FIRST NATIONAL STUDENT CONFERENCE

NASA UNIVERSITY RESEARCH CENTERS
AT MINORITY INSTITUTIONS

"Shaping Young Minds through Aerospace Research and Education Partnership"

FIRST NATIONAL STUDENT CONFERENCE

NASA UNIVERSITY RESEARCH CENTERS
AT MINORITY INSTITUTIONS

MARCH 31 - APRIL 2, 1996
GREENSBORO, NORTH CAROLINA, U.S.A.

Editors:

Endwell O. Daso

Stacie Mebane*

Center for Aeorspace Research
North Carolina A&T State University

*Undergraduate Student Assistant

TSI Press Albuquerque, New Mexico 1997

About the logo on the cover: The logo represents the partnership of the member research centers of NANURC in aerospace research and education.

Logo Designer: The Center for Aerospace Research, North Carolina A&T State University, Greensboro, NC.

TSI Press
Division of TSI Enterprises, Inc.
P.O. Box 14126
Albuquerque, NM 87191-4126, USA
Phone : 505 298 5817 FAX : 505 291 0013
Email : **jsj@swcp.com http://www.tsie.com/~jsj**

ISBN: 1-889335-04-5

Table of Contents

Keynote Address, Invited Papers & Presentations

Tuskeegee University

Morehouse School of Medicine

Fisk University

Prairie View A&M University

University of Puerto Rico

Tennessee State University

University of New Mexico

Alabama A&M University

University of Texas at El Paso

Howard University

Clark Atlanta University

North Carolina A&T State University

Florida A&M University

Hampton University

Message from the Chancellor

Greetings:

It was a pleasure for us, representing one of the Historically Black Universities involved in the NASA Core program, to host the recent meetings held on our campus in Greensboro as associated with major NASA driven endeavor. The commitment which this federal agency has demonstrated, as associated with the efforts of North Carolina 11 State University, and its sister institutions - in the development and operationalizing of programs geared to insure instructional excellence in science, engineering and mathematics - is representative of a major success story for the Academy.

These campuses, in concert with state-of-the-art activities on the part of their students, faculty/scientists and administrators, have joined forces with NASA administrators and scientists to produce models of excellence which can now serve as benchmarks, nationally. We are convinced that these efforts will go a long way toward insuring the continued increased productivity of outstanding minority engineers and scientists and mathematicians. North Carolina A&T State University and its sister institutions commend this federal agency for its long-standing support and commitment as associated with the effort of HBCUs to make an impact upon this nation and its economic development ethos.

Edward B. Fort
Chancellor
North Carolina A&T State University

"Reach for the Stars"

Dr. Philip J. Sakimoto*
Team Leader, Minority University Research & Compliance
Minority University Research and Education Division
Office of Equal Opportunity Programs
NASA Headquarters
Washington, DC

Space is exciting

You were born at the right time. You were born at the right time to change the very nature of human civilization, because during your lifetimes I predict that human beings will leave the planet of their birth, and set foot on Mars. This is not a pipedream. The NASA Strategic Plan clearly states that in the years from 2010 to 2020 and beyond we will "conduct international human missions to planetary bodies in our solar system." The current mood of the country and Congress does not presently allow funding of such an effort, but--like all things--that mood will someday change. And then we will go. We will go where "no one has gone before."

The very nature of our relationship with space is changing. In the 60's and 70's we showed that human beings could go to the Moon--a mere 250,000 miles away. In the 80's we debuted a reusable space transportation system, the Space Shuttle, and in the 90's we showed that such a system is indeed reliable and safe for routine use. As we move into the 2000's, we are building an international space station--a permanent habitat in space--so that in the next century, living and working in space will become routine, and we will be prepared to go to Mars--a daunting 200 *million* miles across the solar system.

We are here at this conference to invite you to live and work in space with us, or to engage in the myriad of related activities which we perform on the ground to preserve the Earth's environment, to understand the origin and evolution of the universe, to build reliable aircraft systems, to gift America with new technologies, and to develop space systems so that a lucky few of us can actually go into space.

Opportunities are large, but vanishing

I can read your minds. You are sitting there wondering if there really is a place for you in the space program. My answer is a resounding, "Maybe." Within NASA we are rapidly moving to change the complexion of the space program. Our Administrator, Mr. Daniel Goldin, has made a strong commitment to diversify the NASA workforce from top to bottom. He has appointed to the highest ranks, senior managers who are African American, Hispanic, Asian American, and female. He has challenged us, the people charged with leading NASA's efforts with minority institutions, to double, triple, and quadruple the rates at which minority students are getting advanced degrees in technical fields. I would like to say that the door is now wide open for you--those very students--to come into the space program. But it is not. The door is slamming closed at an alarming rate. The door is being slammed closed by forces we cannot control.

