSC Internal Note 85-SP-1).

The reflectance at 80 degrees will be entirely diffuse and this value is k_d. The reflectance at 0 degrees will be a mixture of diffuse and specular, with the specular component being dominant. The dot product term is unity at this angle and the total reflectance is equal to $k_d + k_s$ at this point. Since k_d is already known, k_s can be found directly. With a fixed light source position for the experiment, the only variation expected in the reflectance curve would be due to specular effects. Each experimental point can be then used to find a value of n, the specular power factor, which is then averaged for the value used in the Phong shading formula of Equation 1.

Surface colors and lighting positions are found in NASA documents for the light source intensities and positions in the Phong shading formula. The geometric database needed to derive the surface normals, given by the vector N in Figure 2, is taken directly from the PLAID polygon based description of any structure. For example, the PLAID package can be used to build a description of a specific shuttle mission with the appropriate satellites in the

payload bay. The vector V in Figure 2 is interactively derived using the PLAID package or can be found in the documented camera mount locations for most NASA structures.

EXPERIMENTAL STUDIES

Experimental studies were done on a planned shuttle mission STS-46, which will carry the Tethered Satellite System (TSS) in 1991. Scene geometry was taken directly from the PLAID database for this mission. The image was generated using the Craig Kolb rayshade package from Yale University (available via anonymous ftp at weedeater.math.yale.edu). Parameters for this preliminary test are (1) diffuse surface materials; (2) image resolution of 512 rows by 512 columns; (3) all shuttle bay lights including the forward payload bay bulkhead light on; (4) shuttle in the dark zone of its Earth orbit; and (5) view position of Camera A on the payload bay forward bulkhead. The result of the ray tracing of this scene is shown in Figure 4. Although specular reflections are not present in the scene, most of the visibility related effects are included, with shadows cast along the length of the payload bay being of primary importance.

Figure 4 shows the scene as it would appear to a human observer through the payload bay forward bulkhead window. Incorporation of the camera and monitor effects is easily accommodated using the techniques described in the beginning of this report. Application of the point spread function and the Gaussian white noise distribution to Figure 4 gives the image seen in Figure 5. The point spread function is derived using the published CCTV monitor resolution of 300 rows by 400 columns. Detail on the TSS is lost in this view due to the blurring.

Visibility studies using a different camera for this scene can be done by changing the viewpoint. Switching to Camera B mounted on the aft bulkhead of the shuttle bay should give an indication of what the visibility is in the shadowed region of the payload bay. This view is shown in Figure 6, and it can be seen that there is little improvement in visibility with a change in camera position.

Visibility in a region of shadow can be enhanced if the mission is scheduled during the day portion. Substantial illumination is obtained from reflection of sunlight from the Earth.

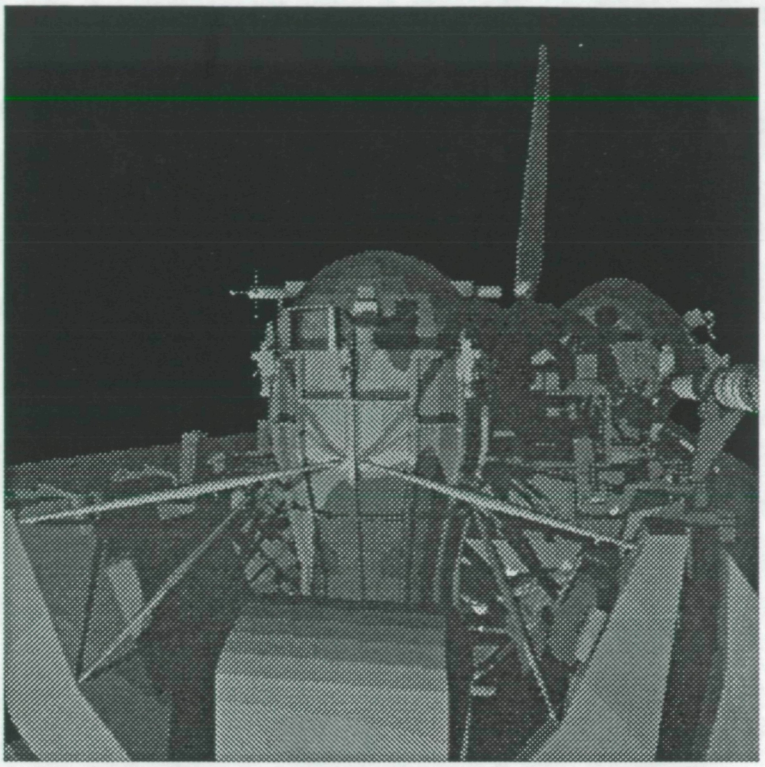


Figure 4. Ray traced image of STS-46 mission

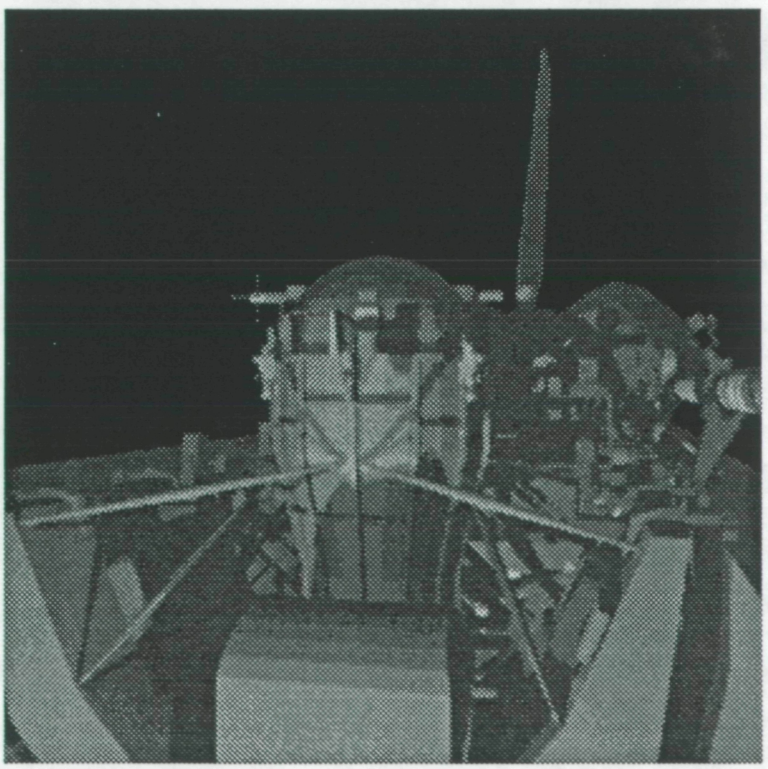


Figure 5. Blurred ray traced image of STS-46 mission

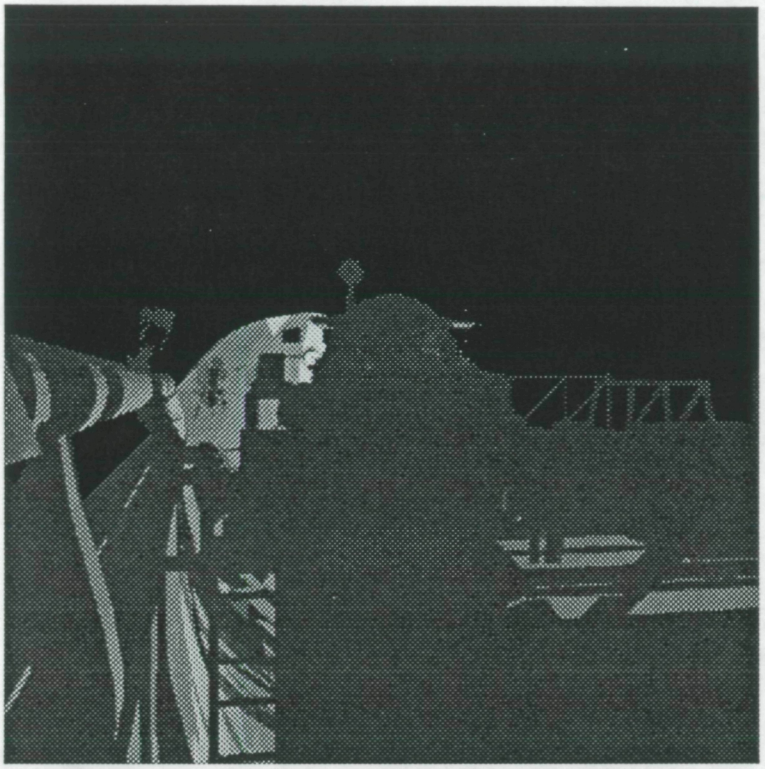


Figure 6. Shadow region ray traced image of STS-46 mission

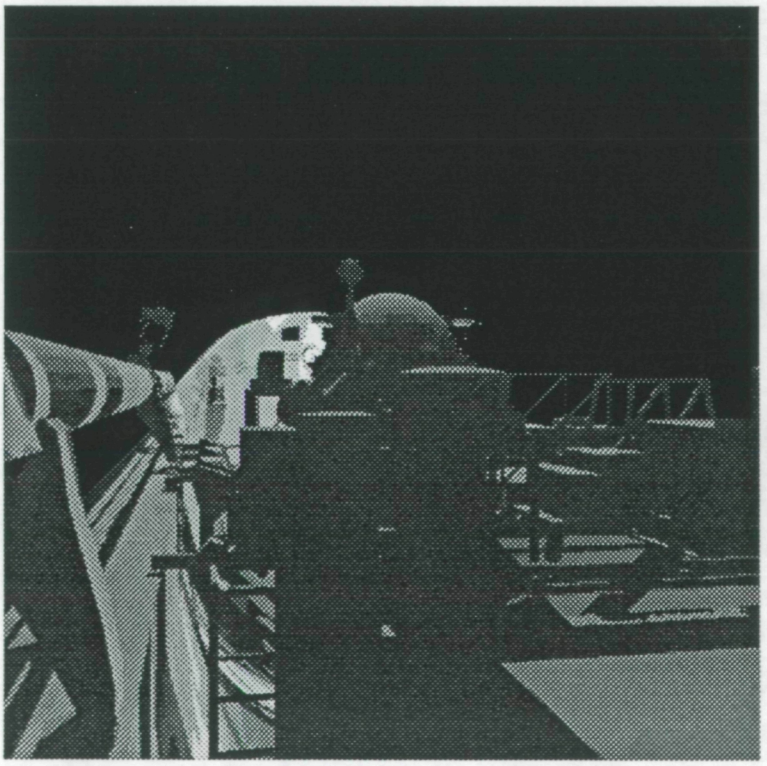


Figure 7. Earthshine model ray traced image of STS-46 mission

To accurately model this effect would require a radiosity treatment of the imaging geometry. An approximation can be found using a light source that is distributed over a large area during the ray tracing image generation process. A ray traced scene with the same viewing parameters as used in Figure 6 but with a distributed light approximation to the Earth is shown in Figure 7. More of the surfaces of the other payloads can be seen. The distributed light approximation gives an indication of the better viewing conditions and can, with increased time needed for the ray tracing, realistically model the actual scene.

DISCUSSION

The PLAID modeling package used in the Man-Systems Division at Johnson maintains an up-to-date geometric database of NASA structures including the shuttle and Space Station Freedom. The work discussed in this report extends the PLAID system to include image degradation processes in the image generation process. The point spread blurring function and noise characteristics were taken directly from the shuttle CCTV monitor and camera documentation, and integrated into a post-processing step for image rendering.

In addition, this report presents some preliminary studies of the use of ray tracing techniques for image generation. The PLAID model database was extended to include surface material characteristics which were obtained from NASA documentation of experimental studies. Placement and strength of light sources, information which is vital for visibility studies, was also obtained from available NASA documentation. An experimental study of shuttle mission STS-46 was done. Figure 4 in the main body of the report shows an unblurred view of the shuttle payload bay for this mission. After the degradation model is applied, detailed feature observation is difficult, as can be seen in Figure 5. Visibility in shadow regions was investigated using the same mission profile. Figures 6 and 7 show the result of this study, where a night view is shown in Figure 6 and a day view is shown in Figure 7. The models used for the Earthshine phenomena are still in development, but an fairly accurate idea of potential visibility problems can be found in these first studies. The images shown in this report have lost

contrast due to the printing process, and can be found in their original form in the videotape *PLAID Graphics Lighting Studies* available as NASA Reference Master 903959 from the Television Office at the Johnson Space Center or from the PLAID Graphics Laboratory in the Man-Systems Division at Johnson Space Center.

Future work will include further extensions of the PLAID modeling framework for faster ray traced previews of scenes. This will be accomplished using a distributed parallel implementation of the ray tracing algorithm. Better modeling of light source geometry will also be studied. With these extensions, the PLAID model database will be a suitable candidate for simulations of visible and range sensor views of scenes in a semi-autonomous robotics environment for NASA development of computer vision algorithms.

TRAINING EFFECTIVENESS OF AN INTELLIGENT TUTORING SYSTEM FOR A PROPULSION CONSOLE TRAINER

Final Report

NASA/ASEE Summer Faculty Fellowship Program - 1990

Johnson Space Center

Prepared by:	Debra Steele Johnson, Ph.D.
Academic Rank:	Assistant Professor
University & Department:	University of Houston Department of Psychology Houston, Texas 77204
NASA/JSC	
Directorate:	Information Systems
Division:	
Branch:	Software Technology
JSC Colleague:	Robert T. Savely
Date Submitted:	August 10, 1990
Contract Number:	NGT-44-005-803

ABSTRACT

A formative evaluation was conducted on an Intelligent Tutoring System (ITS) developed for tasks performed on the Propulsion Console. The ITS, which was developed by AFHRL primarily as a research tool, provides training on use of the Manual Select Keyboard (MSK). Three subjects completed three phases of training using the ITS: declarative, speed, and automaticity training. Data was collected on several performance dimensions, including training time, number of trials performed in each training phase, and number of errors. Information was also collected regarding the user interface and content of training. Suggestions for refining the ITS are discussed. Further, future potential uses and limitations of the ITS are discussed. The results provide an initial demonstration of the effectiveness of the Propulsion Console ITS and indicate the potential benefits of this form of training tool for related tasks.

INTRODUCTION

Intelligent Tutoring Systems (ITS's) have been developed for a variety of tasks, ranging from geometry to LISP programming.[1] However, little systematic evaluation has been conducted on these training systems. Additional research is needed to systematically examine the effectiveness of ITS's both during (formative evaluation) and upon completion (summative evaluation) of software development.[2] Conducting a formative evaluation enables the developer to determine whether the tutor is operating as planned and make program modifications as necessary. A summative evaluation is focused on assessing the training effectiveness of a completed ITS.

The ITS under consideration is still under development. Thus, a formative evaluation was conducted to provide information on the functioning of the system and the effects of the ITS on student learning. Although there is not general agreement yet about which specific evaluation methods are preferred and how to implement them, the information collected at least partially addresses internal and external evaluation issues.[2] Internal evaluation addresses how the tutoring systems functions. External evaluation addresses the educational impact of the tutoring system on students. The primary focus of the current project is on external evaluation, although some information on internal evaluation is also provided. In addition, given that a formative evaluation was being conducted, the focus of data collection was more on process measures than outcome measures. Formative evaluations tend to rely more on process measures (e.g., patterns of task activities) rather than outcome measures (e.g., task performance upon completion of training).[3]

BACKGROUND ON PROPULSION CONSOLE ITS

An ITS is under development by AFHRL which simulates use of the Manual Select Keyboard (MSK) on the Propulsion Console used by flight controllers. The purpose of the ITS development project was to develop a tutoring system for a high performance task.[4] A high performance task is one in which the knowledge required is small, but extensive practice is required to proceduralize the set of skills involved. This type of task is often performed in situations in which high risk or expense is involved.[4] Thus, it may be difficult to provide extensive training in the actual work environment. An ITS provides students an opportunity to proceduralize a set of skills in a safe and relatively inexpensive environment.

The MSK was selected for the training domain because it represents a high performance task. Although little task information is required, extensive practice is required to automate performance. Flight controllers need to automate use of the MSK so that most or all of their attention is available for performing other important console tasks. In addition, the MSK ITS provides a demonstration of a training system that could be expanded to other Propulsion Console tasks, highlighting future

potential training benefits for Propulsion flight controllers. Finally, the MSK ITS has implications for other flight controllers because the MSK is used not only on the Propulsion Console but also on other flight control consoles to perform similar tasks. Thus, this tutoring system has potential training benefits for flight controllers in general.

The MSK ITS includes a domain expert (i.e., an expert model), a trainee model, a training session manager, a scenario generator, and an user interface.[4] The domain expert includes information on how to perform the task. The trainee model includes a record of student performance. The training session manager provides information to the student on performance accuracy and speed. The training session manager also determines the amount and form of remediation to provide. The scenario generator provides variations of the task actions to the student. Finally, the user interface enables the student to interact with the system—for this ITS using a 3-key mouse and five function keys. (The function keys were only used during automaticity training.)

As mentioned above, the training session manager provides information on errors and determines the remediation. The error messages and remediation provided depend on the phase of training. Training is provided in three phases: declarative, speed, and automaticity training. In the declarative phase, task action steps are first described; guided examples are then provided, followed by unguided examples. Guided examples require students to complete an action step following a prompt. Unguided examples require students to complete all steps in an action without being prompted at each step. To complete declarative training, the student must correctly perform two consecutive guided examples, then two unguided examples. Speed training requires students to perform actions correctly and within a specified amount of time. Finally, automaticity training requires students to perform dual tasks correctly and at a specified speed. The primary task is the performance of MSK actions. The secondary task involves correctly responding to patterns of beeps. For both speed and automaticity training, training is completed when the student has accurately performed each task action twice and within a specified amount of time.

Error messages are provided immediately following an incorrect step during initial training (i.e., during guided examples) and following completion of a set of action steps during later training (i.e., during unguided examples, speed, and automaticity training). If an error is made during training, the student is remediated to the previous level of training. For example, if an error is made during speed training, the student is given an unguided example; if an error is made on an unguided example, s/he is provided a guided example to perform. In addition, the amount of tutoring content provided increases for successive occurrences of a given error during guided examples.[5, 6] This is consistent with recommendations made by other researchers. Remediation during automaticity training occurs only if speed and accuracy criteria are not met, rather than after each occurrence of an error.

METHOD

Task Overview

The MSK ITS trains students to perform five console operations: TV Channel (TV Chan), Display Request (Disp Req), Display Decoder Drive (DDD), Analog Event System (AES), and Flight Select (Flt Sel). In addition, two operations have variations. The AES operation includes: Select (AES Sel) and Deselect (AES Des). The DDD operation includes: Select (DDD Sel), Release (DDD Rel), Reset Operation (DDD Reset Op), Reset Critical (DDD Reset Crit), Select Drive (DDD Sel Drive), Select Datatype (DDD Sel Data), and Select Lamp (DDD Sel Lamp). Thus, the student learns a total of 12 task actions relating to 5 console operations. In the current project, the criterion for promoting students from speed training to automaticity training is two actions completed without error and in less than 20 seconds on each of the five console operations. The criterion for completion of automaticity training is two actions completed without error and in less than 40 seconds on each of the five console operations. Moreover, the actions must be completed with 100% accuracy on the secondary task (responding to beep patterns). Students were asked to respond to two target beep patterns (e.g., long-long-short, short-long-short) and not respond to false alarms (i.e., any of the remaining five beep patterns). Beep patterns were administered at 3-second intervals.

Subjects and Procedure

Three students completed training on the MSK ITS. Two students were flight controllers: one was a certified flight controller in on-board navigation with 3 years experience; one was a novice flight controller on the trajectory console with 6 months experience. Both were familiar with the MSK as used on their console, but unfamiliar with its use on the Propulsion console. The third student was a researcher in ITS's with no console experience. Students were asked to complete training on the ITS. All instructions were provided by the tutoring system. Additional informal observations and comments were collected from the students on ITS content, functioning, and the user interface.

Measures

Performance data was collected by the ITS. For Level 1 (Declarative) training, performance measures included number of trials and number of errors. For Level 2 (Speed) training, performance measures included number of trials, number of errors, number of successful trials, number of trials during remediation, and number of errors during remediation. A successful trial was operationalized as completing an action correctly in less than 20 seconds. For Level 3 (Automaticity) training, performance measures included number of trials, number of errors, number of successful trials, number of remediation

cycles, number of trials during remediation, number of errors during remediations. A successful trial was operationalized as completing an action correctly in less than 40 seconds with 100% accuracy in responding to beep patterns. Two successful trials of each of the five operations was required to complete speed training and to complete automaticity training. Results will be reported by task action for Level 1 training, but across actions for Levels 2 and 3. In addition, performance data was collected by action type during Levels 2 and 3 training to assess performance speed.

RESULTS

For Level 1 (Declarative) training, the three students performed between 45 and 54 trials with between 6 and 9 errors, averaging 49 trials and 7.3 errors (see Table 1). Further, the students required between 90 and 100 minutes to complete Level 1 training. Moreover, each student completed additional training on DDD tasks beyond the two consecutive, correct trials. It is unclear why the ITS administered the additional task trials. In some cases the additional trials involved actions on which students had previously made no errors. Further information is needed to clarify this issue.

For Level 2 (Speed) training, the three students performed an average of 30 trials (see Table 2). Results are reported across task actions for this training level. Students required an average of 48.3 minutes to complete the training. During this time, students made 3.7 errors on average. Remediation trials were administered after each error. On average, students completed 21.7 trials during remediation, making an average of 3.7 errors during the remediation trials.

One interesting point is that the remediation provided following an error did not necessarily correspond to the action on which the error was made. For Flt Sel, TV Chan, and Disp Req, an error was followed by remediation trials on the same action. However, an error on AES Sel was often followed by remediation trials on AES Des. Similarly, an error on one of the 7 DDD actions was often followed by remediation trials on other DDD actions but not on the DDD action on which the error was made. It would seem more beneficial to provide remediation on the action on which the error was made.

Another interesting point is that students were not returned to speed training following two consecutive, correct actions although this was the criterion stated. For example, Student 1 correctly performed the Disp Req action 7 consecutive times and DDD Res action 5 times before returning to speed training. Student 2 correctly performed the Disp Req action 9 consecutive times before being returned to speed training and correctly performed the Disp Req action 3 times following a second error and the AES Sel action 3 times. Student 3 also correctly performed actions 3 to 4 consecutive times during remediation trials before returning to speed training. Additional information is needed on the decision rules used by the ITS for providing remediation.

TABLE 1. – PERFORMANCE IN LEVEL 1 (DECLARATIVE) TRAINING.

	Student 1		Student 2		Student 3	
Operation/ Variation	# of Trials	# of Errors	# of Trials	# of Errors	# of Trials	# of Errors
Flt Sel	8	2	7	3	11	3
Disp Req	4	0	4	0	6	2
TV Chan	4	0	4	0	4	0
AES Sel	6	2	5	1	4	0
AES Des	3	0	7	1	2	0
DDD Sel	4	0	2	0	3	1
DDD Rel	2	0	4	1	6	1
DDD Reset Op	2	0	4	0	3	0
DDD Reset Crit	4	1	2	0	2	0
DDD Sel Drive	2	0	2	0	6	1
DDD Sel Data	7	2	2	0	3	1
DDD Sel Lamp	2	0	2	0	5	1

For Level 3 (Automaticity) training, the three students performed an average of 35.3 trials (see Table 3). They required an average of 75 minutes to complete the training. During this time, students made an average of 8.3 errors on actions. In Level 3 training, remediation trials were administered not after each error, but rather if a performance criteria was not met. The criteria involved beep response accuracy, performance speed (i.e., > 40 seconds), and action errors. Only one student received remediation during Level 3 training, performing 18 trials and making 3 errors across 2 remediation cycles.

It is interesting to note that unlike previous training levels, students required substantially different amounts of time to complete Level 3 training. The dual task paradigm was quite novel for the two flight controllers, at least partially explaining the differing time requirements. These two students also reported finding performing two

tasks at one time difficult. The more experienced task controller, however, appeared to have less difficulty. This may be due to greater familiarity and experience with MSK tasks and console use.

One other issue relating to automaticity training is that the ITS did not terminate training upon satisfying the performance criteria for two of the students. The performance criteria was two successful trials of each of the five operations (see above). Additional information is needed to determine why training was not terminated when expected.

TABLE 2. - PERFORMANCE IN LEVEL 2 (SPEED) TRAINING.

	Student 1	Student 2	Student 3
# of Trials	28	37	25
# of Errors	2	3	6
# Successful Trials	11	11	10
# Remediation Trials	16	18	31
# Remediation Errors	3	2	6
Time Required (Min.)	35	50	60

TABLE 3. - PERFORMANCE IN LEVEL 3 (AUTOMATICITY) TRAINING.

	Student 1	Student 2	Student 3
# of Trials	29	53	24
# of Errors	9	10	13
# Successful Trials	10	21	13
# Remediation Cycles	0	2	0
# Remediation Errors	NA	3	NA
Time Required (Min.)	50	150	25

Two additional analyses were conducted. First, performance speed was plotted against amount of task practice. One would expect students' performance to fit a learning curve during Level 2 and possibly Level 3 (as they learn to perform the secondary task). However, logarithmic functions did not explain as much of the performance variance as expected, ranged from 4% to 58% variance accounted for (see Figure 1).

Thus, a second analysis was conducted to determine whether discrepancies from the expected learning curve could be explained in terms of amount of practice and response speed on specific task actions. The 12 actions each had 6 task steps—with the exceptions of TV Chan (5 steps) and Disp Req (7 steps). However, it may have taken students longer to perform actions they were less familiar with (i.e., had received fewer trials of practice on). For Level 2 training, though,

TABLE 4. - PERFORMANCE PRACTICE AND SPEED IN LEVEL 2 TRAINING

	Student 1		Student 2		Student 3	
Operation/ Variation	# of Trials	Average Response Time	# of Trials	Average Response Time	# of Trials	Average Response Time
Flt Sel	3	22.8	7	25.8	4	18.7
Disp Req	9	24.0	3	21.9	4	22.3
TV Chan	3	17.9	3	26.6	2	14.9
AES Sel	3	22.8	6	23.6	3	23.2
AES Des	0	NA	2	25.5	2	20.6
DDD Sel	5	24.1	3	19.0	0	NA
DDD Rel	1	21.9	5	20.7	2	19.0
DDD Reset Op	0	NA	1	42.5	0	NA
DDD Reset Crit	0	NA	0	NA	0	NA
DDD Sel Drive	0	NA	1	25.4	0	NA
DDD Sel Data	2	22.8	0	NA	1	21.5
DDD Sel Lamp	0	NA	2	35.6	1	23.7

the results do not indicate clear differences between response times for either amount of practice (i.e., number of trials) or action type (see Table 4). Similarly, for Level 3 training, the results do not indicate clear differences between response times for either amount of practice or action types (see Table 5). No specific pattern of response time differences were observed across students in either training Levels 2 or 3 with the exception that speed on a specific action type increased with additional task practice. For example, Student 1 increased response time on Disp Req from 39.9 to 16.1 seconds across 7 trials of practice.

TABLE 5. - PERFORMANCE PRACTICE AND SPEED IN LEVEL 3 TRAINING

	Student 1		Student 2		Student 3	
Operation/ Variation	# of Trials	Average Response Time	# of Trials	Average Response Time	# of Trials	Average Response Time
Flt Sel	3	32.8	7	41.3	3	25.3
Disp Req	3	24.5	10	33.6	4	25.0
TV Chan	6	20.4	9	18.4	2	29.0
AES Sel	2	33.0	4	24.8	3	32.8
AES Des	1	30.2	2	28.2	1	19.0
DDD Sel	1	34.8	4	26.9	0	NA
DDD Rel	1	40.5	3	22.0	1	29.5
DDD Reset Op	1	26.3	0	NA	0	NA
DDD Reset Crit	0	NA	2	43.4	1	16.6
DDD Sel Drive	0	NA	1	50.1	1	51.2
DDD Sel Data	0	NA	0	NA	1	31.0
DDD Sel Lamp	2	27.6	1	26.3	1	41.7

The lack of systematic differences in response times in different action types was unexpected given that smaller amounts of practice were received on some actions—most notably the 7 DDD actions. Indeed, as

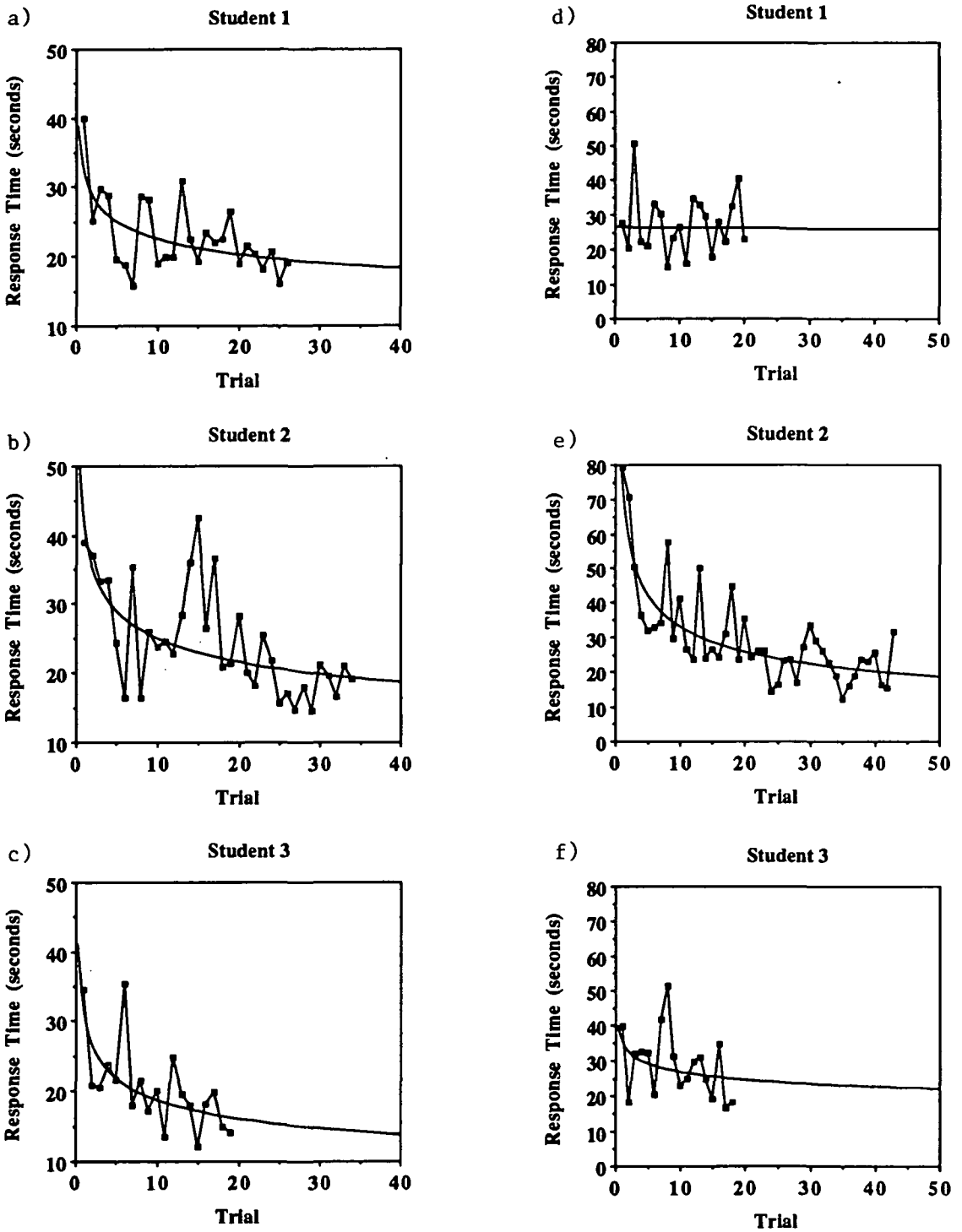


Figure 1.- Response time for task actions across trials. Speed training in panels a, b, & c. Automaticity training in d, e, & f.

shown in Tables 4 and 5, no practice at all was received on some actions in Levels 2 and 3 training. It was expected that the smaller amount of practice would result in substantially increased response times. This result was not observed. However, students did report more difficulty performing the 7 DDD actions and suggested the provision of additional training on these actions.

Finally, additional informal observations and comments were obtained from the students on ITS content, functioning, and the user interface. Several comments addressed training content. Specifically, the students noted that action steps did not have to be performed in the trained sequence on the job. Flight controllers using the MSK on the Propulsion or other consoles may perform the steps of an action in a variety of acceptable sequences. This was known by the software developer. However, it was necessary to require action steps to be performed in a specific sequence to facilitate the automaticity training. The required step sequences did not affect the novice flight controller although the experienced flight controller reported difficulty performing the steps in the required sequence. She had learned to use the MSK using alternate but acceptable sequences on the job and reported that her previous experience interfered with task performance on the ITS. Due to the small sample size (n=1) it is not possible to draw conclusions, though, regarding the possible interference between previous task experience and current ITS task performance. In addition, the experienced flight controller noted that it is unnecessary to perform some of the steps required for different actions after the console has been initialized for a flight (Flt Sel). Another comment addressed the amount of task practice provided on the DDD actions, suggesting that additional practice be provided for each action. This system currently treats the 7 DDD actions (and the 2 AES actions) as part of one operation, which may explain why remediation following an error in speed training did not necessarily match the erroneous action. Finally, the novice flight controller reported that the training was useful, providing information and experience she had not yet obtained on the job. Similarly, the experienced flight controller reported the ITS had potential training benefits for Propulsion Console and other flight controllers although she recommended modifying the task content to more closely resemble the job and address additional components of the job.

The students also commented on ITS functioning. One issue raised was that it was unclear what the criterion was for being promoted from one phase of declarative training to the next. For example, the flight controllers expressed some frustration about having to complete multiple guided and unguided trials on a given action before moving to the next action. This resulted in part from feedback messages stating that the student was demonstrating effective performance and then stating that additional practice would be provided on that action. It may be appropriate to indicate to students how much additional practice they can expect (e.g., they will be asked to complete one additional trial or to successfully complete two consecutive trials). Additional

explanation may also be appropriate during speed and automaticity training. For example, students did not initially realize during speed training that the trial started as soon as the "Goal" (i.e., the action assigned) was displayed on the screen. One student thought the trial began (the timer started) when she clicked the mouse the first time during the action. Moreover, students did not realize what the performance criteria were for successful completion of speed or automaticity training. It may be appropriate to give students more information about what performance levels are necessary to complete speed and automaticity training. Other student comments indicated that students did not understand the purpose of the secondary task during automaticity. Additional explanation could be provided regarding the purpose of secondary task performance.

Finally, student comments addressed the user interface. One issue raised was the use of scrolling rather than refreshing the tutoring/ information window. Declarative information and task assignments were made in a window at the lower right portion of the screen. Students reported difficulty reading the instructions provided, often rereading a portion of the window because it was unclear where new information or instructions appeared in the window. Refreshing the window when additional information or instructions appear would resolve this issue. A second issue related to the use of color. That is, students reported difficulty seeing the red cursor (an arrow) against the purple background. A third issue involved use of the mouse. Students were initially unclear regarding the different functions of the left, middle, and right mouse keys; the keyword descriptions provided in the MSK display window were apparently not sufficient. A brief statement explaining this could be provided at the start of training. A related issue was that two mouse keys (left and right) were required to key in numbers. Students suggested allowing the numbers to cycle from 9 to 0 (and 0 to 9) so that one mouse key could be used to change numbers, although students appear to want one key (e.g., left key) to cycle downward and a second key (e.g., right key) to cycle upward.

DISCUSSION AND CONCLUSIONS

The results indicate that students learned to perform 12 MSK actions using the ITS. Further, they were able to successfully complete training within approximately four hours. However, it is not clear that students received sufficient task practice to automatize the skill. Using the current 20-second and 40-second speed requirements during speed and automaticity training, respectively, subjects performed any given action a maximum of 37 times and as few as 2 times. To ensure automaticity it may be necessary to implement more stringent speed constraints which would result in additional task trials. Further, additional information is needed to determine how the ITS functions in terms of promoting students from one training level to the next and terminating training. Also, some revisions to training content may be appropriate to make ITS tasks more closely resemble job actions.

In terms of the ease of use, students required little or no assistance in using the ITS. The ITS provided instructions and task assignments which students could follow without outside assistance. However, some clarification or additional explanation may be appropriate to ensure that students understand the performance expectations and progress of training. Some modifications may also be appropriate to improve the interface, especially in terms of window refreshing and use of color.

Finally, the results and student comments provide an indication of potential training benefits of the ITS and modifications which may further improve this training system. The results indicate that students learn the training content. In addition, both flight controllers reported that the training content was useful, especially for novice flight controllers. Moreover, the ITS has potential benefits for flight controllers on other consoles given the similarity of MSK use across consoles. These potential benefits could be further increased by expanding the training content to other console activities.

REFERENCES

[1] Wenger, E. (1987). Artificial Intelligence and Tutoring Systems. Los Altos, CA: Morgan Kaufmann.

[2] Littman, D. & Soloway, E. (1988). Evaluating ITS's: The cognitive science perspective. In M. C. Polson & J. J. Richardson, Foundations of Intelligent Tutoring Systems. Hillsdale, NJ: Erlbaum.

[3] Goldstein, I. L. (1986). Training in Organizations: Needs Assessment, Development, and Evaluation. Monterey, CA: Brooks/Cole.

[4] Fink, P. K. & Overby, M. A. (1987). Research in Intelligent Tutoring Systems for Knowledge Poor Domains. (SwRI Proposal 14-4874a). San Antonio, TX: Southwest Research Institute.

[5] Burton, R. R. & Brown, J. S. (1982). An investigation of computer coaching for informal learning activities. In D. Sleeman & J. S. Brown (Eds.), Intelligent Tutoring Systems. NY: Academic Press, 79-88.

[6] Reiser, B. J., Anderson, J. R., & Farrell, R. G. (1985). Dynamic student modelling in an intelligent tutor for LISP programming. Proceedings of the 9th International Joint Conference on Artificial Intelligence (Vol. 1, pp. 8-14).

S10-34
20210

N91-27098 P-16

RE-EXAMINATION OF METMAN, RECOMMENDATIONS ON ENHANCEMENT OF LCVG, AND DEVELOPMENT OF NEW CONCEPTS FOR EMU HEAT SINK

Final Report

NASA/ASEE Summer Fellowship Program -- 1990

Johnson Space Center

TV 551589

Prepared by:	Amir Karimi, Ph.D.
Academic Rank:	Associate Professor
University & Department:	The University of Texas at San Antonio Division of Engineering San Antonio, Texas 78285
NASA/JSC	
Directorate:	Engineering
Division:	Crew and Thermal Systems
Branch:	Thermal Systems
JSC Colleague:	Chin H. Lin, Ph.D.
Date Submitted:	November 1, 1990
Contract Number:	NGT-44-005-803

ABSTRACT

METMAN is a 41-node transient metabolic computer code developed in 1970 and revised in 1989 by Lockheed Engineering and Sciences, Inc. This program relies on a mathematical model to predict the transient temperature distribution in a body influenced by metabolic heat generation and thermal interaction with the environment. A more complex 315-node model (Wissler model) is also available that not only simulates the thermal response of a body exposed to a warm environment, by including in the model the perfusion rate in muscle and the magnitude of counterflow heat exchange between large arteries and veins, but is also capable of describing the thermal response resulting from exposure to a cold environment . It is important to compare the the two models for the prediction of the body's thermal response to metabolic heat generation and exposure to various environmental conditions. Discrepancies between the two models may warrant an investigation of METMAN to ensure its validity for describing the body's thermal response in space environment.

The Liquid Cooling and Ventilation Garment is a sub-system of the Extravehicular Mobility Unit (EMU). Designed by ILC-Dover in the 1960's, this garment, worn under the pressure suit, contains the liquid cooling tubing and gas ventilation manifolds; its purpose is to alleviate or reduce thermal stress resulting from metabolic heat generation. There is renewed interest in modifying this garment through identification of the locus of maximum heat transfer at body-liquid cooled tubing interface.

The sublimator is a vital component of the Primary Life Support System (PLSS) in the EMU. It acts as a heat sink to remove heat and humidity from the gas ventilating circuit and the liquid cooling loop of the LCVG. The deficiency of the sublimator is that the ice, used as the heat sink, sublimates into space. There is an effort to minimize water losses in the feedwater circuit of the EMU. This requires developing new concepts to design an alternative heat sink system.

We directed our efforts to review and verify the heat transfer formulation of the analytical model employed by METMAN. We recommend a conceptual investigation of regenenerative non-venting heat-sink subsystem for EMU.

INTRODUCTION

Considerable effort has been made by physical scientists and engineers to develop mathematical models to describe natural phenomena through mathematical models. Those efforts include attempts to model the human thermal system. The models are designed to simulate the effects of heat generation, resulting from physical activities, and the heat exchange between the body and its environment on heat transfer within the body.

Although our understanding of many important physiological phenomena remains incomplete, it is important to use mathematical modeling for prediction of human response to a variety of stressful environmental conditions. These applications become essential in situations for which it is at least difficult, if not impossible, to conduct experiments on human subjects. Examples include the thermal effects on astronauts during lunar missions, or extravehicular activities (EVA), thermal response of humans while diving in cold water and the effect of thermal stress on workers who must perform their assigned duties under thermally hostile environments. Applications could be extended to various survival situations involving accidental immersion in cold water or military personnel who have to perform their duties under chemical or biological warfare environmental situations. The threat posed to humans by these conditions has long been recognized.

Mathematical modeling of human thermal systems dates back to 1947, when Machle and Hatch [1] introduced the concept of core and shell temperatures (rectal and mean skin temperatures, respectively). In their model the stored energy in the body was expressed as linear function of core. Pennes' cylindrical model [2] (1948) provided a reasonable temperature profile in the forearm. Wissler's first model [3] extended Pennes' model by employing six cylindrical elements to describe the entire human thermal system.

Extensive application of mathematical modeling of human thermal systems began with NASA's Apollo program. Stolwijh [4] developed a model to predict the astronauts' thermal response during the extravehicular activities (EVA). To achieve this a transient metabolic man computer program (METMAN) was developed to describe the temperature profile in the body. METMAN's initial construction was based on a two-model point model and later was extended to an eight-node, a 14-node, a 25-node and to the present 41-node model [5]. This model has been adequate to predict thermal behavior in the body resulting from thermal stress. Wissler's latest 250-node and 315-node physiological models [6,7] were mainly designed to describe the human thermal response resulting from exposure to cold environments. They have been used to analyze the performance of divers working in water as deep as 450 meters. In recent years they have also been used to describe human thermal experience working or exercising in a hot humid environment.

The advantage of Wissler's model over METMAN is that it has a wider range of applications. It can be used for predication of thermal response for humans exposed to hot or cold environments, while METMAN's range of application is limited to hot environments. Wissler's model considers factors which are important in cold subjects. They include the perfusion rate in muscle and counter flow heat exchange between large arteries and veins. This thermal model contains mass balances for oxygen, carbon dioxide, and lactic acid required in equations for perfusion and ventilation. However, Wissler's model is poorly documented and is difficult to use. METMAN's advantage over Wissler's

model is that it is well documented and easy to use. It also has been developed specifically for space application.

The problem of rapid increase in body temperature becomes significant when an individual has to wear special protective clothing to be isolated thermally and/or otherwise, from hazardous surroundings. Although such clothing protects the individual from a hostile environment, it also places additional thermal stress on the individual, as the generated metabolic heat cannot be directly dissipated into the environment by natural means. Therefore, to make such micro-environments habitable, fluid-conditioned suits have been designed to reduce thermal stress. A fluid-cooled garment is usually warn and is placed in direct contact with the skin.

A fluid-cooled garment may consist of cold liquid flowing in a network of flexible tubes which come in direct contact with the heat source (body), or conditioned air flowing against the body or a combination of both. A cold fluid (typically, in the range 5 to 20 °C) is used to remove a portion or all of the metabolically generated heat from the body. On return cycle, heat is rejected to a heat sink medium. The ILC Dover liquid cooled conditioned vest has been used in space applications. A modified (half-length) version of this suit is also been used by USAF for military application. In recent years USAF has directed its efforts in intermittent method of ambient air cooling. References [8-11] provide typical performances of liquid cooled garments and air conditioned suits.

A heat sink for a life support system may consist of an ice pack heat exchanger, a small vapor-compression cooling unit. A thermal heat sink is a vital component of the Primary Life Support System (PLSS) in the Extravehicular Mobility Unit (EMU). It is required to remove heat and humidity from the gas ventilating circuit, to cool the liquid cooling loop of the LCVG, and to eliminate other EMU equipment thermal load. The present heat sink subsystem in the EMU consists of a sublimator. The ice formed in the sublimator is used as the heat sink for heat removal and the vapor product is vented into space. To minimize the water losses in the feedwater circuit of the EMU, NASA has identified a need for developing a regenerable non-venting heat sink (RNTS) subsystem for advanced extravehicular activities (EVA). The subsystem must address the EMU thermal control requirements of the existing Shuttle Orbiter, the Space Station EVA mission, and the Man-Mars mission.

The Shuttle EVA missions require RNTS operations of up to four hours duration with a cumulative heat removal capacity of up to 7240 Btu. They require heat removal rates in the range 310-3445 Btu/hr, with an average rate of 1810 Btu/hr for the entire EVA mission. The liquid-cooling and ventilation garment (LCVG) inlet water temperature should not exceed 45°F at maximum thermal load during the Shuttle EVA mission. The Space Station EVA mission has a maximum duration of 8 hours with a cumulative heat removal capacity requirement of up to 11,680 Btu. It requires an RNTS operation with a heat removal rate in the range of 310-2695 Btu/hr, with an average rate of 1460 Btu/hr for the entire mission. The liquid-cooling and ventilation garment (LCVG) inlet water temperature should not exceed 60°F at maximum thermal load during the space station EVA mission.

A series of feasibility studies have been conducted [12-14] for the development of a RNTS thermal control for the EMU. The studies suggested a thermal control subsystem utilizing a water/alcohol phase change thermal storage for the Shuttle EVA mission and a thermal control subsystem consisting of a vapor compression heat pump/space radiator

integrated with a water/ethanol phase change medium thermal sink for the Space Station EVA mission. The proposed thermal control subsystems are heavy for the lunar and Mars EVA missions.

This report concentrates on a review of mathematical modeling and physical concepts used in the METMAN. Its purpose is to point out shortcomings of the model and, when possible, to make recommendations for improvement.