In California, Governor Pete Wilson signed an executive order to end affirmative action in California, and in response the University of California Board of Regents voted to stop considering race and gender in admissions or hiring . In Maryland, an

*Dr. Philip Sakimoto is the Team Leader for NASA's Minority University Research Programs. He has previously served as the NASA Headquarters Minority University Research and Education Division's Program Analyst for Evaluation and Assessment, University Affairs Officer for NASA Goddard Space Flight Center, Assistant Director of the Johns Hopkins Space Grant Consortium, and Visiting Assistant Professor of Physics and Astronomy at Whitman College. He holds a B.A. in Physics from Pomona College, and an M.A. and Ph.D. in Astronomy from the University of California at Los Angeles.

Appeals court ruled that the University of Maryland's Banneker Scholarship Program for high-achieving African Americans is illegal and discriminatory. Two weeks ago, the U.S. Circuit Court of Appeals struck down admissions policies at the University of Texas law school, ruling that it is a violation of the Constitution to give preference blacks and Hispanics. And as we speak, the U.S. Congress is considering draft legislation entitled "The Equal Opportunity Act of 1996," which threatens to make illegal the granting of preferences of any kind based on race, gender, or ethnicity. This act does preserve programs for HBCU's, but it does not preserve programs for Hispanic Serving Institutions or for any other class of minority institutions.

In NASA, we are phasing out undergraduate scholarship programs targeted at minorities and graduate fellowship programs targeted at minorities. We do still have a minority university program--that is why you are here today--but how long will this country allow us to retain such a program? Some days it seems like not long at all.

Does this mean that it is time for you to give up? NO! It means that it is time for you to do what we have always had to do. We have to get ahead by forging our own pathways, by rising to overcome all roadblocks, by refusing to quit no matter how hard the going gets.

The Prescription

There still are opportunities for you in the space program. *You are taking advantage of one of them right now.* You are here at this conference by virtue of our University Research Centers at Minority Institutions program, and at home you are conducting research work related to the space program under the same sponsorship. Other opportunities may also come, but they are not likely to just appear at your doorstep. You will have to make them yourself. Here is my prescription for doing so:

*1. **Perfection.*** How many of you have ever gotten 100% on a college exam? 98%? 95%? 90%? What is the difference between those who got 90 or 95 or 98, and those few who got 100? It's perfection. Almost perfect is not good enough. Just getting an A is not good enough. And the way that you achieve perfection, is by refining and checking and re-checking your work until *no one* can find a flaw in it. It takes a lot of time, and it takes a lot of dedication. But that's what makes the difference between 98% and 100. Does this difference matter? In space, it's all that matters. Would you climb on a shuttle that was designed 98% right and 2% flawed? Of course not. So if you want to get into the space program, you must routinely perform at 100%. You must teach yourself to routinely achieve perfection.

*2. **Entrepreneurship*** Opportunities will not just appear on your front door. You have to make them happen. You have to sell yourself. Here's how it works. First, you have to learn what it is that your customer wants. You have to study their literature (in our case, the NASA Strategic Plans), you have to visit them and ask just what it is that they do, you have to talk with anybody you know that knows them, to find out all the hidden details and all the unwritten rules. Second, you have to assess what it is that you do that will help out what it is that they do. Third, you have to discover where the potential entryways are. Finally, you have to make an opportunity to present yourself to offer your services within the context of that entryway. It sounds simple, but it is a lot of work. Typically, what I have just outlined represents perhaps six months of effort, to find a good opportunity.

*3. **Focus**. It is a jealously guarded secret that it is actually much easier to get a job doing exactly what it is that you want to be doing than it is to get a job doing something else.* That is so contrary to what most students think, that its worth repeating [repeat]. The reason is that when you are doing what you want most to be doing, you are usually better at it, more interested in it, and more dedicated to it, than you are at other things. That is exactly what an employer wants: high quality, interested, dedicated employees. So figure out exactly what it is that you really want to be doing, and then set out to find the person who needs someone to

do just that. If you don't believe this, think about the curse of doing the opposite: if you try to get a job you don't really want to be doing, you might actually end up with a job you don't really want to be doing. Why strive for that?

4. Flexibility and Context. Of course, you won't always walk right into the right opportunity the first time you try. More likely, you will find that there are some additional skills you need to pick up, some roundabout pathways you must take to get there, or some modifications you must make to fit into the context of the opportunities you have. So be willing to be flexible, but never lose sight of your destination.

Rewards

Is all of this worth the effort? I think it is. Your rewards will be genuine job satisfaction, a chance to make a real mark on the world, and the feeling of peace with yourself.

So stop just sitting around here. When today's sessions are over, go back to your hotel room. Polish up your presentation. Time yourself to ensure that you will not be stopped before you are through. Make it just a little bit closer to perfect. Make yourself ready to begin selling yourself to your audience.

And what if you don't make it? What if you do all theses things and you still don't make your ultimate dream of being an astronaut or of designing a hypersonic transport? That's OK. Your training and efforts will still lead you into a satisfactory place in industry, in the university community, in teaching, or in other places not yet imagined. There is no shame in failure, there is only shame in not giving it your absolute best.

When you reach for the Stars, you might make it as far as Mars--and that's not doing too badly at all.