DESCRIPTION OF THERMAL MODEL

Both METMAN and Wissler's models are transient-state models which predict thermal and metabolic changes that occur during a certain period of time. Given initial values for all dependent variables and specification of independent variables, they both evaluate nodal temperatures at various time periods and compute thermal storage within the body. Each program is capable of predicting body-thermal behavior subject to different modes of operation. The modes of operation of METMAN include shirt-sleeve, normal-suited intravehicular activity (IVA), extravehicular activity (EVA) and helmet off. The METMAN and Wissler's models are based on similar descriptions of the human body. In METMAN, as shown in Figure 1, the human form is divided into 10 cylindrical elements: head, trunk, right and left arms, right and left legs, and right and left feet. Each element is further divided into four concentric compartments. They include a central core, muscle, fat layer and skin. Each compartment is represented by a nodal point. Therefore, including the central blood as a compartment, the METMAN model is constructed around 41 nodal points. The temperature throughout each compartment is assumed to be uniform, equaling its nodal point.

The Wissler model as shown in Figure 2 is based on 15 major cylindrical elements: head, upper torso, lower torso, right and left proximal arms, right and left medial arms, right and left distal arms, right and left proximal legs, right and left medial legs, right and left distal legs. Within each element there are 15 nodal points representing a conglomeration of tissue, bone, fat, skin and a vascular system containing arteries, veins and capillaries. The model includes up to six additional nodal points, external to the body, for computation of temperature and accumulation of sweat within clothing. Thus, the nodal points in this model may exceed 300, depending upon the mode of operation.

DISCUSSION

A human thermal model must account for the following factors. It should be capable of describing temperature as a function of position and time. It should account for the geometry of the human body. It should consider that properties vary with position throughout the body. It should account for metabolic heat generation resulting from various types of work and environmental conditions. It should account for heat and mass transfer in the respiratory tract. And it should account for all thermal interactions of body with its environment. Thus it is important that the physical modeling of thermal behavior in the body also include the effects of fluid-cooled garment-body interactions.

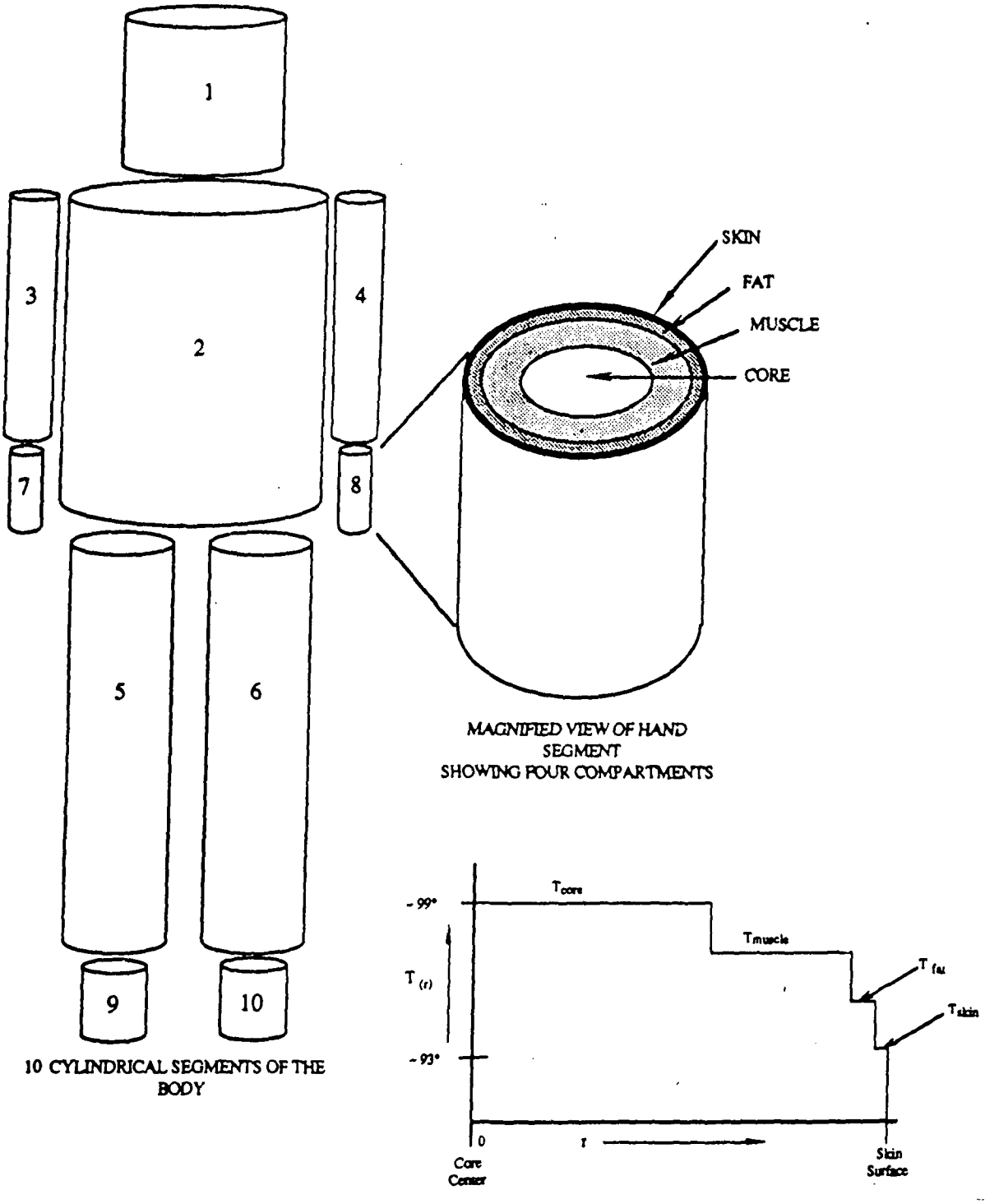


Figure 1. Geometric Representation of Human Body by METMAN Model

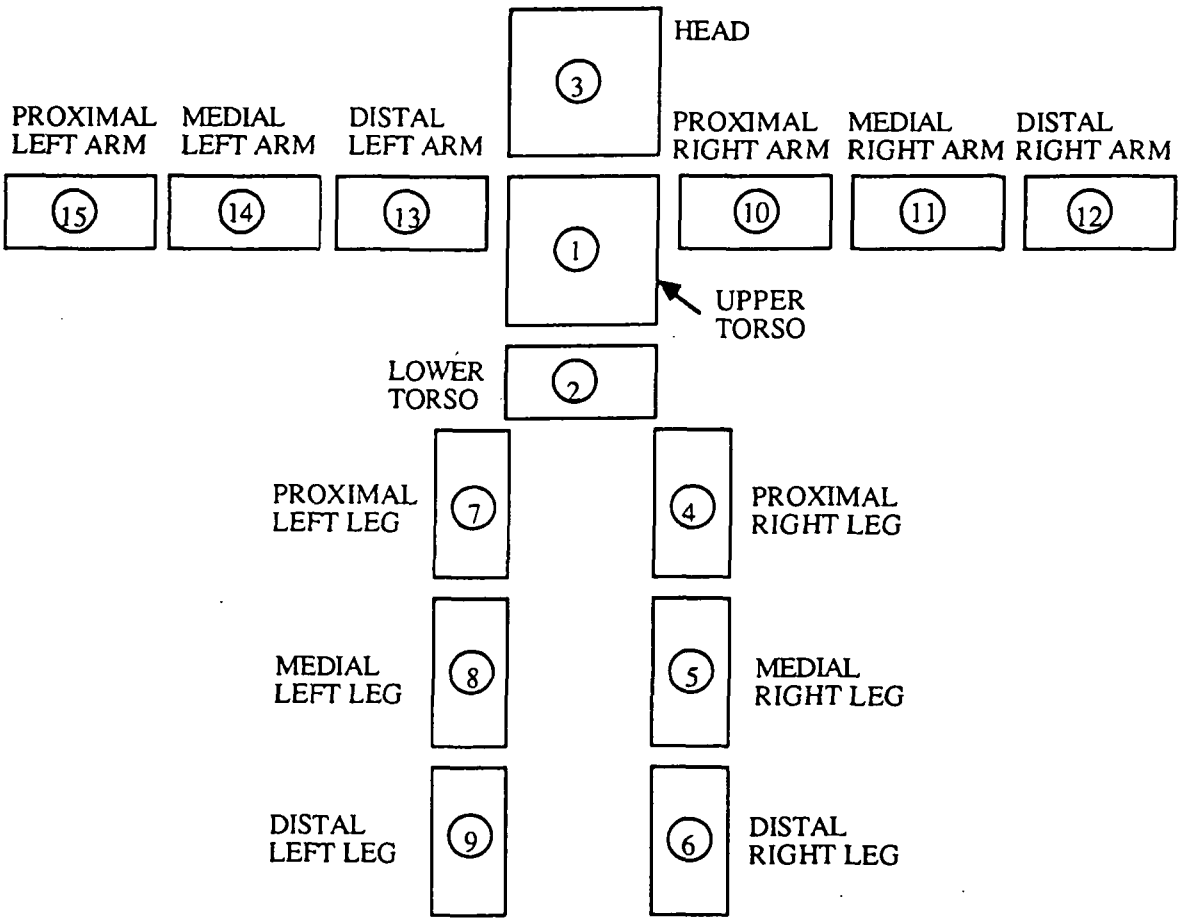


Figure 2. Geometric Representation of Human Body by Wissler's Model

The heat transfer formulation of METMAN model is described in detailed in [5]. Therefore, we limit this report to few equations necessary, and refer the reader to this document for detailed description of thermal system.

The energy balance for each compartment is based on the following general relationship.

$$Q_{st} = Q_m \pm Q_c \pm Q_r - Q_e \pm Q_k \pm Q_{res} \pm Q_{LCG} \qquad (1)$$

Where Q_{st} is the rate of energy storage, Q_m is the metabolic heat generation, Q_c is convective surface heat transfer, Q_r is the radiation heat transfer at the surface, Q_e is the rate of evaporation losses, Q_k is the rate of conductive heat transfer, Q_{res} is the rate of heat transfer through respiratory tract and Q is the heat transfer with the liquid-cooled garment.It should be noted that Q_c, Q_r and Q_e are zero for internal compartments and they are important at skin compartment.

The rate of energy stored in each compartment can be expressed by:

$$Q_{st} = \left(m\, c_p \frac{\partial T}{\partial t} \right) \qquad (2)$$

The rate of heat generation, Q_m, in the body is the sum of the basal metabolic rate and heat generated through work or shivering by muscles. METMAN model considers the effect of the height and the mass of an individual in evaluating the basal metabolic rate. The relations used in evaluation of heat generation are given in [5]. METMAN uses a distribution coefficient, K_{mw}, to evaluate the heat produced by work in each muscle compartment. These distribution coefficients are presented in Table 1. It indicates that the leg muscles' contribution to heat generation dominates all other body elements. These distribution coefficients are reasonable for work in a gravitation field. However, in a micro-gravity environment the major work is done by the arms. Therefore, it appears that the METMAN uses unreasonable distribution coefficients for the evaluation of heat production in the muscle compartments.

The heat transfer between internal body compartments occurs through conduction and vascular convection. The METMAN model assumes a uniform temperature, equal to the nodal point temperature, throughout each compartment. It employs the thermal resistant analogy, as shown in Figure 3, to evaluate the conductive heat transfer between two adjacent compartments. The internodal conductances are calculated based on the following relationship:

$$G_j^{i \to i+1} = \frac{2 \pi l_j}{\frac{\ln \left(\frac{r_i}{r_{n,i}} \right)}{g_i} + \frac{\ln \left(\frac{r_{n,i+1}}{r_i} \right)}{g_{i+1}}} \qquad (3)$$

The variables in Equation (3) are defined in Figure 3.

Table 1. DISTRIBUTION COEFFICIENT, K_{mw}, OF HEAT PRODUCED BY WORK AMONG THE MUSCLE COMPARTMENTS

Body Element	K_{mw}
Head	0.00
Trunk	0.30
Arms	0.08
Legs	0.60
Hands	0.01
Feet	0.01
	1.00

A uniform temperature distribution assumption for each compartment would present temperature discontinuities at the interface of two adjacent compartments. To improve the model we recommend the addition of appropriate boundary conditions for each compartment. Along the axis of each cylindrical element the boundary condition (central core) is

$$\left(\frac{\partial T_j}{\partial r}\right)_{r=0} = 0 \qquad (4)$$

Equation (4) indicates that along the element's central axis the temperature is either a minimum or a maximum. Subscript j refers to j^{th} element. At the interface between the compartments, both temperature and thermal conduction must be continuous. This implies that

$$(T_j)_{r_i} = (T_j)_{r_{i+1}} \qquad (5)$$

and

$$\left(k_j \frac{\partial T_j}{\partial r}\right)_{r_i} = k\left(\frac{\partial T_j}{\partial r}\right)_{r_i} \qquad (6)$$

subscripts i and i + 1 refers to the i^{th} and $(i + 1)^{th}$ compartment of the j^{th} element.

Intrinsic convection is the result of heat transfer between the tissue compartments and the vascular system. The heat transfer between skin and environment depends on the micro or macro environments. The mechanism for this mode of heat transfer may include extrinsic convection, radiation, latent heat transfer (evaporation of sweat), heat transfer to the undergarment, heat transfer to the liquid-cooled garment and respiratory heat transfer. These modes of heat transfer are described in detail in [5].

Four active controllers are included in the METMAN model. These are sweat production, shivering, vasodilation and vascodilation. They play an important role in maintaining the body in an isothermal state. The relations used to describe the human thermoregulatory system are based on experimental results.

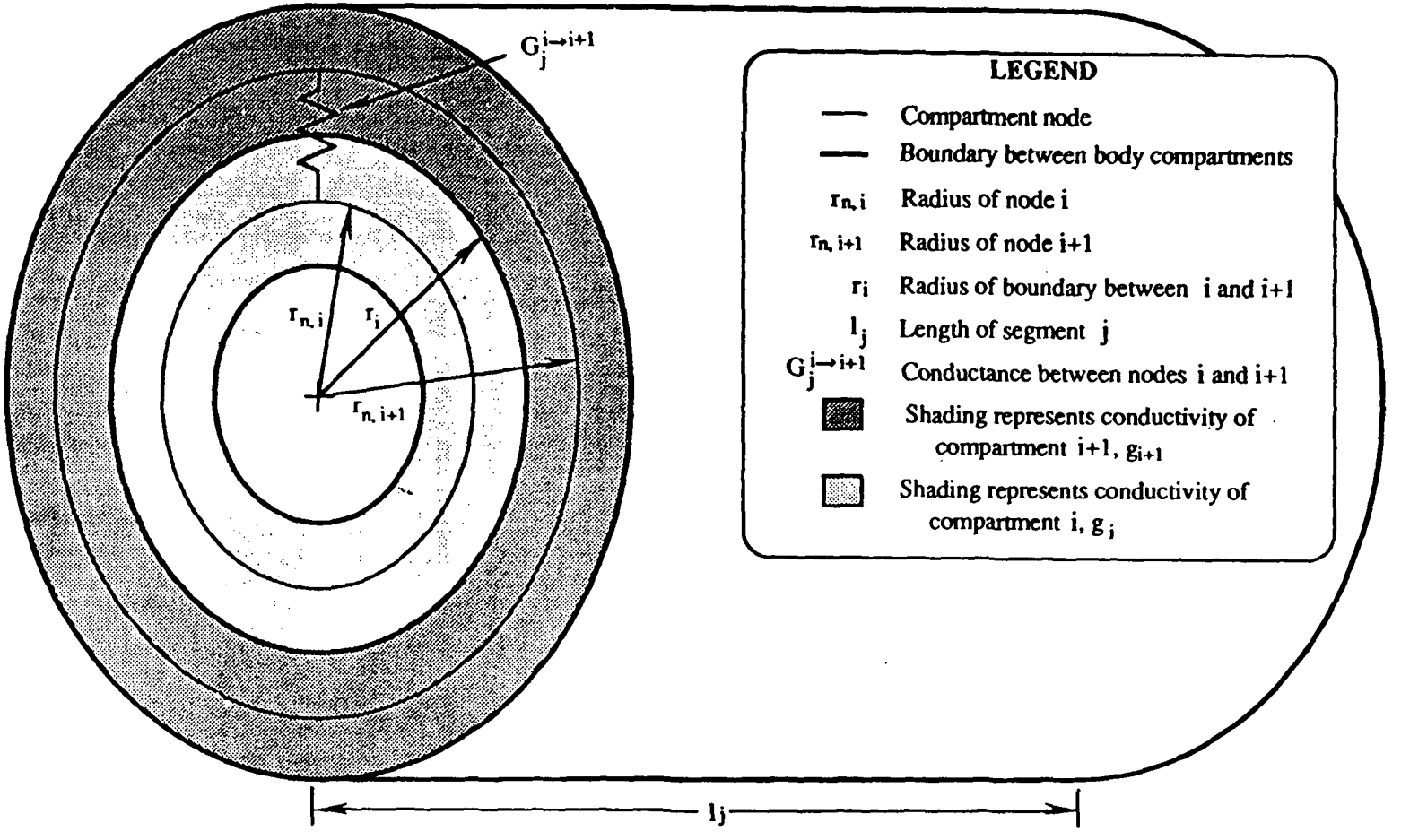


Figure 3. Radial Conductance Between Two Body Compartments

In various mode of operations METMAN compares the effect of natural convection with forced convection and chooses the greater value for convective heat transfer. The forced convection is presented by:

$$q_{senc} = 0.212\ (P_{cab}\ V_{cab})^{1/2}\ \kappa_1\ A_c\ (T_{ug} - T_{cab}) \tag{7}$$

and the free convection is given as:

$$q_{senc} = 0.06\ [\ G\ (P_{cab})^2\ |\ T_{ug} - T_{cab}\ |\ ]^{1/4}\ \kappa_2\ A_c\ (T_{ug} - T_{cab}) \tag{8}$$

However relation are available for combined effects of free and forced convection. The following relationship is used to establish the dominant regions of free and forced convection.

$$\frac{Gr}{Re^2} = \frac{g\,\beta\,(T_w - T_\infty)\,L}{U_o^2} \tag{9}$$

When the LHS of Eq. 9 >> 1 then the free convection dominates. If the LHS of Eq. 9 << 1, then forced convection dominates. in all other cases forced and free convections are of comparable magnitude. References [15-23] describe the combined free and forced convections for various geometric configurations.

In general heat transfer between the skin and environment depends on the mode of operation. The METMAN model can be applied to various modes of operation. This includes the shirt-sleeve model which emulates a man in a cabin, or in any environment, wearing an undergarment, as well as the suited mode which describes a man in a space suit. The space suit model itself can be applied to several modes of operation. The EVA suited mode is used to model extravehicular activity. The IVA model is used to identify intravehicular activity. METMAN is also capable of modeling the suited modes of operation for both the helmet off and the purge flow activity.

The METMAN program utilizes a number of subroutines and functions which are generally used to determine physical parameters and thermophysical properties. Among these subroutines are Function DEWPT and Function VVP. Function DEWPT evaluates the saturation temperature for a given vapor pressure. Function VPP determines saturated vapor pressure for a given temperature. Both functions are based on old relations in which three separate equations are used to describe the vapor pressure curve. Each equation is valid only within a limited range of temperature and pressure. In addition, the equation used by METMAN for the evaluation of dewpoint at very low pressures (0.0<P<0.0185 Psia) is highly inaccurate.

There are modern equations of state and fundamental equations [23-27] available now that describe a wide range of thermodynamic regions with a single relation. The fundamental equation developed by Haar et. al [27] is the most recent and accurate relation for thermodynamic properties of water. A FORTRAN computer code is included in [14] for the evaluation of the thermodynamic properties of water. Inclusion of this computer code will enhance the METMAN model.

There are many parameters that should be considered in selection and design of a thermal control subsystem for the EMU. These include the size and the weight of the subsystem, thermophysical and thermal transport properties, and operational safety during an EVA mission. For example, the selection of a heat sink medium should not be based solely on the total heat capacity of the heat sink material, but should also consider the size of equipment required for the freezing and melting processes.

We must identify a heat sink concept that will minimize the size and the weight of the thermal control subsystem required for the EMU. To achieve this goal we must to re-examine and extend the existing alternative technologies for the heat sink medium. This includes a feasibility study of:

- o phase change materials with a suitable melting-freezing temperature
- o heat pipe technologies
- o eutectic binary solutions
- o thermoelectric technologies
- o heat of solution of mixtures
- o endothermic and exothermic heat of chemical reactions

CONCLUSIONS AND RECOMMENDATIONS

The METMAN thermal model in its present form is adequate to describe the heat transfer phenomena in a warm subject. However, there are deficiencies present in the current model. Following is a summary of the shortcomings of the present model and recommendations for improvement.

1. The distribution coefficients used by METMAN to evaluate the heat generation in the muscle compartments are unreasonable for space application. These coefficients should be adjusted for EVA missions in the microgravity environment. Further physiological study is recommended in this area.

2. The METMAN model can be further improved by including appropriate boundary conditions along the central axis of each cylindrical element and along the interface of two adjacent compartments. These boundary conditions are presented in Equations (4-6) in the text.

3. The present equations used by METMAN for the evaluation of dewpoint temperatures and water vapor pressures are inaccurate. We recommend that the FORTRAN computer code given in [14] for the evaluation of the thermodynamic properties of water be incorporated into the METMAN model.

4. Addition of external nodal points for the evaluation of temperatures and the amount of sweat accumulation will enhance the METMAN model

5. The heat transfer formulation of METMAN is based on old knowledge. We recommend that METMAN be revised to include advances in heat transfer research.

A conceptual investigation in desired to determine alternative heat sink for the EMU.

REFERENCES

1. Machle, W., and Hatch, T.F. Heat: Mass exchanges and physiological responses. Physiol. Rev. 27, 200-227, 1947.

2. Pennes, H.H. Analysis of tissue and arterial blood temperatures in the resting forearm. J. Appl. Physiol. 1, 93-122, 1948.

3. Wissler, E.H. Steady-state temperature distribution in man. J. Appl. Physiol. 16, 734-740, 1961.

4. Stolwijk, J. A. J., and Hardy, J.D. Temperature regulation in man - A theoretical study. Pflugers Arch. 291, 129-162, 1966.

5. Bue, G.C., Milliken, A.H., Cook, D., and Franklin, G.C., "Computer program documentation-41-Node Transient Metabolic Man Program," Lockheed Engineering and Sciences Co., LESC-27578, CTSD-0425, October 1989.

6. Wissler, E.H. Comparison of results obtained from two mathematical models - A simple 14-node model and a complex 250-node model. J. Physiol. (Paris) 63, 455-458, 1970.

7. Wissler, E.H. A mathematical model of the human thermal system. Bull. Math. Biophys. 26, 147, 1964.

8. Burton, D. R., "Performance of Water Conditioned Suits," J. Aerospace Medicine, (May 1966), pp. 500-504.

9. Chen, T. Y., Bomalaski, S.H. and Constable, S. H., "Development of a Light Weight Ambient Air Cooling Unit for Use During Operations in a Chemical Defense Environment," USAF/Sam Brooks AFB., 1990.

10. Constable, S. H., "Alleviation of Thermal Stress in Ground Crew Supporting Air Operations During a Chemical Warfare Scenario," USAF/SAM Brooks AFB.

11. Harrison, M. H. and Belyavin, A.J., " Operational Characteristics of Liquid-Conditioned Suits," J. Aviat. Environ. Med., 49(8) (August 1978), pp. 994-1003.

12. "Subsystem Conceptual Study Report for the Regenerable Non-venting Thermal Sink (RNTS)," United Technologies, Hamilton Standard, Final Report, Contract No. NAS9-16609, December, 1982.

13. "Development of a Prototype Thermal Control Subsystem for Advanced Extravehicular Activity Portable Life Support System (Regenerable Nonventing Thermal Sink - RNTS I),", United Technologies, Hamilton Standard, Final Report, NAS9-16609, DRL No. T-1700, March 1986.

14. "The Advanced Nonventing Extravehicular Activity Cooling System (RNTS II),", United Technologies, Hamilton Standard, Final Report, NAS9-17301, DRL No. T-1833, January 1988.

15. Lloyd, J.R.., and E.M. Sparrow: "Combined Forced and Free Convection Flow on Vertical Surfaces, Int. J. Heat Mass Transfer, 13:434-438 (1970).

16. Wilks, G. "Combined Forced and Free Convection Flow on Vertical Surfaces," Int. J. Heat Mass Transfer, 16:1958-1964 (1973).

17. Redekopp, L.G., and A.F. Charwatt: "Role of Buoyancy and the Boussinesq Approximation in Horizontal Boundary Layers," J. Hydronautics, 6:34-39 (1972).

18. Robertson, G.E., J.H. Seinfeld, and G.E. Leal: "Combined Forced and Free Convection Flow Passed a Horizontal Flat Plate," AIChE J., 19:998-1008 (1973).

19. Yuge, T.: "Experiments on Heat Transfer from Spheres Including Combined Natural and Free Convection," J. Heat Transfer, 82C:214-220 (1960).

20. Mucoglu, A., and T.S. Chen: "Mixed Convection across a Horizontal Cylinder with Uniform Surface Heat Flux," J. Heat Transfer, 99C:679-682 (1977).

21. Newell, P.H., and A.E. Bergles: "Analysis of Combined Free and Forced Convection for Fully Developed Laminar Flow in Horizontal Tubes, J. Heat Transfer, 90C:83-93 (1970).

22. Sabbagh, J.A., A. Aziz, A.S. El-Ariny , and G. Hamad: "Combined Free and Forced Convection in Circular Tubes," J. Heat Transfer, 98C:322--324 (1976).

23. Patankar, S.V., S. Ramadhyani, and E.M. Sparrow: "Effect of Circumferentially Nonuniform Heating on Laminar Combined Convection in a Horizontal Tube," J. Heat Transfer, 100C:63-70 (1978).

24. Irvine, T.F., and Lilely, P.E., Steam and Gas Tables with Computer Equations, Academic Press,New York, 1984.

25. Karimi, A. and Lienhard, J., "A Fundamental Equation Representing Water in Stable, Metastable and Unstable States", EPRI NP-3328, RP 1438-2, Final Report, December 1983, 322 pp.

26. Keenan, J.H., Keyes, F.G., Hill, P.G., and Moore, J.G., Steam Tables (S.I. Units), John Wiley & Sons, New York, 1978.

27. Haar,L., Gallagher, John S., and Kell, G.S.,NBS/NRC Steam Tables, McGraw Hill, New York, 1984.

BIBLIOGRAPHY

1. Cook, D.W., Nguyen, N.H., "Liquid Cooled Garment Overall Heat Transfer Coefficient Definition," Technical Memorandum, Lockheed Engineering and Management Services Company, LEMSCO-15410 (August 1980).

2. Cook, D.W., and Nguyen, N. H., "Shirt Sleeve Comfort as Affected by Body Size," Lockheed Electronics Company, LEC-13401 (May 1979)

3. Cook, D. W., "Addition of Solar Effect Option to 41-Node Man Program," Technical Memorandum, Lockheed Electronics Co., LEC-10195 (March 1977).

4. Fanger, P.O., Thermal Comfort, McGraw-Hill, New York (1970).

5. Ganong, W. F. Review of Medical Physiology, Lange Medical Publications, 1971.

6. Houdus, Y. and Ring, E. F., Human Body Temperature, Its Measurement and Regulation, Plenum Press, New York, 1982.

7. Inouye, J., Hick, F.K., Tesler, S.E., and Keetan, R.W., "Effect of relative humidity on heat loss of men exposed to environments of 80, 76, and 72°F," ASHRAE Trans., Vol. 59 (1953), p. 329.

8. Iovine, J.V., "Breakdown of Liquid Cooled Garment Overall Heat Transfer Coefficient," Technical Memorandum, Lockheed Engineering Management and Sciences Company, LEMSCO-25665, (June 1988).

9. Karimi, A., "A Thermal Evaluation of the 'LSSI' Liquid-Cooled System: An Engineering Perspective", USAF Summer Faculty Program, Technical Report, Vol. 1/2, Report No. 48, September 1983, 23 pp.

10. Karimi, A. and Patel, P., "A Liquid-Cooled Garment-Body-Interfaced Model Study for Intermittent Cooling", Final Report, Brooks AFB (In preparation).

11. Karimi, A., "Design and Evaluation of Intermittent Liquid Cooling System for Ground Crew Personnel", Final Report, Brooks AFB, 1988.

12. Karimi, A., "Thermodynamic Analysis of Liquid Cooling Technology", Final Report, SCEEE/ARB/85-63, May 1986, 52 pp.

13. Karimi, A., "A Thermal Evaluation of a Portable Battery-Operated Vapor Compression Cooling System", Final Report UES/F49620-85-C-0013, September 1985, 25 pp.

14. Karimi, A.," An Experimental Study of the Portable Liquid-Cooled System", Final Report, SCEEE/RPI Subcontract No. 83 RPI 20, December 1984, 35 pp.

15. Keslake, D. M., "Errors Arising from the Use of Mean Heat Exchange Coefficients in the Calculation of the Heat Exchanges of a Cylindrical Body in Transverse Wind," Temperature, Vol. 3, Biology and Medicine, J.D. Hardy (Editor), New York, Reinhold (1963), pp. 183-190.

16. Lin, C.H. and Berenson, P.J., "Study of the Thermal Processes for Man in Space," Air Research Manufacturing Co., NASA CR-216 (1965).

17. Nguyen, N. H., Hu, A.S., "Body Size Effect on Shuttle EMU Comfort Curves," Technical Memorandum, Lockheed Engineering and Management Services Company, LEMSCO-16363, (March 1981).

18. Nguyen, N. H., "Shuttle EMU Comfort as Affected by Body Size," Technical Memorandum, Lockheed Engineering and Management Services Co., LEMSCO-14567 (March 1980).

19. Nguyen, N. H., "Updated Shuttle EMU Comfort Curves, " Technical Memorandum, Lockheed Electronics Company, AMD/173 (December 1978).

20 Olschewski, H. and Bruck, K., " Thermoregulatory, Cardiovascular, and Muscular Factors Related to Exercise after Precooling," J. Appl. Physiol., 64(2), 1988, pp. 803-811.

21. Pimental, N. A., Cosimini, H. M., Sawka, M. N., and Wenger, C. B., "Effectiveness of an Air-Cooled Vest Using Selected Air Temperature and Humidity Combinations," Aviation, Space, and Environmental Medicine, (February 1987), pp. 119-124.

22. Shitzer, A. and Chato, J.C. "Analytical Solutions to the Problem of Transient Heat Transfer in Living Tissue," J. Biomedical Engineering, Vol. 100, (November 1978), pp. 202-210.

23. Smith, E., "Establishing a Minimum UA for LCG," Internal Memorandum, Hamilton Standard, I.D. # H 78-003-GC (March 1978).

24. Stolwijk, J.A. J, " A Mathematical Model of Physiological Temperature Regulation in Man," NASA CR-1855, (August 1971).

25. Stolwijk, J.A.J., "Expansion of a Mathematical Model of Thermoregulation to Include High Metabolic Rates," NASA CR-1855 (1971).

26. Werner, J., and Buse, M., "Temperature Profiles with Respect to Inhomogeneity and Geometry of the Human Body'" J. Appl. Physiol, 65 (3), 1988, pp. 1110-1118.

A NONRECURSIVE ORDER N PRECONDITIONED CONJUGATE GRADIENT / RANGE SPACE FORMULATION OF MDOF DYNAMICS

Final Report

NASA/ASEE Summer Faculty Fellowship Program - 1990

Johnson Space Center

Prepared by:	Andrew J. Kurdila
Academic Rank:	Assistant Professor
University & Department:	Texas A&M University Department of Aerospace Engineering College Station, Texas 77843
NASA/JSC	
Directorate:	Engineering
Division:	Navigation, Control and Aeronautics
Branch:	Guidance and Control Systems
JSC Colleague:	John Sunkel Tuyen Hua Larry McWhorter
Date Submitted:	August 10, 1990
Contract Number:	NGT-44-005-803

ABSTRACT

While excellent progress has been made in deriving algorithms that are efficient for certain combinations of system topologies and concurrent multiprocessing hardware, several issues must be resolved to incorporate transient simulation in the control design process for large space structures. Specifically, strategies must be developed that are applicable to systems with numerous degrees of freedom. In addition, the algorithms must have a *growth potential* in that they must also be amenable to implementation on forthcoming parallel system architectures. For mechanical system simulation, this fact implies that

(ii) Algorithms are required that induce parallelism on a fine scale, suitable for the emerging class of highly parallel processors.

(iii) Transient simulation methods must be automatically load balancing for a wider collection of system topologies and hardware configurations.

This paper addresses these problems by employing a combination range space / preconditioned conjugate gradient formulation of multi-degree-of-freedom dynamics. The method described herein has several advantages. In a sequential computing environment, the method has the features that:

(i) By employing regular ordering of the system connectivity graph, an extremely efficient preconditioner can be derived from the "range space metric", as opposed to the system coefficient matrix.

(ii) Because of the effectiveness of the preconditioner , preliminary studies indicate that the method can achieve performance rates that depend linearly upon the number of substructures, hence the title "Order N".

(iii) The method is non-assembling, i.e., it does not require the assembly of system mass or stiffness matrices, and is consequently amenable to implementation on workstations.

Furthermore, the approach is promising as a potential parallel processing algorithm in that

(iv) The method exhibits a fine parallel granularity suitable for a wide collection of combinations of physical system topologies / computer architectures.

(v) The method is easily load balanced among processors, and does not rely upon system topology to induce parallelism.

INTRODUCTION

There is no doubt that an effective design process for the space station absolutely requires that high fidelity simulations of the transient response to control inputs be rapidly attainable. Much research has been carried out over the past few years that concentrates on improving the performance of methods for simulating the dynamics of nonlinear, multibody systems [4],[5],[14]. The research has primarily been devoted to

(i) the derivation of more efficient formulations of multibody dynamics, and to

(ii) the derivation of parallel processing algorithms.

Perhaps the most significant research addressing these two areas has been the introduction of the recursive, Order N algorithms in [6], and their subsequent refinements in [2], [14] for systems of rigid bodies. As noted in [14], these methods have the feature that the computational cost of the solution procedure is linear in the number of degrees of freedom N of the system, while conventional Lagrangian formulations are of cubic order . The conclusion that the Lagrangian methods are of cubic order derives from the fact that a system generalized mass/inertia matrix of dimension N X N must be factored at each time step. Just as importantly, the computational structure of the recursive Order N algorithms is amenable to parallel computation for some system topologies. If the system to be modelled has many independent branches in its system connectivity graph, the computational work required by the algorithm can be distributed among processors by assigning branches to independent processors. As an example, [2] considers the simulation of an all terrain vehicle. Because of the system connectivity and specific hardware architecture, excellent performance improvements and processor utilization are achieved in [2].

Due to these successes for rigid body simulations, it is well-known that many research institutions are presently investigating adaptations of the original recursive method to model systems comprised of flexible bodies. No doubt, the result will be highly efficient algorithms that perform well. Still, three key goals must be resolved before a general parallel processing algorithm can be obtained.

(i) Algorithms are required that induce parallelism on a finer scale, suitable for the emerging class of highly parallel processors.

(ii) Concurrent transient simulation methods must be automatically load balancing for a wider collection of combinations of mechanical systems and concurrent multiprocessing hardware.

(iii) The transient simulation method should also be amenable to vector processing implementation on each independent concurrent multiprocessor.

Based upon preliminary investigation, these goals should be very challenging if the algorithm is based upon an recursive Order N formulation.

An innovative strategy based upon these goals is derived in this paper. In part, its foundation can be traced to element-by-element methods already in use in finite element solution procedures [7]. As regards sequential computing environments:

(i) The combination range space formulation / PCG solution is an extremely efficient sequential algorithm for a class of problems described in the paper. The

efficiency is primarily due to the selection of a Block Jacobi preconditioner that is rapidly convergent.

(i) The method is non-assembling, i.e. , it does not require a large amount of in-core storage, and consequently is also attractive as a candidate for implementation on workstations.

(iii) Preliminary studies indicate that due to the rapid convergence achieved by using the selected preconditioner, the method can achieve performance rates that depend linearly upon the number of substructures.

Moreover, the method should be readily implemented on parallel processors:

(iii) A vast literature exists on the amenability of the PCG solution procedure to both concurrent and vector processing.

(iv) The method is relatively easily load balanced among processors, and does not rely upon system topology to induce parallelism.

This paper focuses on the fundamental dynamical formulation using a combination range space / PCG solution, and its performance on sequential computing machines. Although the potential application of the method on parallel architectures is outlined, the details of a concurrent implementation are presented in a forthcoming paper.

RANGE SPACE / PRECONDITIONED CG EQUATIONS

The range space formulation of dynamics has been derived in the aerospace and mechanism dynamics research literature in [1,11,12]. Its theoretical foundation can be traced to the range space formulation of constrained quadratic optimization [3]. Still, despite the fact that it is often less computationally expensive than the nullspace methods, the nullspace method seems to have received more attention in the literature [1,9,13,15...]. If the dynamics of a nonlinear, multibody system are governed by the collection of differential-algebraic equations

$$[M(q)]\ddot{q} = f(q,\dot{q},t) + [C(q)]^T \lambda$$

subject to constraints in linear, non-holonomic form

$$[C(q)]\dot{q} = 0$$

the range space solution of these equations are given by explicitly solving for the multipliers

$$\lambda = -([C(q)][M(q)]^{-1}[C(q)]^T)^{-1}\{[C(q)][M(q)]^{-1} f(q,\dot{q},t) - e(q,\dot{q},t)\}$$

and substituting to achieve a govering system of ordinary differential equations.

$$\ddot{q} = [M(q)]^{-1}\{f(q,\dot{q},t) - [C(q)]^T \{([C(q)][M(q)]^{-1}[C(q)]^T)^{-1}\{[C(q)][M(q)]^{-1} f(q,\dot{q},t) - e(q,\dot{q},t)\}\}$$

In the above equations, the constraints have been differentiated twice to yield

$$[C(q)]\ddot{q} = -\frac{d}{dt}([C(q)])\dot{q} = e(q, \dot{q}, t)$$

Any standard explict-predictor / implicit-corrector, or Runge-Kutta integration scheme can be applied to these equations provided that the condition number

$$\kappa(A) = \frac{\lambda_{max}(A)}{\lambda_{min}(A)}$$

of the constraint metric

$$[C(q)][M(q)]^{-1}[C(q)]^T$$

does not become too large. The restriction that the condition number above remains small precludes the possibility of redundant constraints (for example, as associated with singularities arising from closed loops) and remains an underlying assumption throughout the rest of the paper.

One advantage of the range space equations for systems having many independent structures to be assembled is that the system coefficient matrix is block diagonal and, consequently, the factorization and back-substitution required to form the product of the inverse of the mass matrix and a given vector is relatively inexpensive to calculate. It requires that one calculate the factorization of the individual substructure mass matrices alone. In fact, one need not even assemble the system mass matrix, and the factorizations can occur in parallel. Unfortunately, if one subdivides the overall system into finer collections of substructures (to facilitate the factorization of the system coefficient matrix), numerous constraints are introduced into the model.

The approach taken in this paper is to finely subdivide the system to be modelled, and thus accrue the benefits of having a system coefficient matrix with smaller block diagonals, but also employ a solution procedure that ameliorates the cost associated with the increasing dimensionality of the constraint metric. Specifically, the calculation of the Lagrange multipliers in

$$\lambda = -([C(q)][M(q)]^{-1}[C(q)]^T)^{-1}\{[C(q)][M(q)]^{-1}f(q, \dot{q}, t) - e(q, \dot{q}, t)\}$$

is carried out using the preconditioned conjugate gradient procedure.

THE PRECONDITIONED CONJUGATE GRADIENT SOLUTION

The preconditioned conjugate gradient procedure is an "accelerated" variant of the classical conjugate gradient procedure. If it is required to solve the linear system of equations

$$Ax = b$$

the procedure can be summarized as follows:

$x_0 = 0$

$r_0 = b$

For $k = 1, \ldots n$

  If $r_{k-1} = 0$

    then

      $x = x_{k-1}$

    else

      Solve $Qz_{k-1} = r_{k-1}$

      $\beta_k = z_{k-1}^T r_{k-1} / z_{k-2}^T r_{k-2}$

      $p_k = z_{k-1} + \beta_k p_{k-1}$

      $\alpha_k = z_{k-1}^T r_{k-1} / p_k^T A p_k$

      $x_k = x_{k-1} + \alpha_k p_k$

      $r_k = r_{k-1} - \alpha_k A p_k$

Careful inspection of the algorithm shows that the most computationally expensive tasks in the procedure are the

(i) calculation of the product of the coefficient matrix A and a given residual vector,

(ii) and the solution of a linear system of equations requiring the factorization of the preconditioner Q.

The rate of convergence of the preconditioned conjugate gradient algorithm is accelerated by employing a user-defined "preconditioning matrix." This matrix must have two properties to be an effective preconditioner:

(i) It must be relatively easy to factor.

(ii) It must be an approximate inverse to the constraint metric in a sense to be made precise below.

The reason for employing the preconditioned conjugate gradient solution method is that the convergence rate of the *conjugate gradient algorithm* (that is, with Q = I) is governed by

$$\frac{\|x - x_k\|_A}{\|x - x_0\|_A} \leq \left[\frac{1 - \sqrt{\kappa(A)}}{1 + \sqrt{\kappa(A)}}\right]^{2k}$$

Thus, the rate of convergence of the algorithm improves as the condition number

$\kappa(A)$

decreases. It has been shown in several publications that the convergence of the preconditioned conjugate gradient method is governed by the same expression, but with A replaced with

$$\tilde{A} = Q^{-\frac{1}{2}} A Q^{-\frac{1}{2}}$$

Clearly, if the preconditioner is identical to the coefficient matrix, then the condition number of $\tilde{A}$ is minimized. Hence, the preconditioner is sought such that its inverse approximates the inverse of the coefficient matrix. Many methods exist for the calculation of preconditioners . It should be noted that while the motivation for the use of many of these preconditioners is mathematically sound, the final choice invariably involves some heuristic.

THE CHOICE OF THE PRECONDITIONER

The choice of the preconditioner employed in this paper is based upon the following assumptions regarding the structural/mechanical system to be modelled:

(i) The system closely resembles a series of chains of bodies

(ii) The number of interface degrees of freedom is small relative to the number of interior degrees of freedom for a substructure.

(iii) The system does not contain any closed chains.

To a large extent, these assumptions have been driven by the physical structure of the space station in its assembly complete configuration.

The preconditioner for the system constraint metric is based upon the topology of a chain of substructures , such as those comprising the early configurations of the space station. If

$$[C_i(q)] \in R^{d_i \times N}$$

denotes the constraint matrix connecting two bodies at the ith interface, the system constraint matrix has the form

$$[C(q)] = \begin{bmatrix} [C_1] \\ [C_2] \\ \vdots \\ [C_k] \end{bmatrix} \in R^{D \times N}$$

The system constraint metric can then be written

$$\begin{bmatrix} [C_1][M]^{-1}[C_1]^T & \cdots & [C_1][M]^{-1}[C_k]^T \\ [C_2][M]^{-1}[C_1]^T & & \\ [C_3][M]^{-1}[C_1]^T & & \\ \vdots & & \\ [C_k][M]^{-1}[C_1]^T & \cdots & [C_k][M]^{-1}[C_k]^T \end{bmatrix}$$

Based upon the structure of the constraint metric above, the preconditioner is selected to be the block diagonal matrix

$$\begin{bmatrix} [C_1][M]^{-1}[C_1]^T & & & \\ & [C_2][M]^{-1}[C_2]^T & & \\ & & & \\ & & & [C_k][M]^{-1}[C_k]^T \end{bmatrix}$$

Although the off-diagonal blocks

$$[C_i(q)][M(q)]^{-1}[C_j(q)]^T = 0$$

(for i not equal to j) are not generally identically equal to zero, this choice of preconditioner is shown to be extremely efficient for the class of problems described in the next section. Furthermore, this preconditioner satisfies the two essential criteria of good preconditioners:

(i) It is block diagonal, with small diagonal blocks, and is relatively easy to factor.

(ii) It has an inverse that provides a good approximation to the inverse of the full system coefficient matrix.

This latter conclusion results from the well-known fact [16] that the directed graph representing the connectivity of an open loop system can be regularly ordered. The regular ordering results in a system constraint metric that has a reduced bandwidth. That is, many of the off-diagonal blocks

$$[C_i(q)][M(q)]^{-1}[C_j(q)]^T = 0$$

are identically zero for i>>j. The choice of preconditioner shown above is often denoted the Block Jacobi preconditioner and is known to be highly effective for classes of systems of equations arising from elliptic partial differential equations.

SPACE STRUCTURE SIMULATIONS

Numerous simulations have been carried out to verify the attributes of the algorithm cited earlier. In this section, only two simulations are described. More detailed simulations can be obtained from a forthcoming presentation at the 32nd Structures, Dynamics and Materials Conference.

Figure (1) depicts a space mast simulation in which z-truss substructures having 63 degrees of freedom each are assembled end to end. As shown in figure (2), the preconditioned conjugate gradient method is rapidly convergent using the block Jacobi preconditioner derived from the constraint metric. In fact, the number of iterations required for convergence remains constant independent of the number of degrees of freedom. Figure (3) shows that the method is indeed Order N in computational cost in that the total time per integration time step increases linearly as a number of the degrees of freedom.

A second example simulation of the space station in its permanently manned capability is depicted in figures (4) through (6). Figure (4) depicts the components comprising the station and their relative positions in the overall assembly. Figures (5) and (6) summarize the performance of the algorithm for the entire assembly. Figure (5) plots the total time per integration time step evaluation versus the number of degrees of freedom. As demonstrated earlier, the simulation time does grow as a linear function of the total number of degrees of freedom. Figure (5) illustrates the important fact that the primary computational cost of the algorithm is associated with

(i) system coefficient matrix multiplication and

(ii) preconditioner application,

both of which are trivially parallelizable. Figure (6) shows that the number of PCG iterations required for convergence is nearly independent of the number of substructures/dof as in the previous case.

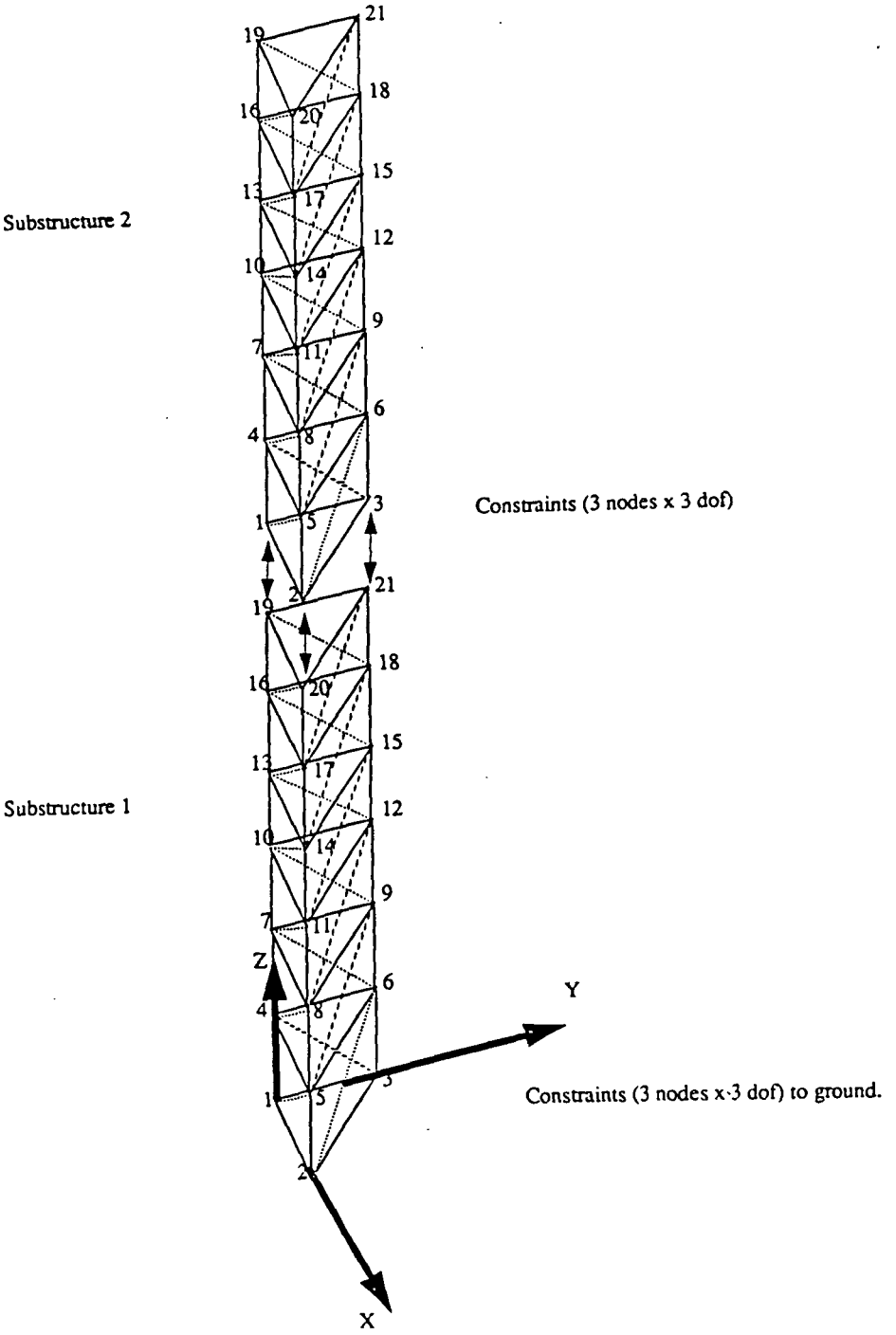


Figure 1.- Z Truss / Space Mast Assembly.

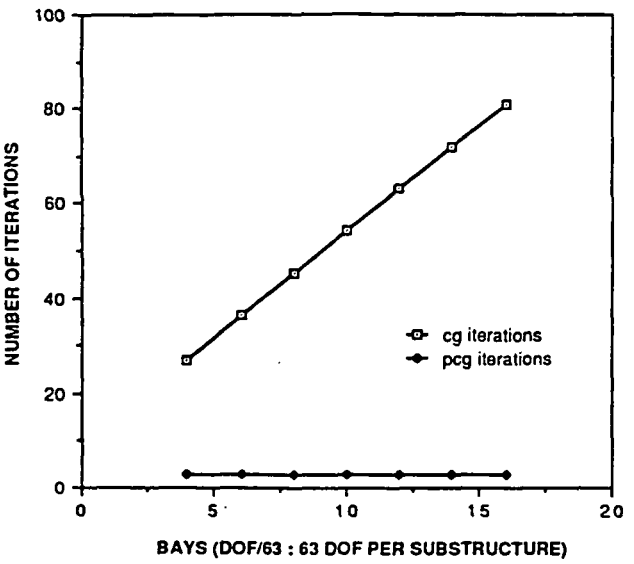


Figure 2.- Iterations Versus Degrees of Freedom for Z-Truss.

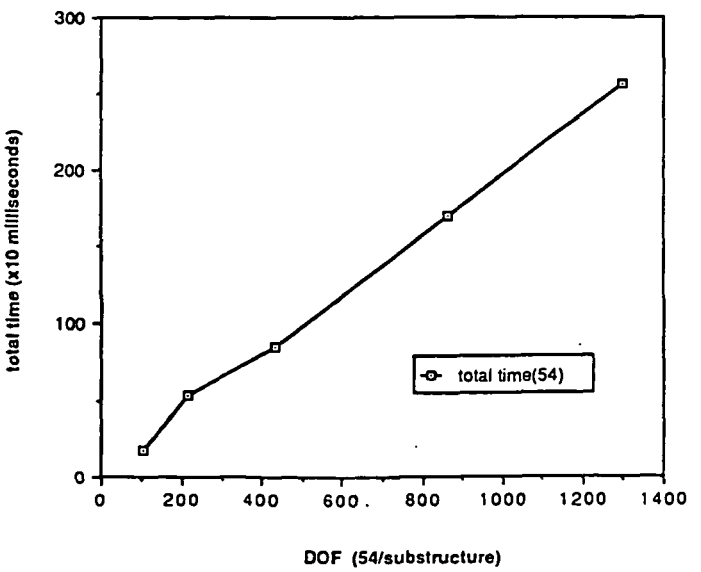


Figure 3.- Computational Cost per Integration Step for Z-Truss.

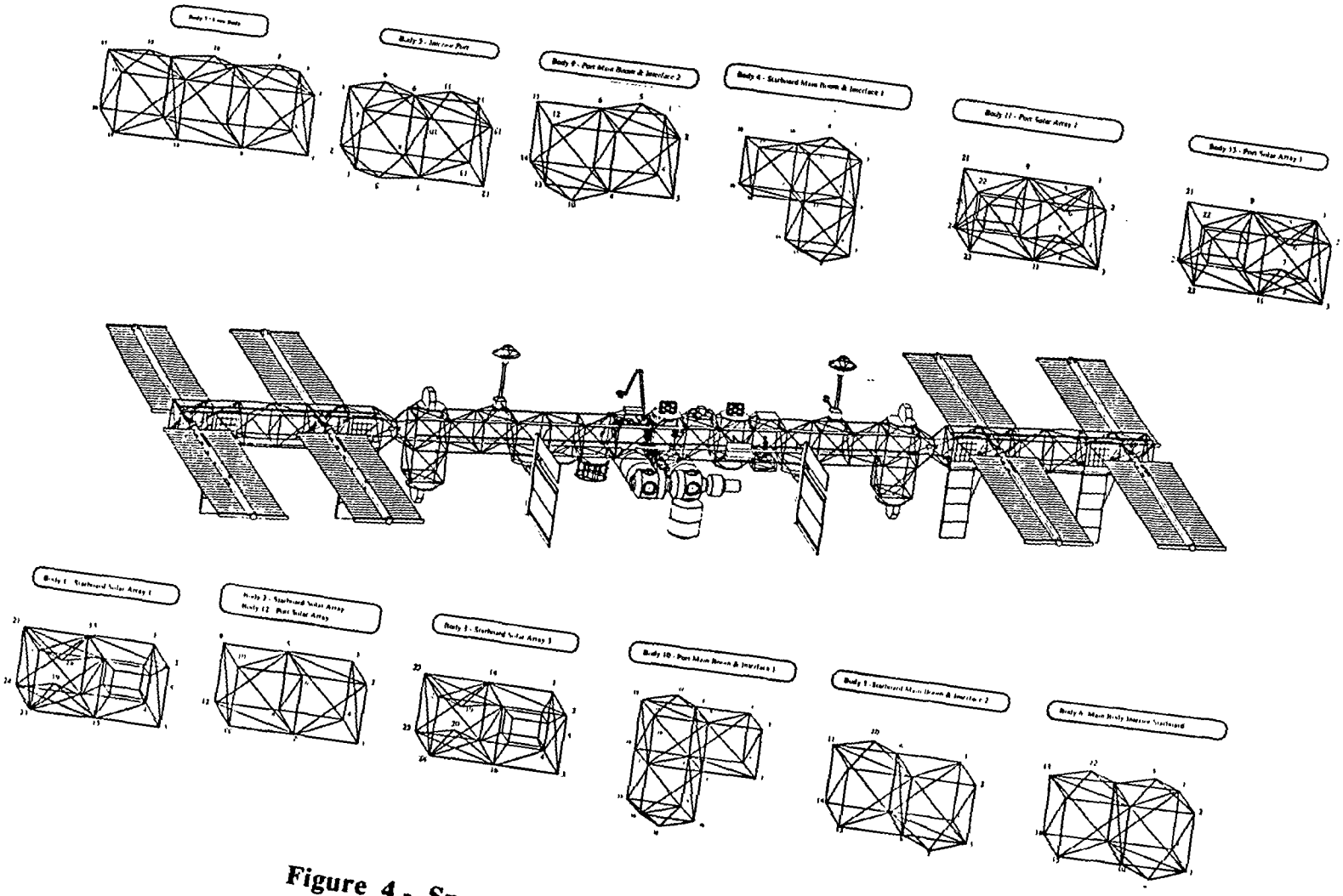


Figure 4.- Space Station 13 Substructure Assembly.

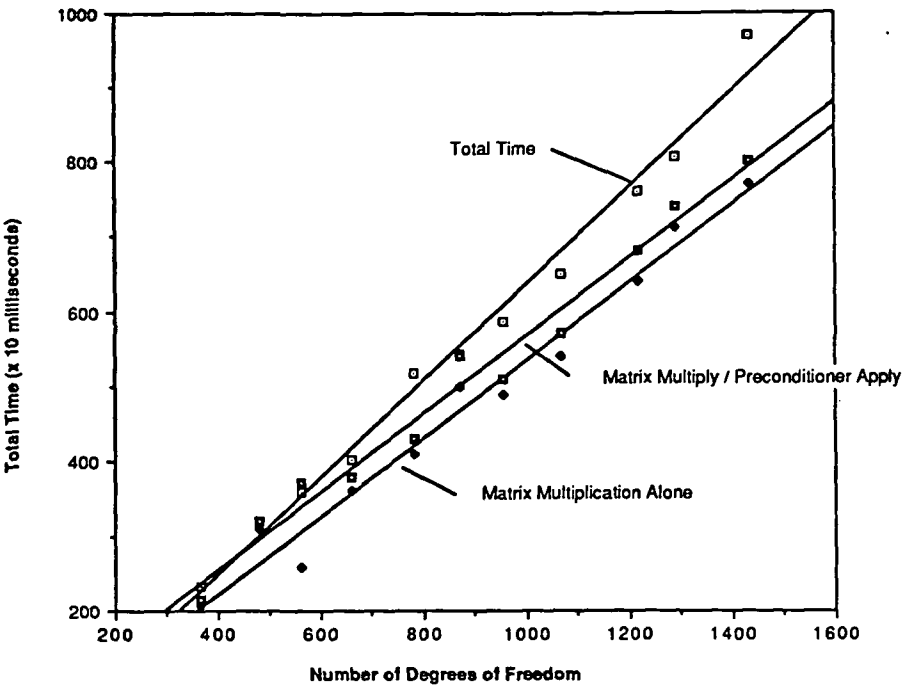


Figure 5.- Computational Cost per Integration Step for Space Station

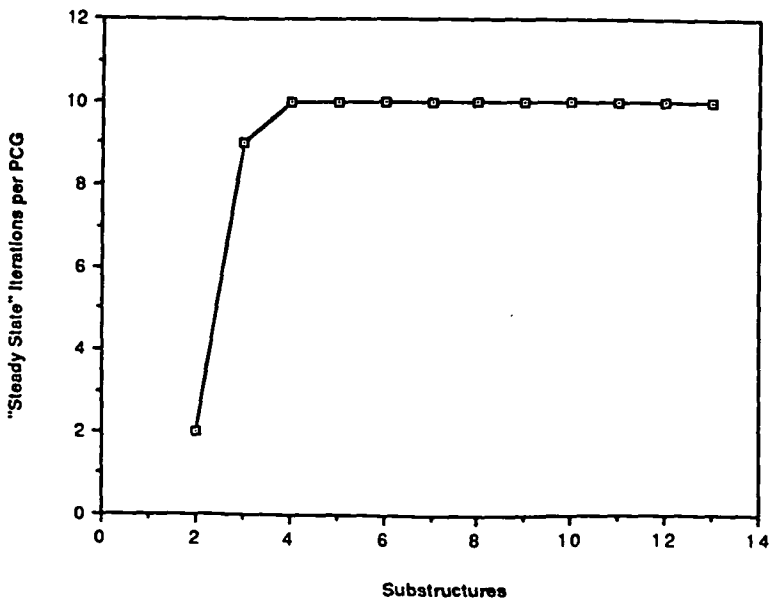


Figure 6. - Iterations for PCG Convergence Versus Number of Substructures

CONCLUSIONS / RECOMMENDATIONS FOR FUTURE WORK

The primary conclusions of this report can be summarized as follows:

(I) Although the recursive, Order N multibody dynamics formulations can yield excellent performance in many simulations, they are not a panacea as regards applications to all classes of problems in multibody dynamics.

(II) It is distinctly counterproductive to limit research to recursive, Order N methods. The order N algorithm traces its roots to simulation methods for low dimensionality robotic simulations. The benefits of the method for simulating systems with thousands of degrees of freedom, such as the space station, have yet to be firmly established.

(III) There are three reasons why alternative formulations to the recursive Order N algorithm should be pursued:

(i) trends in parallel computing architectures
(ii) underutilization of numerous concurrent multiprocessors
(iii) difficulties in load balancing

(IV) An alternative nonrecursive, Order N algorithm has been presented in this report that has many advantages for a sequential computing environment:

(i) It is rapidly convergent.
(ii) It can achieve Order N computational cost.
(iii) It is non-assembling.

In addition, the method addresses the issues noted above for a parallel computing implementation.

(i) It exhibits a fine parallel granularity suitable for emerging computer architectures.
(ii) The method is easily load balanced.
(iii) A vast literature exists on parallel preconditioned conjugate gradient methods.

The research described in this report has provided a promising new avenue for further research. In particular, the range space/PCG formulation of multibody dynamics should undergo further research, to be carried out in three primary phases:

(i) The extremely promising computational cost estimates for concurrent multiprocessing should be validated by implementing the method for linear simulations of the space station. The class of potential concurrent multiprocessing architectures could include

BBN Butterfly	32 processors
N-Cube	32+processors
Capps 8064	32 processors

(ii) While research in this report has been concerned with the feasibility of an alternative concurrent method, the work has been limited to linear systems. The formulation should be extended to include nonlinear, multibody effects

(iii) The resulting linear/nonlinear simulation capability should be incorporated in a controls design procedure package for carrying out the tasks of computational control and control design required for the development of attitude control systems for the space station.

REFRENCES

[1] Agrawal, O. P., "Dynamic Analysis of Multi-body Systems Using Tangent Coordinates", In E. Bautista, J. Garcia-Lomas, and A. Navarro, editors, *The Theory of Machines and Mechanisms*, pp. 533-536, Pergamon Press, Oxford, 1987.

[2] Bae, D. and Haug, E. J., "A Recursive Formulation for Constrained Mechanical System Dynamics: Parts i, ii, iii. *Mechanics of Structures and Machines*, 1988. To appear.

[3] Gill, P. E., Murray, W. and Wright, M. H., *Practical Optimization*, Academic Press, London, 1981.

[4] Gluck, R., "Hope for Stimulating Flexible Spacecraft", *Aerospace America*, pp. 40-44, November 1986.

[5] Haug, E. J., "Elements and Methods of Computational Dynamics", In E. J. Haug, editor, *Computer Aided Analysis and Optimization of Mechanical System Dynamics*, pp. 3-37, Springer-Verlag, Berlin, 1984.

[6] Hollerbach, J. M., "A Recursive Lagrangian Formulation of Manipulator Dynamics and a Comparative Study of Dynamics Formulation Complexity", *IEEE Trans. on Systems, Man, and Cybernetics*, 10:730-736, 1980.

[7] Hughes, T. J. R., *The Finite Element Method*, Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

[8] Huston, R. L., "Useful Procedures in Multibody Dynamics", In G. Bianchi, editor, *Dynamics of Multibody Systems*, pp. 69-78, Springer-Verlag, Berlin, 1986.

[9] Kim, S. S., and Vandorploeg, M. J., "QR Decomposition for State Space representation of Constrained Mechanical Dynamic Systems", *Journal of Mechanisms Transmissions, and Automation in Design*, 108:103-188, June 1986.

[10] Kurdila, A., Kamat, M. P., "Concurrent Multiprocessing Methods for Calculating Nullspace and Range Space Bases for Multibody Simulation", AIAA Journal, April 1990, accepted for publication.

[11] Kurdila, A., Papastavridis, J. G. and Kamat, M. P., "On the Role of Maggi's Equations in Computational Methods for Constrained Dynamical Systems", AIAA Journal of Guidance, Control and Dynamics, Vol. 13, No. 1, Jan. 1990, pp. 113-120.

[12] Placek, B., "Contribution to the Solution of the Equations of Motion of the Discrete Dynamical System with Holonomic Constraints", In E. Bautisa, J. Garcia-Lomas, and A. Navarro editors, *The Theory of Machines and Mechanisms*, pp. 379-382, Pregamon Press,

Oxford, 1987.

[13] Singh, R. P., and Likins, P. W., "Singular Value Decomposition for Constrained Dynamical Systems, *Journal of applied Mechanics*, 52:943-948, Dec. 1985.

[14] Singh, R. P., Schubele, B. and Sunkel, J. W., *Efficient Algorithm for the Dynamics of Multi-Link Mechanisms*, AIAA Paper No. 89-3527-CP, 1989.

[15] Webage, R. A., and Hung, E. J., "Generalized Coordinate Partitioning for Dimension Reduction in Analysis of Constrained Dynamic Systems", *ASME Journal of Mechanical Design*, 104:242-255, January 1982.

[16] Wittenburg, J., *Dynamics of Systems of Rigid Bodies*, B. G. Teubner, Stuttgart, 1977.

APPLICATIONS OF FORMAL SIMULATION LANGUAGES IN THE CONTROL AND MONITORING SUBSYSTEM OF SPACE STATION FREEDOM

Final Report

NASA/ASEE Summer Faculty Fellowship Program - 1990

Johnson Space Center

Prepared By:	R. C. Lacovara, Ph.D.
Academic Rank:	Assistant Professor
University & Department	Stevens Institute of Technology Dept. of Electrical Engineering and Computer Science
NASA/JSC	
Directorate:	Engineering
Division:	Tracking and Communications
Branch:	Communications Performance and Integration
JSC Colleague	James C. Dallas
Date Submitted	August 17, 1990
Contract Number	NGT-44-005-803

1. Abstract

The notions, benefits, and drawbacks of numeric simulation are introduced. Two formal simulation languages, Simscript and Modsim are introduced. The capabilities of each are discussed briefly, and then the two programs are compared.

The use of simulation in the process of design engineering for the Control and Monitoring System for Space Station Freedom is discussed. The application of the formal simulation languages to the CMS design is presented, and recommendations are made as to their use.

2. Simulation and the Scope of this Paper

This paper discusses certain formal simulation techniques and tools, and makes observations and recommendations on the use of these techniques in a specific environment: the Control and Monitoring Subsystem (CMS) for Space Station Freedom.

The concept of simulation and varieties of simulation are discussed briefly. After that discussion, the application of simulation to the CMS is presented. The formal languages are presented and compared. Lastly, some recommendations for the use of these tools are advanced.

2.1. Rationale for Simulation

Simulation is one of many methods of obtaining information about physical or conceptual systems. The chief feature of simulation is that the information is not obtained by methods of formal analysis. Instead, a model of the system which includes the pertinent behavior is constructed and the model is exercised. Possibly, the model is exercised using pseudo-random input conditions, and the model may be exercised many times to smooth out variations in results due to the pseudo-random inputs, or internal psuedo-random events.

(In this paper, the term "random" will be used to describe the behavior of pseudo-random sequence generators. The differences and implications of this choice, and the definition of pseudo-random are beyond the scope of this work.)

The chief advantage of simulation accrues from the chief feature. Simulation is an interesting methodology precisely when (a) no formal analysis is possible or (b) no physical system is available from which to make observations. Any number of reasons will occur to the reader which may cause case (b). Case (a), the lack of formal analysis, occurs whenever the system to be studied has no known system of equations which describe it, or when the complexity, inherent non-linearities, pathologies, or plain old intractability of the system make the search for analytical description impractical. A cold appraisal of the world leads workers to conclude that case (a) is the rule, not the exception.

In brief, simulation is applied to problems for which formal analysis is unavailable or for which no physical system is available for observation.

3. Categorization of Simulation

Simulations may be usefully categorized in several ways. The process domain of a simulations refers to the description of time in which the system operates. Discrete time and continuous time descriptions will be familiar to most electrical engineers.

The computational domain describes the facility used to model the system. Although most facilities are digital computers, some analog computers are used in simulations.

The simulation scheme refers to the method in which the simulation passes time. This may be synchronously, in which the program advances time in small quanta and determines what, if anything, should occur. Otherwise, the simulation might be asynchronous, in which the program maintains a list of scheduled events and proceeds directly from one scheduled event to another regardless of the intervening time.

3.1. Process Domain

Process domain refers to the basic view of time taken by the simulation. A natural view of the progression of time is that of continuous time. The time variable may take on any value, and this is certainly the most common view of processes such as ballistic bodies, electronic circuits and similar systems. The other common description of the passage of time is discrete time, in which the time as an independent variable takes on only specific values, usually of the form

$$t = nT$$

where T is the smallest identifiable duration, and n is the actual independent variable.

The difference between continuous and discrete time signals is more fundamental however. Continuous time systems are those which may be described by differential equations, whereas discrete time systems are described by difference equations. Analytic solutions to continuous time systems are therefore the solutions to ordinary and partial differential equations with boundary conditions, usually facilitated by the use of Laplace transforms. Analytic solutions for discrete time systems satisfy difference equations with forcing functions, and a common technique is that of z-transforms.

Often it is convenient to model a naturally continuous system with a discrete time model. This may entail a choice of T such that no significant events occur in a period of time less than 2T. Unfortunately, it is not always clear what constitutes a "significant event", and further, other subtleties (such as solution stability concerns) may intrude.

As a practical matter, most formal simulations are in fact discrete time. Philosophical arguments aside, digital computers (the most common vehicle of simulations) represent quantities discretely (in finite precision.)

3.2. Computational Domain

Calculations may be performed in many different ways. Mechanical and hydraulic systems have been designed to perform logic and arithmetic. The only interest here is in electronic methods.

3.2.1. Analog Computers

Analog "computers" are good examples of analog simulators. These devices are actually arrays of operational amplifiers configured as summers, integrators and differentiators. As a result, analog computers are well suited to the solution of systems of linear differential equations. As an illustration, note that voltage or current is made the analogue of some physical quantity, and time is made the analogue of the independent variable, usually time. This is in contrast to something like a phonograph recording, in which vertical mechanical displacement of a grove is made the analogue of acoustic pressure, and linear displacement along the groove is made the analogue of time.

Analog computers are most often set up on plugboards to represent a particular differential equation. An applicable forcing function is applied, and the behavior of the system as a function of time is observed and scaled.

3.2.2. Digital Computers

Digital computers are the most common vehicle for calculation and simulation. Inherently, these machines represent discrete variables, and are naturally suited to represent systems which are characterized by difference equations. Various strategies allow the system to represent continuous time systems with acceptable results.

Digital "computers" could be constructed in the same manner as analog computers. Collections of adders, subtractors and delays could be assembled to exactly represent some particular difference equation. This is uncommon, although such an arrangement might be one of the fastest methods of obtaining a solution to a set of equations.

A slightly less hardware-intensive method is to use digital signal processing chips to implement the difference equations. These single-chip computers use specialized architectures and instruction sets to efficiently evaluate difference equations.

The majority of computational applications are performed on general-purpose computer architectures. This includes formal simulation languages.

3.3. Simulation Methodology

There are two primary simulation methods. These differ in the manner in which simulated time is advanced. An important distinction is that simulation time is unrelated to run time, or actual time. Simulation time is advanced either by identical small quanta (synchronous simulation) or advanced from event to event listed in a process list (asynchronous simulation.)

3.3.1. Synchronous Simulation

Synchronous simulations are used in systems which require interaction with exterior hardware. Examples are systems which are interfaced to hardware or certain recording systems. These systems have stringent timing requirements: the simulation must be able to advance simulation time at least as fast as real time.

Synchronous simulations are common even when there are no exterior requirements. This is not surprising, since synchronous simulations are often easier to code than asynchronous simulations.

Coding a synchronous simulation proceeds as follows. Prior to coding, the analyst determines the largest time slice which will suit the problem. An initial state of the problem is chosen. Then, for each element of the system, the analyst computes (guesses?) the probability (formally, a transition probability) that a particular element of the simulation changes state in during a time slice.

Code is written which can store the state of the system during any particular time slice, and update the state during the next slice. This is done by making a number of draws from a suitable probability distribution, and mapping the results into the new state of the system.

This technique has the advantage of simplicity. If the code is hosted on a sufficiently fast platform, the simulation is by default "real-time". (For example, if the time quanta chosen is 0.1 second, and the program completes all transition probability draws in less the 0.1 second, then the program is not a rate limiting part of a larger system, hence is real-time.)

There are drawbacks to this arrangement. If real-time performance is required, and the code does not run fast enough, then there is little recourse except a full re-write, or a port to a faster host. Further, a view of a process from the level of the time slice makes is difficult to implement a "process-wide" view of the system. Interactions between processes in the simulation are not obvious, and not necessarily easy to implement.

For certain classes of problems, the synchronous approach is sparing of a programmer's time, and may allow real-time operation in some cases. A discussion and exposition of a non-trivial synchronous simulation is found in [Lacovara 87].

3.3.2. Asynchronous Simulation

For more complex systems, or systems in which there is significant interaction between parts of the system, synchronous simulations present non-trivial coding problems. Further, the size of the time slice determines to a great extent the rate at which the simulation runs. A choice of small time slices will slow the system, and many slices may pass without any activity. Large time slices will allow simulated time to pass faster, but large slices may not allow sufficient fidelity in the simulated world.

Asynchronous simulation is an approach which resolves the time slice dilemma. In an asynchronous simulation, the time of occurrence of an event is predicted by a draw from an appropriate distribution. The predicted event is then placed as an event notice in a linked list of events, often called an event list. After all predictions which can be made from the present state and simulation time of the system have been make, simulation time is advanced to the

event notice closest in simulation time. Any changes needed to the state of the system implied by the current event are made, any new event notices are posted to the event list, and simulation time is advanced again.

The principal feature of asynchronous simulation is that simulation time advances from event to event. No compute time is spent evaluating time slices in which no events occur, and the "view" of the simulation is at the process level.

The advantages of asynchronous simulation are often great. A more intuitive view of the system may be use to create, maintain, and interpret the simulation. The simulation will permit interaction between its component processes. The disadvantage, however, is that the creation and maintenance of event lists and process notices are non-trivial, and certainly not advised for casual, even if experienced, programmers.

The formal simulation tools discussed below have the great advantage that the mechanics of the simulation, linked lists, queues, priority chains and so on, are hidden from the programmer. With the use of these tools, powerful simulation constructs are available to users who do not plan to spend their entire career writing code.

4. Formal Languages: Simscript and Modsim

In this study, two formal asynchronous simulation tools produced by CACI Products Company of La Jolla, California were used. These languages are Simscript and Modsim. They are entirely different approaches to the problem of simulation, and have overlapping but not identical domains of applicability.

4.1. Simscript

Simscript (from simulation script) is an old (1962) hence mature product. It imposes on the programmer a "view" of a system divided into processes. This is not altogether artificial, as the Simscript model of a bank consists of processes which represent a) customers, b) tellers, and c) customer arrival generator.

Simscript maintains a complex internal system of queues, lists, and other data structures. The maintenance of these structures is transparent to the programmer for the most part. In the bank example, customer arrivals are

"generated" by introducing event notices of arrival in an event list. The bank teller is represented by a teller "resource", a process which incorporates a waiting queue and other flags to indicate teller status. When a customer arrives at a teller, the Simscript system determines whether the teller is "busy" with a previous customer. If busy, the customer is placed in a waiting queue until the teller is free to process the customer.

Simscript manages the teller's waiting queue transparently. It checks the teller's status flags and determines whether the customer may be served or must wait. In addition to manipulation of these internal structures, Simscript monitors selected parameters of the simulation automatically. Quantities such as queue size, maximum, minimum, average and other statistics are accumulated without explicit programmer intervention.

As a result of the internal features, a certain class of simulation may be implemented in Simscript in an elegant and straightforward fashion. From a descriptive point of view, Simscript is well suited to systems which involve queuing theory: processes in which requests for service contend for limited resources.

Simscript includes facilities for screen graphics. Pre-defined screen constructs include graphs, indicators and clocks. Moving icons may be designed and animated by the simulation. These are relatively straightforward to use, and seem to be quite useful. Some are quite impressive.

4.2. Modsim

Modsim is a relatively recent language product. It is syntactically based on Modula 2, which is similar to Pascal. Basically, Modsim adds formal object-oriented structures to Modula-2, and provides simulation capability by providing a library of objects which can pass simulated time. Due to the Modula 2 underpinning, Modsim has considerable general purpose characteristics. It would be perfectly feasible to use Modsim wherever Modula 2 or Pascal is employed.

A description of object-oriented programming is beyond the scope of this paper. However, it is necessary to note that the pertinent features of "objects" are these. Objects are data structures comprised of typed-fields, and a list of procedures (called methods) which are the only procedures which act

on the data structure. The implications are these: objects may interact only through a object's allowed methods. An object's external fields may be read by other objects, its internal fields remain invisible. An object's fields may not be modified directly, but only through the agency of its allowed methods. The purpose of this strict control of access is to impose a discipline on the construction of the program, and to provide a type of safety in the control of actions performed on data structures.

Modsim extends the notions of object-oriented programming in the following manner. Methods are available in two classes, "ASK" and "TELL". An ASK method corresponds to the original methods noted above. When invoked upon an object, some action ensues in instantaneous simulation time. TELL methods, however, are asynchronous: simulation time may elapse before the desired changes occur. This is accomplished by the mechanism of an event-list. As in Simscript, the control and maintenance of the event list for the asynchronous simulation aspects of Modsim are implicit.

Modsim provides a complex and general programming environment. (Translation: Modsim is powerful, difficult, and sometimes obscure.) It would appear to be suitable for a very wide range of simulations. Like Simscript, it contains provisions for complex screen graphics.

4.3. Comparison of Simscript and Modsim

Be advised that the author's programming experience is as follows: considerable expertise in assembly languages, Fortran, Pascal, and C, and operating systems. The author has written numeric simulations on several systems, but is new to Simscript, Modsim, and object-oriented programming.

It seems to be easier to write simulations in Simscript than Modsim. In many ways, Simscript seems to bridge the conceptual gap between the system under consideration and the programming model in a more direct fashion. Modsim's object constructs however correspond to identifiable processes in the real world. Modsim's power and complexity initially interfere with the process of writing code: Modsim requires a substantial amount of "overhead" code to just get started.

The author designed a simple multiplexer as a simulation example for both Modsim and Simscript. The multiplexer accepts "data packets" about every 75

milliseconds. The packets face two servers. One is a high speed server, the other is a low speed server. The multiplexer enforces the following service discipline. If the queue size for the low speed server is twenty-five or fewer packets, the incoming packet joins the low speed queue. Otherwise, the packet joins the high speed queue. The simulation accumulates statistics on the sizes of the queues, and the distribution of the time taken by every packet to traverse the system.

The Simscript version of this code is about 40 lines. These 40 lines compile into about 250K bytes for a Sun Sparcstation. The Modsim version is only about 50% larger, but compiles into about 1M bytes for the same machine. Modification of the server discipline is about the same difficulty in either language. As features are added to either version, the Simscript object file grows somewhat, but the Modsim does not seem to expand much at all. The simplest interpretation of this behavior is that the Simscript object file is growing proportionally to the additional lines of code. The Modsim compiler seems to be able to perform one of the selling points of object-oriented programming: the ability to reuse many parts of the code through inheritance and recursion.

Of course, the Modsim object starts out about four times the size of the Simscript model, but this is not really serious on large machines.

Both Modsim and Simscript seem to utilize whatever processor power is available, as judged by a Sun performance meter.

The screen graphics package for Modsim is very similar to that of Simscript, but available in a more advanced version, and is therefore preferable.

Overall, Modsim seems to be a more general and versatile simulation system than Simscript. It carries with it corresponding penalties is size and programmer learning curve. Simscript is almost as versatile, and seems to be a very good choice if only one package is to be available to the general analyst working with small to medium sized systems. For very large projects, Modsim may be a better choice, since in large systems the organizational advantages of object-oriented programming will begin to be felt.

5. Simulation of Control and Monitoring Subsystem

The Control and Monitoring Subsystem (CMS) for Space Station Freedom is a collection computers and busses which determine the configuration and operate

various hardware systems. Some tasks include health observation, fault diagnosis and recovery, and other system-wide tasks.

5.1. Functional Simulation

The process of design of this system includes several generations of simulation of the CMS. However, these simulations are not operational, but functional. As an example, other applications on the Space Station communicate with the CMS via a local area network. The present simulations of CMS accept these commands, and provide a simulated response. This simulation system does not depend heavily on stochastic processes. Instead, the simulation must "merely" accept communication from exterior agencies and provide a sensible response. As a result the current simulations are written primarily in ADA and C.

5.2. Use of Operational Simulations in the CMS

There are several areas in which operational simulations may be of use in the CMS scheme. These are primarily incidental, but of some interest.

The current fault generation in the CMS simulator is manual. If there is no automatic fault diagnosis and isolation, this is probably adequate. If automatic detection, diagnosis, isolation, and correction are to be implemented, good testing practice would dictate that the simulated flaws occur without warning to the simulation operators.

Some aspects of the simulation are time-varying. Some require knowledge of the state of other components of the simulation. Either simulation tool would be of some use here. If complex time-dependent behavior was required, the tools would have some advantages over C or ADA.

5.3. Disadvantages in the CMS Simulation Environment

Simscript and Modsim have common properties which do not fit well with current simulation and software philosophy on Space Station. Neither tool resembles ADA, is written in ADA, or is likely to be available in ADA in the foreseeable future. An ADA ideology permeates the Space Station software engineering efforts, possibly to the exclusion of other useful concepts and directions in computer science.

A secondary item of dogma is object-oriented programming. Modsim fills this

bill nicely, and brings with it all of the overhead and performance penalties built in to this paradigm. However, Simscript is not strictly object-oriented. Simscript is modular, but it retains some characteristics from its early origins. (A new line is read from a disk file by the keywords "START NEW CARD".) The disadvantages, quirks and structural awkwardness of either program are far outweighed by the advantages they bring to simulation.

6. Recommendations

In the context of the current software requirements of EE7 and in support of engineering efforts for Space Station Freedom, I offer the following recommendations.

a) One civil service employee of the branch should learn Simscript or Modsim well enough to write a small, but non-trivial simulation. This will bring some experience in formal simulation tools into the branch. I suspect that this will pay off rather sooner than later, as requirements and tasks evolve.

b) A portion of the CMS demo which could use graphics output might be identified and coded in Modsim. Modsim has convenient and useful graphics capability, and it is possible that it could be the most efficient means of generating many graphics displays.

c) Some demonstration of Simscript's or Modsim's capabilities should be shown to staff involved in analysis tasks in other branches. These tools are potentially most useful to people who write operational simulations or perform formal analysis. They should know that the tools exist on-site.

7. Reference

Lacovara 87

R. C. Lacovara, Dissertation: Adapted Packet Speech Interpolation, Stevens Institute of Technology, University Microfilm 88-17309 Issue 4907 DAI January 1989.

A STUDY OF MURINE BONE MARROW CELLS CULTURED IN BIOREACTORS WHICH CREATE AN ENVIRONMENT WHICH SIMULATED MICROGRAVITY

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1990

Johnson Space Center

Prepared By:	Brother DeSales Lawless, Ph.D.
Academic Rank:	Associate Professor
University and Department	Fordham University Division of Science and Mathematics New York, N.Y., 10023
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	MEDICAL SCIENCES
Branch:	Biomedical Operations and Research
JSC Colleague:	Steve Gonda, Ph.D.
Date Submitted:	August 2, 1990
Contract Number:	NGT 44-005-803

ABSTRACT

This study is a continuation of work undertaken at Johnson Space Center in the summer, 1989. That research indicated that mouse bone marrow cells could be grown in conditions of simulated microgravity. This environment was created in rotating bioreactor vessels, which were fabricated by the Biotechnology Instrument Laboratories of the Johnson Space Center.

On three attempts mouse bone marrow cells were grown successfully in the vessels. The cells reached a stage where the concentrations were doubling daily. Phenotypic analysis using a panel of monoclonal antibodies indicated that the cell were hematopoietic pluripotent stem cells. They did not show surface markers characteristic of differentiated myeloid, lymphoid or erythroid cells. Furthermore, they stained positively for an antibody which has been shown to be present on stem cells. These cells formed colonies when grown on soft agar. Morphologically, they appeared clonic. One unsuccessful attempt was made to reestablish the immune system in immunocompromised mice using these cells.

Since last summer, several unsuccessful attempts were made to duplicate these results . It was determined by electron microscopy that the cells successfully grown in 1989 contained virus particles. It has been suggested that these virally parasitized cells had been immortalized,

The work of this summer is a continuation of efforts to grow mouse bone marrow in these vessels. A number of variations of the protocol have been introduced. Certified pathogen free mice have been used in the repeat experiments. In some attempts the medium of last summer was used; in others Dexture Culture Medium containing Iscove's Medium supplemented with 20% horse serum and 10-6 M hydrocortisone. The Dexter medium proved to be unsuitable for growth in the vessels but was excellent for growth of static cultures at unit gravity.

Our efforts this summer were directed solely to repeating the work of last summer. We had planned many investigation if we succeeded in isolating the stem cells. We would attempt immortalization of the undifferentiated stem cell by transfection with oncogenic vector. We would induce selective differentiation in the stem cell line by growing it with known growth factors and immune response modulators. We were particularly interested in identifying any surface antigens unique to stem cells that would help in their characterization.

Another goal was to search for markers on stem cells that would distinguish them from stem cells commited to a particular lineage; these are the cells referred to as precursor cells. We believe that the stem cell is self-renewing while the precursor cell obeys the restrictions proposed by Leonard Hayflick (Hayflick, L., 1965).

If we obtained the undifferentiated hematopoietic stem cell we would study the pathways that would terminally convert it to myeloid, lymphoid, erythroid or other cell line. We would like to transfect it with a known gene and then convert it to a terminally identifiable cell.

INTRODUCTION

This research is a study of murine bone marrow cells cultured in simulated microgravity. It has been shown by several investigators that gravity is an environmental factor which affects growth and functionality of cells (Lorenzi, G., 1986). When flown on a space mission where they encountered zero gravity, human lymphocytes showed less than 3% of mitogenic activation to Concanavalin A as similar cells which had been cultured at unit gravity (Cogoli, 1985). It has been further demonstrated that microgravity depresses and hypergravity enhances cell proliferation rates. These effects are particularly strong in cells which are undergoing differentiation. The cellular proliferation rates of several different cell types were found to increase by 30% in hypergravity while the consumption per cell of glucose was lower than at unity gravity (Tschopp, 1983). A summary of the effects of spaceflight on single cell organisms, plant and mammalian cells has appeared in the literature. (Gmunder, F.K. and A. Cogoli, 1988). The effects of spaceflight on levels and activity of immune cells has been reported by Sonnenfeld, et al (1990).

We have undertaken this research encouraged by the work of these investigators and others who have demonstrated that gravity affects cell behavior. Our particular interest in bone marrow is associated with its importance in clinical medicine and several areas of basic research. It is important that the stem cell be isolated and characterized. The investigation of hematopoiesis is limited by the by the lack of a definitive in vitro assay for the most primitive hematopoietic stem cell. (Rowley, et al, 1987). No antigenic determinant has been found to date that is specific for mouse stem cells. The antigens that are positive for stem cells are not specific. (Van de Rijn, et al., 1989) It is estimated that only 0.02% of bone marrow cells are undifferentiated stem cells. It is postulated that these cells are the only ones with self-renewal potentialities. (Spangrude, et al., 1988).

Among the premises upon which this research is predicated are: (a) cell-cell contacts would be minimized in the bioreactors; (b) cell differentiations would be curtailed since the adherent cells that make up the microenvironment would would not thrive in the vessels; (c) cytokine communications could be lessened by periodic changes of the medium; (d) stem cells, unlike differentiated cells, are self-renewing. Stem cells would outlast precursor cells. We consider the precursor cells those that have been committed to a specific differentiation.

Among approaches investigators have used in their quest for the hematopoietic stem cells in bone marrow are: density gradient centrifugation (Yoichi Takave, 1989), elutriation (Nijhof, et al., 1985), cell sorting by flow cytometry (McAlister, I., et al., 1990), immunomagnetic bead separation (Bertoncello, I., et al., 1989), and antibody-complement cytotoxicity (Spangrude, et al., 1988).

An efficient time and material saving protocol for stem cell isolation would help in gene therapy, bone marrow transplants, elucidation of cellular differentiation pathways, etc.

The work reported here includes a review of our work of the previous summer. We were confident that we had achieved the goal of stem cell isolation and characterization. We have since learned that the cells we described were infected with a virus and our conclusions became suspect. We have taken a number of different approaches to be used with the bioreactors including the use of Dexter medium with horse serum and hydrocortisone. We have also added some conditioned medium to the vessels. Our mice were certified pathogen-free.

MATERIALS AND METHODS

Mice: Certified pathogen-free CD2/F1 male mice 6-7 weeks old were purchased from Taconic Farms, Germantown, New York, 12526.

Cells: Mice were sacrificed by cervical dislocation. The bone marrow cells were obtained by flushing tibiae and femora of ten mice with Phosphate Buffered Saline (PBS) without calcium and magnesium and supplemented with 1% Bovine Serum Albumin. These cells were washed twice and resuspended in growth medium. The growth medium in one experiment was Dexter formulation: Iscove's Minimal Essential Medium (Sigma Chemical Co.) containing 20% horse serum (heat-inactivated at 56 deg C for 30 min) (HyClone, Upton, Utah) and 10-6 M hydrocortisone succinate sodium salt (Sigma). This medium was supplemented with 100 U/ml Penicillin-Streptomycin (Gibco, Grand Island, N.Y.). Cells in this medium were grown at 33 deg C.

In a second experiment the medium was that used in the summer, 1989: RPMI-1640 (Gibco), 10% Fetal Calf Serum (Flow Laboratories, Raockville, MD), and supplemented with Penicillin-Streptomycin (Gibco). The Fetal Calf Serum was heat inactivated at 56 deg C for 30 min. Cells in this medium were grown at 37 deg C.

Cell concentrations in both reactor vessels was 1 x 10-6/ml. The vessel volumes were 50 ml. and were rotated at 8 rpm in an incubator in an atmosphere of 5% carbon-dioxide and 95% humidity. The bone marrow used in both controls and in experiments in microgavity initially was partially depleted of monocytes by removal of adherent cells by incubating the bone marrow cells overnight in polystyrene flasks at 37 deg C.in RPMI medium supplemented with 5% Fetal Calf Serum.

Simulated Microgravity:The cells were grown in one of NASA JSC's horizontally rotating culture vessels that simulates some aspects of microgravity. This vessel consists of a motor-driven rotating vessel with a separate air pump connected by tubing to the central shaft of the culture vessel. It operates in an ordinary laboratory tissue culture incubator. The cells are oxygenated by pumping 5% CO2 to 95% incubator air through the internal silicone membrane which surrounds the central vessel shaft.

Cell counts. Cell counts were obtained on a standard Coulter Counter Model ZM (Coulter Electronics, Inc., Hialeah, Fl., 33010) or by counting on a light microscope with a hemacytometer. Viabilities were determined by trypan blue exclusion.

Oxygen consumption and carbon-dioxide production were determined using a Corning 168, pH Blood Gas Analyzer (Corning, Medfield, MA, 02052). Glucose utilization was determined on a Beckman Glucose Analyzer 2 (Fullerton, CA, 92634).

Monoclonal Antibodies: The panel of antibodies included the following from clones purchased from American Type Culture Collection, Rockville, MD: TIB-104 (Lyt-1); TIB-120 (Anti Ia-b,d,q haplotypes); TIB-I28 (Anti-MAC-1, macrophages and granulocytes); TIB-146 (Anti B cell, anti B cell precursors with antigen B220); TIB-150 (Anti Lyt 2.2 expressed on T suppressor-cytotoxic cellsd); TIB-1264 (Anti murine B cell); TIB-207 (anti-L3T4-T helper-inducer subsets). The following antibodies were from Sigma: FITC-anti Thy 1.2 (cell surface differentiation antigen), and the secondary antibodies FITC-anti Rat IgG and FITC Goat anti-mouse. We obtained from Becton-Dickinson Co., Mountain View, CA: Anti-mouse Lyt-1 and anti- mouse I-A. The anti-stem cell antigen antibody E13-161-7 was from ATACC. 2.4G2 (anti-mouse Fc receptor was gift of Dr. MaryAnn Principato of the National Institutes of Health, Bethesda, Md.

Staining: In direct staining, FITC-conjugated antibodies were used. In indirect staining, the unconjugated antibodies were incubated first with the bone marrow cells for 30 min on ice, washed, and incubated with the secondary FITC-conjugated antibody for 30 min on ice. After two washings, the cells were fixed for 10 minutes at 4 deg C with 3.7% formaldehyde, washed and resuspended in PBS-1% BSA/0.02% azide. If we had immediate access to the FACScan, the cells were not fixed but were incubated with a solution of propidium iodide which enabled us to sort out dead cells for the flow cytometric analysis.

Flow Cytometry. Experimental data from the control cell populations grown in static culture at unit gravity and from the cells grown in microgravity in the bioreactors were obtained using Coulter Electronics EPICS V Cell Sorter (Coulter Electronics, Hialeah, FL). 10,000 events were scored on each test. Dead cells and cell aggregates were gated from the cells under study using propidium iodide. The percentage of positive cells was estimated by using bone marrow with irrelevant antibody as controls. Cell size distribution was assayed by forward scatter of unseparated bone marrow cells. The percentage of cells with staining above tghe background is determined for each of the monoclonal antibodies in the panel. For DNA analysis, the cells were stained according to the Method of Krishan as modified by McDivitt, et al. Cells were fixed by suspension in a solution of 67% ethanol for 30 min. The staining solution was made of 0.005% propidium iodide (Calbiochem-Behring, San Diego, CA., 0.002 % RNAse A (Sigma). After 30 min at 4 deg C the cells were washed and rsuspended in PBS.

Stem Cell Bioassay. (Hemopoietic colonies in vitro). Stem cells in complete medium were added to soft agar formed colonies. The ratio of the of colonies formed with the stem cell preparation and normal bone marrow cells was recorded. CFU-c was determined as described by Jacobs and Metcalf (I979). CFU-e was determined according to the method of Mcleod (1974). BFU-e was determined as described by Peschle, et al (1979). CFU-GEMM was determined by the method of Johnson and Metcalf (1977).

Immunological Reestablishment; Three sets of ten mice each were irradiated with 900 rads at MD Anderson Cancer Center, Houston, TX. One set of the mice was kept as controls. A second set received fresh bone marrow injections. The test group was given 100 cell/mouse of the cells which had been circulating in the microgravity vessel for 40 days. Records were kept of animal viabilities.

Phenotypic and functional assays: (a) Among the surface antigens to be studied by our panel of antibodies are: Thy-1, CD3, CD4, CD5, CD8, H-2 Class I and II, sIg, 14.8, SCA and Mac-1. (b) histochemical staining of the cells for monocyte detection will use specific esterase and for granulocyte/monocytes will use peroxidase.. (c) To determie if the cells secrete constitutively or after induction, we will use various hemopoietic growth and diferentiation factors: IL-1, IL-2, IL-3 IL-4, CAF and INF-gamma. These phenotypic and functional assays were performed in the Laboratory of Dr. Chris Platsoucas of the M.D. Anderson Hospital in Houston, Texas.

RESULTS

Cells grown at unit gravity in T-flasks using Dexter formulation medium were our most successful cultures. They continued to produce non-adherent cells through a two month period. The cells grown under the same conditions using RPMI-10% Fetal Calf Serum grew well for only a twelve day period. Cells in both Dexter medium and in RPMI-10%FCS medium in the bioreactor vessels declined in numbers steadily until there were very few viable cells by day 12. This was very unlike our three experiences in the summer of 1989. In the studies of last summer, the cell cultures were doubling every twenty-four hours by day twelve. Phenotypically they were the cells we sought -- the undifferentiated, hematopoietically pluripotent stem cells. We had apparently succeeded last year in producing a self-renewing immortalized cell line. Electron microscope photography indicated that the cells contained a virus. The virus could have accounted for their immortalization.

The cells grown with Dexter medium in the HARV did not survive beyond a few days; however similar cells grown in static culture in T-flasks at unit gravity were growing vigorously after five weeks. The cell suspension in the bioreactor by day twelve contained floating precipitate which we judged to be protein coming out of solution from the horse serum. The cells present had increased in size and resembled blasts. We concluded that the horse serum proteins precipitated which resulted in osmolarity and pH changes. Changes in either of these parameters would cause plasmolysis and lysis of the cells. We repeated the experiment several times and found the same results . We did not detect precipitated protein when the RPMI-10% Fetal Calf Serum was used; however, the cells did not survive after the 12 day incubation period. All of the horse serum in the Dexter composition was from a single batch; it is possible that this material was subject to denaturation in the vessels.

In the 1989 experiment we successfully monitored the conditions of the cultures in both static controls and in the rotating vessels using the parameters: osmolarity, pH, glucose and oxygen consumption and carbon dioxide production. We were also able to follow the depletion of the cells with lineage markers for cell surfce antigens. The remaining, immortalized culture, was frozen away for future studies. When thawed a year later, the cells readily went into culture and continued to double in number daily. The cells stained positively for Thy-1 and for SCA-1. They were negative the Fc-receptor. The cells were negative to staining with the lineage markers.

The cells that were monitored for reestablishment of their immune systems in the 1989 experiments did not give us positive results. It may have been that the 900 rad radiation was too high a dosage; it may have been that we did not inject sufficient numbers of the stem cells from the rotating vessels. This test will be included in any study of stem cells isolated using microgravity.

Our conclusions are that the conditions that produce an environment of microgravity in the bioreactors will be most helpful after some basic science questions have been answered. There are apparently some growth factors needed from the stromal cell environment that are unique for the stem cell. This factor is different from a proliferation factor. There may be needed articulation with some stromal cells in the microenvironment formed by the adherent cell population. Our continued studies will consider each possibility.

DISCUSSION

This research was undertaken to determine the growth characteristics of murine bone marrow cells in the bioreactors of NASA-Johnson Space Center. Since these vessels offer an environment that partially simulates the microgravity environment, the data obtained would be relevant to the space program.

The major premises upon which this research was based are reasonable; if they are sound, we would expect the self-renewing stem stell to survive in the environment of microgravity and we would not expect differentiated cells to do so.

Our investigations have been concerned exclusively to this point with murine bone marrow cells. We want to expand the study to include human bone marrow. We are also anxious to explore the behavior of murine fetal liver cells in microgravity. The proliferative rates of fetal liver cells are considerable greater than bone marrow rates. They are also a self-renewing population.

We are confident that the microenvironment will yield the stem cell clone. The procedure must be "fine tuned" and optimal conditions and media identified. We would use cells from this clone to search for any antigens that might be unique to the stem cell. We would also seek pathways from the stem cell to terminally differentiataed cells - myeloid, lymphoid, erythroid, as well as cells of current interest in medical research, e.g., the dendritic cell.

While we probably unintentionally immortalized the stem cell line in 1989, we are planning to transfect our isolated stem cells with a well defined oncogenic vector to produce an immortalized clone. In one experiment, canine bone marrow cells that were cultured for more than thirty weeks were found to be virally infected. (Schuening, F.G., et al, 1989). Murine bone marrow cells have been transformed in vitro with v-raf/v-myc retrovirus to yield mature B cells and macrophages. (Principato, M, et al., 1988) We would seek the stem cell clone using the same v-raf/v-myc vector. Retroviral vectors have been used with murine long-term bone marrow to transfect human glucocerebrosidase (Nolta, J..A., et al., 1990) and human adenosine deaminase (Wilson, J.M., et al, 1990). We are formulating protocols to transfect the stem cell clone if we succeed in isolating it. Erythropoietin cDNA has been transferred to murine bone marrow stromal cells and yielded hemoglobinized red blood cells. We would transfect our clone with cDNA and study the progeny. (Corey, C.A., et al 1989.)

The stem cell is deserving of further work in microgravity. Its isolation and characterization will be of particular relevance in bone marrow transplantation, immunodeficiency disorders, genetic engineering and long term space flights protection.

We are eager to study cells other than stem cells and self-renewing cells in microgravity, e.g., cells of the neuro-endocrine production.

The background preparation required for this study of murine bone marrow cells in simulated microgravity has been most rewarding. Our studies have prompted to ask many questions and now we must seek the answsers to these questions.

Are bone marrow stem cells indefinitely self-renewing? They appear to be programmed for division. When one of them under stimulus is targeted to become a particular differentiated cell, then another stem cell seemingly undergoes mitosis. This appears to be the case since the percentage of stem cells seems to be the same in infants and in adult humans. The role of the stromal cell microenvironment in hematopoiesis is well documented. Specific stromal cell lines have been isolated and cloned and some their growth factors, differentiation, factors and proliferative factors have been identified. Factors tht inhibit the differentiation of stem cells have not been identified, if they exist. Neither have proliferation factors. been identified that are specific for the stem cells.

It is important that the distinction between stem cells and precursor cells be defined. The literature is confusing. Investigators use the terms interchangeably. It is our conclusion that stem cells are programed and once they are committed to become a specific differentiated cell line, they cease to be stem cells and become precursor cells. The reaction is apparently irreversible.

The bioreactor vessels which provide an environment of simulated microgravity is a major advance in cellular investigations. When the vessels become more widely available to investigators, reports of their applicability will appear in scientific publications. It is understandable that the efforts of the NASA engineers in providing such an apparatus will be a boon to the space program. The environment of space is one of hypogravity and conclusive evidence that some cells behave differently in altered gravity has been well documented. Humans in space will be exposed for longer periods in altered gravity as space exploration advances. The bioreactor vessels can serve as model systems for in vitro studies which could give information that could be extrapolated to situations that will be encountered on missions to Mars and beyond.

BIBLIOGRAPHY

Cogoli, A., 1985. Lymphocyte Reactivity During Spaceflight. IMMUNOLOGY TODAY, 6:1.

Cogoli, A., Tschopp, and P. Fuchs-Bislin. Cell Sensitivity to Gravity. SCIENCE, 13 July, 1984, 228-230, vol 225.

Corey, A.C., Desilva, A., and Williams D.A., 1990. Erythropoiesis in Murine Long-term Marrow Cultures following Transfer of the Erythropoietin cDNA into Marrow Stromal Cells. EXP HEMATOL 18: 201-204.

Dexter, T. Michael. 1987. Growth factor involved in haemopoiesis. J. CELL SCIENCE, 88, 1-6.

Gmunder, F.K., and A. Cogoli. 1988. Cultivation of Single Cells in Space. APPL. MICROGRAVITY TECH. 1, 3, 115-122.

Hayflick, L. 1977. The Cellular Basis for Biological Aging. In Finch, C.D., Hayflick, L. (eds): HANDBOOK OF BIOLOGY OF AGING. New York: Van Nostrand Reinhold Co., 159-179.

Jacobs, P,. and D. Metcalf. 1979. Bone Marrow Culture in Vitro. EXP. HEMAT. 7:177.

Johnson, G.R., and D. Metcalf. 1977. Pure and mixed erythroid colony formation in vitro. PROC. NATL. ACAD. SCI, USA, 74: 3879.

Keever, A.A., Waite, K, Small T, Levick, J, Sullivan, M, Hsauch, M, Evans, R.L., O'Reilly, R.J. 1987. Interleukin 2-activated killer cells in patients following transplants. BLOOD, 70 (6) 1893-1903.

Lorenzi, T. 1986. Effects of Hypergravity on "whole blood" cultures of human lymphocytes. AVIAT. SPACE ENVIRON. MED. 57: 1131-1135.

McDivitt, R.W., Stone, K.R., Meyer, J.S. 1984. A Method for Dissociation of Viable Human Breast Cancer Cells that Produces Flow Cytometric Kinetic Information. CANCER RES., 44, 2628.

Nijhof, W. and Wierenga, P.K. 1985. Characterization of a Cycling Pluripotent Stem Cell Population. EUR J. CELL BIOL. 39, 136.

Nolta, J.A., Sender, L.S., Barranger, J.A., and Kohn, D.B., 1990. Expression of human glucocerebrosidase in murine long-term bone marrow cultures after retroviral vector-mediated transfer. BLOOD, 75, (3) 787-797.

Ploemacher, R.E., and Brons, N.J. 1988. Isolation of hemopoietic stem cell subsets from murine bone marrow. EXP. HEMATOL. 16, 210-26.

Principato, M., Klinken, S.P., Cleveland, J.L., Rapp. U.R., Holmes, K.L., Pierce, J.H., and Morse, H.C , 1988. In vitro transformation of murine bone marrow cells with a v-raf/v-myc retrovirus yields clonally related mature B cells and macrophages. CURR TOP MICROBIOL IMMUNOL, 141, 31-41.

Rowley, S.D., Sharkis, S.J., Hattenburg, C., and Sensenbrenner, L.L., 1987. Culture from Human Bone Marrow of Blast Progenitor Cells. BLOOD, 69, 804-808.

Schuening, F.G., Rainer, S. Meyer, J, and Goehle, S., 1989. Long-term Culture of Canine Bone Marrow Cells. EXP. HEMATOL. 17; 411-417.

Sonnenfeld, G, Mandel, A.D., Konstantinova, I.V., Taylor, G.R., Berry, W.D., Wellhausen S.R., Lesnyak, A.T., and Fuchs, B.B., 1990. Effects of Space flight on Levels and Activity of Immune Cells. AVIA SPACE AND ENVIRON MED, 648-653.

Spangrude, G.J., Heimfeld, S, Weissman, I.L. 1988. Characterization of Mouse Hematopoietic Stem Cells. SCIENCE. 241, 58-62.

Van De Rijn, M, Heimfeld, S., Spangrude, G.J., and Weissman, I.L., 1989. Mouse hematopoietic stem-cell antigen Sca-1 is a member of the Ly-6 antigen family. PROC. NATL. ACAD. SCI. USA, 86, 4634-4638

Wilson, J.M., Danos, O., Grossman, M., Raulet, D.H., and Mulligan, R.C., 1990. Expression of human adenosine deaminase in mice reaconstituted with retrovirus-transduced hematopoietic stem cells. PROC NATL ACAD SCI USA, 87, (10), 439-443.

Worton, R.G., McCulloch, E.A., and Till, J.E. 1969. Physical separation of hemopoietic stem cells from forming colonies in culture. J. CELL PHYSIOL. 74, 171.

SUMMARY

The data summarized graphically are from the successful culture growths of murine bone marrow in summer, 1989. The cells were grown in bioreactor vessels in an environment which simulated some aspects of microgravity.

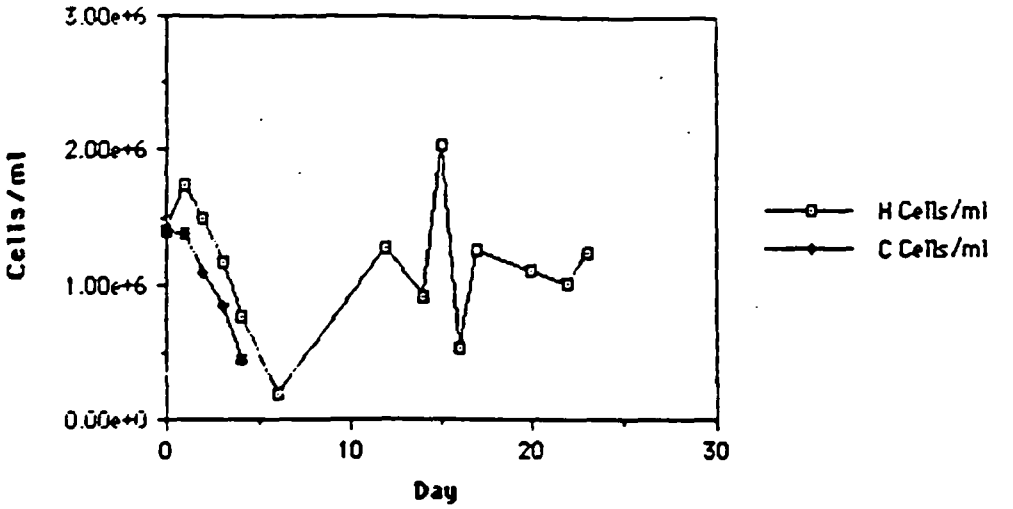
Page 13-13. Test Run #1
These cells were growing in simulated microgravity after 24 days. Their numbers at that time were doubling daily. The number of days of growth is plotted along the x axis and the log of the cell count number along the y axis.

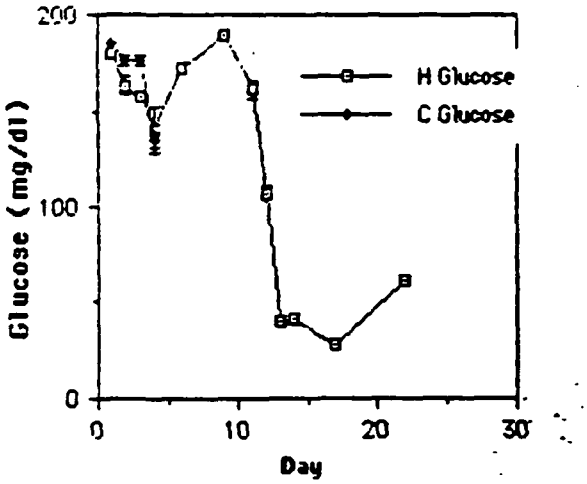
Page 13-14 This graphical data is from the flow cytometer program. The cells monitored are shown to be positive for Thy-1 and negative for FcReceptor antigen. This data is consistent with that expected for stem cells. Fluorescence intensity is plotted on the abscissa and cell numbers on the ordinate.

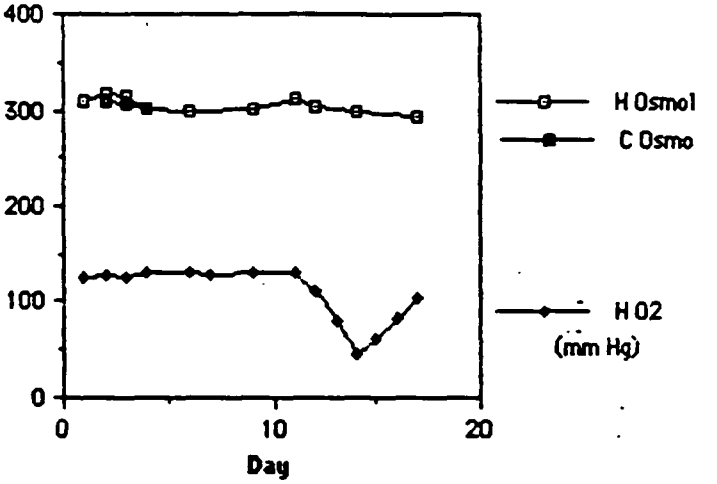
Page 13-15. This data from flow cytometer shows that the cells monitored are positive for the "stem cell antigen" known as SCA-1. The clone which produces this antibody is availabale from ATCC. Fluorescence intensity is plotted against forward light scatter.

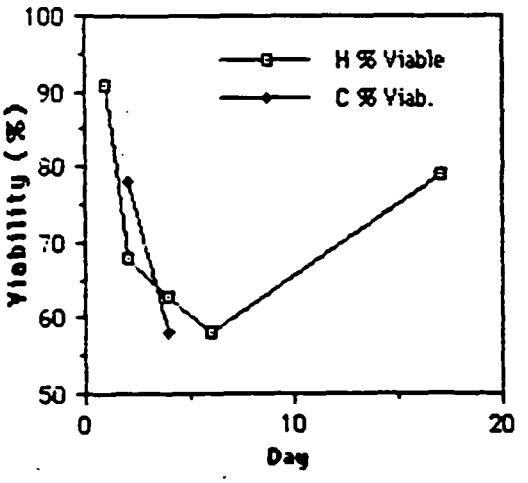
The program was gated to read only the viable cells in quadrant four.

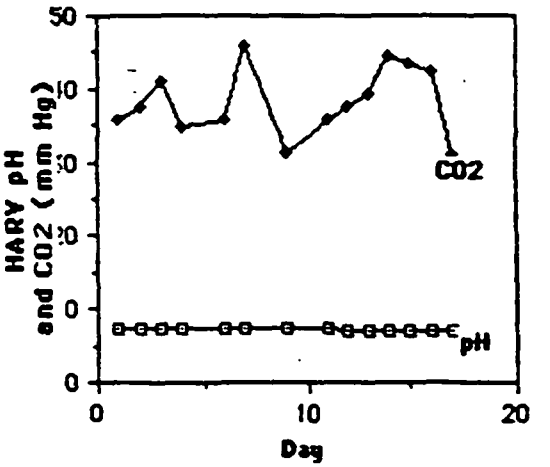
6/19/89

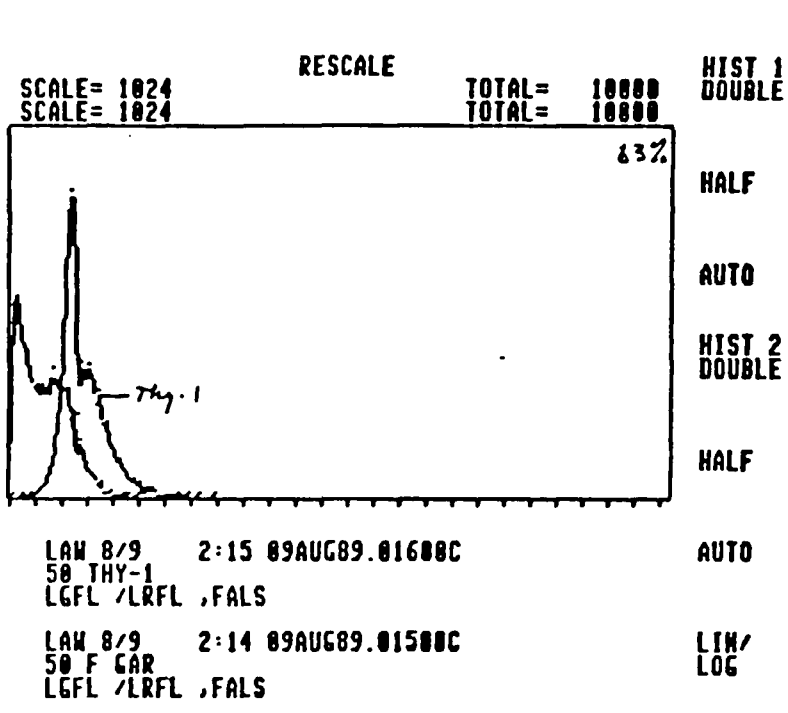

3.00e+6
2.00e+6
1.00e+6
0.00e+0
Cells/ml
0
10
20
30
Day
H Cells/ml
C Cells/ml

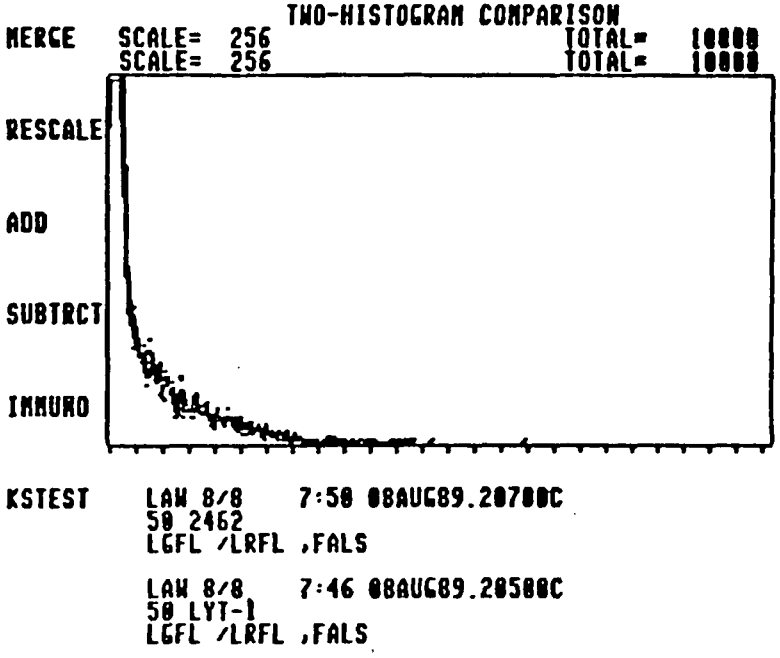

200
100
0
Glucose (mg/dl)
0
10
20
30
Day
H Glucose
C Glucose


400
300
200
100
0
0
10
20
Day
H Osmol
C Osmo
H O2
(mm Hg)


100
90
80
70
60
50
Viability (%)
0
10
20
Day
H % Viable
C % Viab.

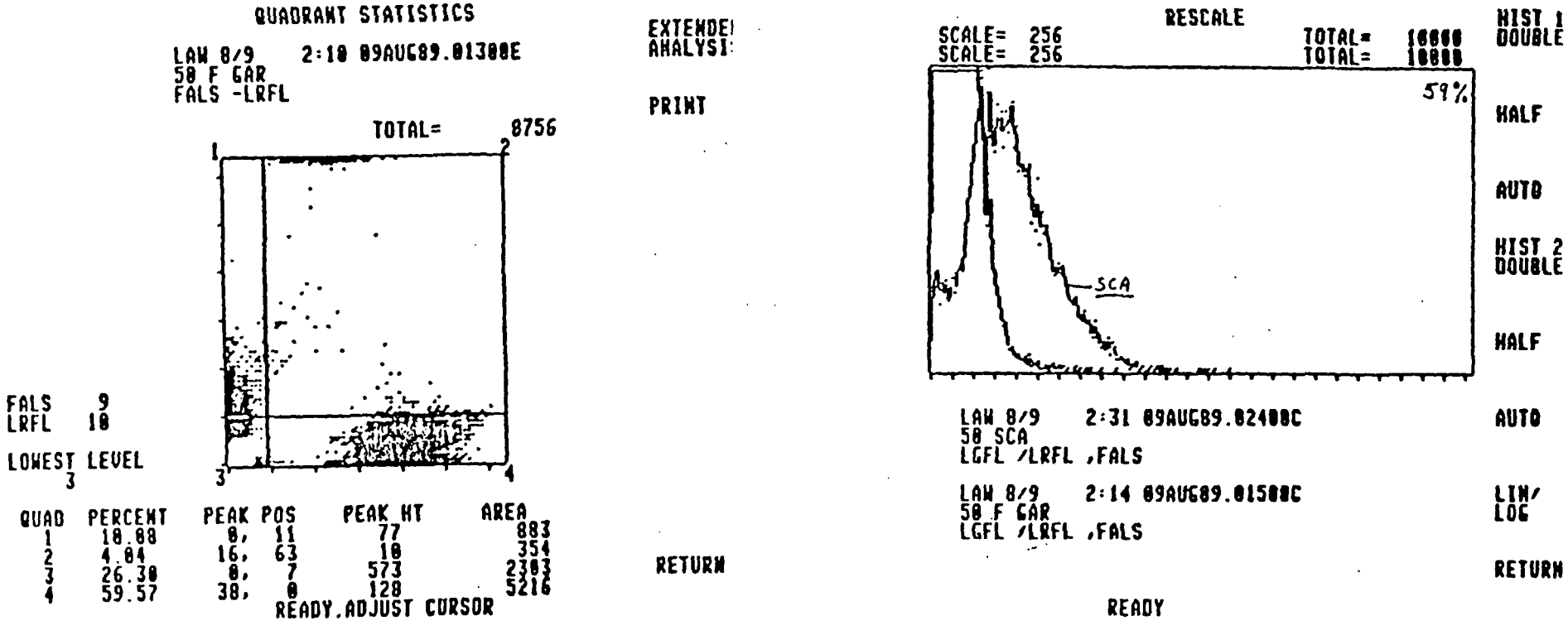

50
0
0
10
20
Day
HARV pH
and CO2 (mm Hg)
CO2
pH





FACS Analysis.

a. Thy-1+ cells
b. Lyt-1 and 2.4G2 (FcReceptor) -



FACS Analysis. Fluorescence Histograms. Fluorescence Intensity on abscissa; cell numbers on ordinate.

a. Cell viability -- Cells in quadrant #4 analyzed.

b. Stem cell antigen+ cells

THE EFFECTS OF DEUTERIUM ON STATIC POSTURE CONTROL

Final Report

NASA/ASEE Summer Faculty Fellowship Program - 1990

Johnson Space Center

Prepared by:	Charles S. Layne, Ph.D.
Academic Rank:	Assistant Professor
University and Department:	Kansas State University Department of Kinesiology Manhattan, Kansas 66506
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Space Biomedical Research Institute
JSC Colleague:	Millard Reschke, Ph.D.
Date Submitted:	July 27, 1990
Contract Number:	NGT-44-005-803

ABSTRACT

A significant operational problem impacting upon the Space Shuttle program involves the astronaut's ability to safely egress from the Orbiter during an emergency situation. Following space flight, astronauts display significant movement problems. One variable which may contribute to increased movement ataxia is deuterium (D^2O), an isotope of hydrogen. Deuterium is present in low levels within the Orbiter's water supply but may accumulate to significant physiological levels during lengthy missions. Deuterium has been linked to a number of negative physiological responses, including motion sickness, decreased metabolism and slowing of neural conduction velocity. The purpose of the present study was to investigate the effects of D^2O on static postural control in response to a range of dosage levels. Nine subjects were divided into three groups of three subjects each. The groups were divided in to a low, (50 g/70 Kg body water), medium (100 g/70 Kg body water), and a high (200 g/70 Kg body water) D^2O dosage group. The subjects static posture was assessed with the use of the EquiTest system, a commercially available postural control evaluation system featuring movable force plates and a visual surround that can be servoed to the subject's sway. In addition to the force plate information, data about the degree of subject sway about the hips and shoulders was obtained. Additionally, surface electromyographic (EMG) data from the selected lower limb muscles was collected along with saliva samples used to determine the amount of deuterium enrichment following D^2O ingestion. Two baseline testing sessions were performed using the EquiTest testing protocol prior to ingestion of the D^2O. Thirty minutes after dosing, subjects again performed the test. Two more post-dosing tests were run with an intertest interval of one hour. Preliminary data analysis indicates that only subjects in the high dose group displayed any significant static postural problems. Future analyses of the sway and EMG is expected to reveal significant variations in the subjects's postural control strategy following D^2O dosing. While functionally significant static postural problems were not commonly observed, subjects in both the medium and high dosage groups displayed significant, and in some cases, severe voluntary movement problems. These problems included locomotion and hand-eye coordination deficits.

INTRODUCTION

It is well documented that Space Shuttle crews are posturally ataxic following flight. This instability results from neurovestibular readaptation to a unit gravity environment, muscle deconditioning, and to some extent, cardiovascular deconditioning. Postural instability is a significant operational problem impacting upon rapid emergency egress from the Orbiter. Understanding the processes involved in movement control problems is critical if countermeasures are to be developed to increase the probability of successful emergency Shuttle egress.

One possible factor contributing to movement problems following space flight is an increase in the levels of deuterium oxide (D^2O) in the crew's physiological systems. Such an increase is likely to occur because of the presence of D^2O in the Shuttle's drinking water at levels exceeding those found in the nation's water supply. While the absolute levels of deuterium oxide in the Shuttle water are limited, it can be expected that a potentially physiologically significant amount may accumulate during extended duration flights. This increase results from the fact that D^2O is not "washed out" of the system with continuous ingestion of the Orbiter's drinking water.

While deuterium oxide has been shown to have a number of negative effects on mammals, few studies have attempted to assess the effects of D^2O on humans. In an effort to understand the impact of deuterium oxide on postural stability, a pilot study was conducted in June, 1990, as part of the Neurophysiology Laboratory's ongoing effort to probe the inherent complexities of the human motor control system. This project was undertaken in conjunction with the Nutritional Biochemistry Laboratory.

BACKGROUND TO THE PROBLEM

Deuterium oxide, the non-radioactive isotope of hydrogen, has been reported to have a variety of effects on physiological systems. Some of the effects include decreased metabolism rate (Reuter, et al., 1985), decreases in peripheral nerve conduction velocity (Thompson, 1963), decreases in maximum muscle twitch tension (Sato and Fujino, 1987) and positional alcohol nystagmus(Money and Myles,1974). While these symptoms are observed in response to

relatively high dose levels of D^2O, it can reasonably be expected that the physiological processes responsible for these symptoms are functioning similarly in response to lower levels of deuterium oxide. Thus, it becomes important to investigate what role, if any, the isotope plays in the disruption of normal functioning.

The majority of studies employing human subjects are designed to investigate the effects of deuterium oxide on the vestibular system. Money and Myles (1974) reported that their subjects experienced dizziness, nausea, and strong sensations of bodily rotations following D^2O ingestion. These symptoms primarily occurred while the subjects were lying on their sides. It was also reported that the subjects displayed a variation of positional alcohol nystagmus (PAN I) while in the lateral position. Nystagmus is an involuntary oscillating movement of the eyeballs in which they repeatedly turn slowly in one direction and then rapidly reverse direction. The direction of the nystagmus is identified based on the direction of the fast phase of the eye movements. Drinking large amounts of ethyl alcohol results in the eyes moving to the right during fast phase of the nystagmus if the head is held with the right side down. PAN I is called positional because it occurs when the head is maintained in a certain position relative to gravity.

After drinking D^2O, Money and Myles subjects displayed what these authors termed as PAN II. PAN II differs from PAN I in that the direction of the fast phase of the nystagmus is in the direction opposite from the side the head held in the lateral position. For example, if the subject is lying on his or her left side the fast phase is to the right. Nystagmus can be observed with the use of Fresnel lenses placed over the subjects eyes. These specially designed spectacles provide an opaque display to the subject while magnifying the subject's eyes to the observer. In this manner, rapid movements of the eyes can readily be detected.

Money and Myles (1974) suggested that deuterium oxide elicits nystagmus by the following mechanism. The cupula of the semicircular canal floats in a surrounding liquid (endolymph) which has the same density. Under normal circumstances, angular accelerations cause the cupula to move as a result of circular endolymph movements. However, because of its neutral buoyancy the cupula is not moved by linear accelerations or by gravity. Following ingestion of deuterium oxide, the cupula acquires D^2O faster than the surrounding endolymph because of the prominence of blood

capillaries near the base of the cupula. The cupula's neutral buoyancy is therefore lost as it becomes heavier than the surrounding endolymph. This imbalance allows gravity to move the cupula, thus, producing positional nystagmus responses when the head achieves specific orientations relative to gravity.

Greven, et al., (1977) have also reported similar responses to D^2O. More importantly, perhaps, is their finding that the speed of the fast phase of the PAN II nystagmus was increased during exposure to G-loads of 2 and 3 G. Thus, increased level of gravity appear to enhance the response of the vestibular system to deuterium oxide.

The reported findings have important implications for astronauts attempting rapid egress from the Space Shuttle as it is well known, from work with vestibular patients, that altered vestibular input is disruptive to both postural and movement control. Greven, et al's. (1977) results suggest that even low levels of D^2O in the astronaut's physiological systems may pose a problem during the high gravity loads experienced during re-entry. This report describes a preliminary attempt to characterize the effects of three different dosage levels upon static postural control.

PROJECT SUMMARY

Methods

The static postural control of nine volunteer subjects was tested both prior to and following ingestion of deuterium oxide. The nine individuals were divided into three groups of three subjects each. Each group was administered a different dose level which were characterized as low, medium or high doses. The absolute amount of D^2O administered to each subject varied as levels were adjusted to percentage of body water levels based upon percentage of body fat measures. The three dosage levels were equivalent to: a) 50 g/70 Kg body water; b) 100 g/70 Kg body water and c) 200 g/Kg body water. The dosage levels were consistent with the range of levels reported in the literature. The D^2O supply was 99.9% atom percent (ICON Services).

The standing static posture of the subjects was tested using the EquiTest protocol (Neurocom International). The EquiTest system consists of a motor-driven footplate, capable of translations in the

anterior-posterior directions and rotations about the ankle joint. A motor-driven visual surround allows for rotational movements about the subject's ankles. The force platform provides information about the ground reaction forces exerted along the X,Y,and Z axes and horizontal sheer. Sway bars interfaced with potentiometers, are attached to the shoulders and hip area and provide information about the degree of movement about these two joints. The testing protocol was comprised of a series of movement coordination and sensory organization tests. The subjects attempted to maintain equilibrium without lifting their feet from the platform while their arms remained folded across their chest.

The movement coordination test was subdivided into four unique tasks.

1. Anterior translations (resulting in backward sway)
2. Posterior translations (resulting in forward sway)
3. Toes-up rotations
4. Toes-down rotations

Subjects experienced three trials of Conditions 1 and 2 and five trials of the two remaining conditions.

The sensory organization test is designed to determine the relative contribution of different sensory systems to the subject's overall equilibrium strategy. This can be accomplished because of two features of the EquiTest system. The visual surround can be made to move in response to the pressure exerted by the subject's feet on the force platform. As a subject sways forward, the visual surround sways forward an equivalent distance. In this situation, visual feedback about the degree of subject sway is nulled out. If a subject were to rely exclusively on visual input to maintain balance he or she would soon fall forward. Thus, to maintain upright stance, the subject must learn to utilize vestibular and proprioception inputs.

The second feature of the EquiTest systems which enables investigators to determine to what extent a subject relies on a specific sensory input is the platform's ability to sway in response to changes in the subject's applied foot pressure. For instance, if the subject begins to exert pressure on the front of the platform the support surface begins to rotate downward. In this situation the subject receives inappropriate proprioception information from the ankle joints. If the subject were to rely exclusively on this inappropriate input, which is indicating to the subject he or she is

not actually swaying forward because the initial ankle angle is being maintained, the subject would soon lose balance. By combining the above two features in either eyes open or eyes closed conditions, a complete profile of the relative weighting each subject gives to the various sensory inputs can be achieved.

The sensory organization test was subdivided into six unique tasks.

1. Normal vision, fixed support surface
2. Absent vision (eyes closed) fixed support surface
3. Sway-reference vision, fixed support surface
4. Normal vision, sway-reference support surface
5. Absent vision (eyes closed), sway-reference support surface
6. Sway-reference vision, sway-reference support surface

Each task was 20 seconds in duration with tasks 3,4,5 and 6 being repeated three times.

The following parameters are routinely derived from the above testing protocol and are a output from the system in a hardcopy format following completion of the testing.

Movement Coordination Tests:

Static Symmetry: Average right/left weight distribution during test.

Force Latency: Time from onset of support surface perturbation until initial active force response. This is calculated separately for each leg.

Strength Amplitude: Average slope of initial active force response to support surface perturbation. This is calculated separately for each leg.

Dynamic Symmetry: Average right/left strength amplitude distribution during test.

Sway Amplitude: Average (rms) and peak-to-peak amplitudes of anterior-posterior sway calculated for center of mass, ankle angle, hip displacement, shoulder displacement, and head displacement.

Adaptation: Relative decrease in sway amplitude occurring with repeated presentation of the ankle rotation stimulus.

Strategy: Relative contributions of ankle and hip joint motions to sway control following support surface translations.

Sway Amplitude: Average (rms) and peak-to-peak amplitudes of sway calculated for center of mass, ankle angle, hip displacement, shoulder displacement, and head displacement during test.

Strategy: Relative contributions of ankle and hip joint motions to sway control during test based on force, transducer and joint position measurements.

Equilibrium: Relative stability of the subject's sway control during each individual test.

Composite Equilibrium:: Composite estimate of the relative stability of the subject's sway control based on weighted sums of individual test equilibrium scores.

In addition to the above parameters, surface electromyography (EMG) was used to monitor the activity of the following muscles of the left lower limb: soleus, gastrocnemius, bicep femoris (hamstring), and rectus femoris (quadricep). EMG data from the movement coordination tests was analyzed by obtaining latency and amplitude variables. Latency measures were obtained by determining the initial increase in muscle activity relative to platform movement. Amplitude measures were defined as the average (rms) amplitude from platform movement onset until the end of the active force response. Both variables were determined independently for each muscle, for each trial. The Pensacola Coriolis Sickness Scale was administered throughout the testing procedures in order to monitor the subject's level of possible physical discomfort. The Pensacola Coriolis Sickness Scale has been extensively used to obtain a subjective measure of motion sickness.

All testing was completed in the Dynamic Posture Laboratory located in the KRUG Life Sciences' portion of the Intermetrics building (1100 Hercules Drive, Houston, Texas). Subjects arrived at

the testing site and were prepared for EMG collection. This involved shaving and lightly abrading the skin above the electrode placement sites. The electrodes where then positioned over the muscles and skin impedance measures were obtained. Before the subjects mounted the platform, they donned a safety harness. The safety harness attached to a safety bar that looped over the platform such that it was impossible for the subjects to collapse if postural equilibrium was completely lost. Once the subjects stepped onto the platform the EMG lead were connected to the amplifiers and EMG signal quality was checked. The testing procedures began by positioning the subject's feet precisely on the platform and asking the subject to fold their arms across their chest. Headphones providing white noise designed to eliminate external cues from the platform, were then placed on the subjects.

In order to obtain baseline values, the motor coordination and sensory organization test protocol were completed two times. There was a one hour intertest interval between the baseline collection sessions. Immediately following the second baseline testing session, the subjects ingested the pre-determined amount of deuterium oxide. The subjects then assumed a comfortable lying position on their left side.

Thirty minutes after dosing, subjects again mounted the posture platform and were subjected to the motor coordination and sensory organization tests. Then subjects then resumed their reclining position. Two additional post-dosing sessions were completed with an intertest interval of one hour. Both prior to mounting the platform and at the conclusion of each testing session the Pensacola Coriolis Sickness Scale was administered. Prior to each of the baseline testing sessions, a saliva sample was collected in order to determine the baseline levels of deuterium. Following D^2O dosing and prior to each of the enrichment of the sample. In some cases, Fresnel lenses were applied and the subject was checked for nystagmus. Following completion of the posture tests, the subjects drank two liters of water. This procedure was employed to increase the rate at which the deuterium oxide was "washed out" of the system.

Results and Discussion

The saliva samples were centrifuged at Johnson Space Center in the Nutrition Biochemisty Laboratory and then sent to University of Chicago Medical Center for stable isotope analysis. The results of this analysis are unavailable at the present time.

Preliminary analysis of the data obtained from the EquiTest system printout revealed that D^2O had little affect upon the subjects static postural control. This was especially true for the low dosage group. However, at the medium and high dosage levels, initial analysis of conditions 5 and 6 of the sensory organization tests, revealed subjects were less stable post-dose relative to baseline measures. This result is consistent with the previously reported findings that vestibular system functioning is altered as a result of deuterium oxide ingestion. Both condtions 5 and 6 are configured such that the subject must preferentially rely on vestibular input to maintain upright stance. Therefore, is was hypothesized that the results from these two conditions would be most sensitive to changes in postural stability following D^2O ingestion.

Results from the Pensacola Coriolis Sickness Scale revealed that none of the subjects in the low dosage group indicated any unusual physical sensations other than a feeling of being chilled. This finding may possibly be explained by the fact that deuterium has been reported to decrease metabolism. Similar sensations were reported by the subjects in the medium dosage group.

On the other hand, subjects in the high D^2O dosage group reported a variety of physical symptoms, including movement illusions both of their body and the environment. All three subjects in this group displayed nystagmus based on Fresnel lense observation. One subject vomited repeatedly. This subject's nausea was most violent following changes in body postion, i.e., moving from either a reclining position to standing or vice versa. Despite these symptoms, this subject was able to achieve static balance measures consistent with his baseline scores. Consistent with the reports of Money and Myles (1974), the only subject to become significantly nauseated refrains from all exposure to ethyl alcohol. The reason for this is unknown, but it may be that drinkers have developed compensatory processes allowing them to function relatively

unimpaired in response to low to moderate levels of alcohol consumption. These compensatory processes may have been operating in response to altered vestibular input resulting from deuterium ingestion.

One interesting phenomenon was observed in all members of this group and in one subject in the medium dosage group. These subjects reported (and displayed) numerous movement coordination problems when asked to make voluntary limb movements. The subjects became very unstable when asked to walk even short distances. This instability was compensated for by adopting wide gait patterns. This behavior closely resembled that of patients with cerebellar disturbances. These subjects also reported disturbances in hand-eye coordination. For instance, one subject displayed great difficulty in securing a buckle similar to a commen seat belt about his waist. It is possible that the above symptoms may be partially the result of neuromuscular processes which were impaired by the relatively high level of deuterium.

Due to both time limitations and technical problems, further analysis of the hip and shoulder sway data and the EMG data has not been completed. Sway data is currently being evaluated by members of the Dynamic Posture Laboratory working group. The EMG data will be analyzed in the coming months in the Neuromuscular Laboratory at Kansas State University.

Conclusions

The preliminary results of the present study indicate that high levels of deuterium in the body's physiological systems can result in significant movement problems, as well as produce epigastric sickness in some subjects. Static postural control appears to be only slightly impaired, even with relatively high dosage levels of deuterium. However, subsequent analysis of body sway parameters and EMG data is expected to reveal that the subjects were, in fact, less stable following ingestion of deuterium oxide. Future research should focus on the processes responsible for the observed voluntary movement problems displayed by the current subjects. Obviously, any factor which contributes to impaired movement control greatly impacts upon an astronaut's ability to safely egress from the Space Shuttle in an emergency situation.

REFERENCES

Sato, Y. and Fujino, M. Inhibition of arsenazo III Ca transient with deuterium oxide in frog twitch fibers at a resting sarcomere length. Japanese Journal of Physiology, 37 (1) 149-53, 1987.

Money, K. E. and Myles, W. S. Heavy water nystagmus and effects of alcohol. Nature, 247 (5440) 404-405, 1974.

Reuter, H. D. Fischer J. H., and Thiele, S. Investigations on the effects of heavy water (D^2O) on the functional activity of human platelets. Haemostasis, 15 (3) 157-63, 1985.

Greven, A. J., Oosterveld, W. J. and Samson, G. The influence of heavy water on the vestibular system: A study of heavy water nystagmus in Advances, in Oto-Rhino-laryng, vol 22, p.152-159.

Thomson, J. F. Biological Effects of Deuterium, Pergamon Press: New York, New York, p. 80, 1963.

NASA
National Aeronautics and Space Administration

REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR 185637		

4. Title and Subtitle	5. Report Date
NASA/ASEE Summer Faculty Felllowship Program-)-1990) Volume 1	December 1990
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Richard B. Bannerot and Stanley H. Goldstein, Editors	
9. Performing Organization Name and Address	10. Work Unit No.
University of Houston-University Park Houston, TX 77204	11. Contract or Grant No. NGT 44-005-803
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
National Aeronautics and Space Administration Washington, D.C. 20546	Contractor Report
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The 1990 Johnson Space Center (JSC) National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program was conducted by the University of Houston-University Park and JSC. The 10-week program was operated under the auspices of the ASEE. The program at JSC, as well as the programs at other NASA Centers, was funded by the Office of University Affairs, NASA Headquarters, Washington, D.C. The objectives of the program, which began nationally in 1964 and at JSC in 1965, are (1) to further the professional knowledge of qualified engineering and science faculty members; (2) to stimulate an exchange of ideas between participants and NASA; (3) to enrich and refresh the research and teaching activities of participants' institutions; (4) to contribute to the research objective of the NASA Centers.

Each faculty fellow spent at least 10 weeks at JSC engaged in a research project commensurate with his/her interests and background and worked in collaboration with a NASA/JSC colleague. This document is a compilation of the final reports on the research projects performed by the faculty fellow during the summer of 1990. Volume 1 contains reports 1 through 14, and Volume 2 contains reports 15 through 28.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
	Unclassified - Unlimited

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified		

For sale by the National Technical Information Service, Springfield, VA 22161-2171