Message from the Vice Chancellor

Dear Colleagues:

I am pleased to have been a part of the First National Student Conference that was hosted by the NASA Center of Research Excellence (CORE) at North Carolina A&T State University. The enthusiastic participation of numerous government and industry representatives, and administrators, faculty, and students from universities throughout the country was overwhelming. This Proceedings reflects the quality, depth, and breadth of the research presented during the Conference. It was especially rewarding to observe the outstanding participation of so many students who have been engaged in the highest quality of research through the fourteen (14) NASA-sponsored Research Centers of Excellence.

This outstanding Conference was sponsored by the National Alliance of NASA University Research Centers at Minority Institutions. These fourteen (14) university research centers were established by NASA to foster new science and technology concepts, expand the nation's base for aerospace research and development, develop mechanisms for the increased participation by faculty and students of Historically Black Colleges and Universities and Minority Institutions in mainstream research, and to increase the production of underrepresented minorities with advanced degrees in NASA-related fields.

The goals of the Alliance, made up ot these fourteen universities, are to enhance collaborative research, promote aerospace engineering education at member institutions, and to develop new science and technology concepts through innovative research in the framework of NASA's mission and Strategic Enterprises.

This Proceedings represents the efforts of the Alliance through its *First National Student Conference.* I hope that the enthusiasm of the participants in the Conference is evident through this document.

Harold L. Martin
Vice Chancellor for Academic Affairs
North Carolina A&T State University

National Alliance of NASA University Research Centers
at Minority Institutions (NANURC)

Alliance Members

Center for Nonlinear and Nonequilibrium Aerosciences (CeNNAs)
Florida A&M University, Dr. Joseph J. Johnson, III, Director and Chairman of NANURC

Research Center for Optical Physics (RCOP)
Hampton University, Dr. Doyle Temple, Director

Center for Photonic Materials and Devices (CPMD)
Fisk University, Dr. Enrique Silberman, Director

High Performance Polymers and Ceramics Research Center (HiPPAC)
Clark Atlanta University, Dr. Eric Mintz, Director

Center for Aerospace Research (CAR)
North Carolina A&T State University, Dr. Endwell O. Daso, Interim Director

Center for the Study of Terrestrial and Extraterrestrial Atmospheres (CSTEA)
Howard University, Dr. Arthur N. Thorpe, Director

Center for Food Production, Processing and Waste Management in Controlled Ecological Life Support Systems (FPPWM in CELSS)
Tuskegee University, Dr. Walter A. Hill, Director

Space Medicine and Life Sciences Research Center (SMLSRC)
Morehouse College of Medicine, Dr. Gary L. Sanford, Director

Center for Applied Radiation Research (CARR)
Prairie View A&M University, Dr. Thomas A. Fogarty, Director

Center for Automated Space Science (CASS)
Tennessee State University, Dr. Michael Busby, Director

Center for Hydrology, Soil Climatology and Remote Sensing (CHSCaRC)
Alabama A&M University, Dr. Tommy A. Coleman, Director

Center for Autonomous Control Engineering (ACE)
University of New Mexico, Dr. Mo Jamshidi, Director

Tropical Center for Earth and Space Studies (TCESS)
University of Puerto Rico at Mayaquez, Dr. Rafael Fernandez-Sein, Director

Pan American Center for Earth and Environmental Studies (PACE)
University of Texas at El Paso, Dr. Scott Starks, Director

Message from NANURC

The National Alliance of NASA University Research Centers (NANURC) is a consortium of fourteen Historically Black Colleges and Universities and Minority Institutions. The member Centers of the consortium were established and are supported by NASA Headquarters under the Minority University Research and Education Division. These Research Centers were established to foster new science and technology concepts, expand the nation's base for aerospace research and development, develop mechanisms for the increased participation by faculty and students of Historically Black Colleges and Universities and Minority Institutions in mainstream research, and increase the production of socially and economically disadvantaged students with advanced degrees in NASA-related fields.

The goals of NANURC are to enhance collaborative research, promote aerospace engineering education at member centers/institutions and develop new science and technology concepts through innovative research in the framework of Strategic Enterprises and NASA's mission. Also, through broad-based research programs in support of NASA, the Alliance will facilitate technology transfer to strengthen science and engineering and utilize the scientific and technical expertise of the member institutions to serve the aerospace industry, NASA and the nation.

NANURC is "Shaping Young Minds through Aerospace Research and Education Partnership" to "fill the pipeline" by educating and graduating students with aerospace-related backgrounds to meet the technical and diversity needs of NASA, other government agencies and the aerospace industry.

We welcome your support in helping us accomplish these and other important goals of NANURC.

Joseph A. Johnson, III

Message from the Director

Center for Aerospace Research
NASA Center of Research Excellence (NASA-CORE)
College of Engineering
North Carolina A&T State University, Greensboro, NC

In January 1992, NASA established NASA-CORE as a multidisciplinary research center, and one of seven NASA University Research Centers at Historically Black Colleges and Universities (HBCUs) under the Minority University Research and Education Division at NASA Headquarters. These Research Centers, now a total of fourteen at HBCUs and Other Minority Institutions (OMIs), were established to foster new science and technology concepts, expand the nation's base for aerospace research and development, develop mechanisms for the increased participation by faculty and students at HBCUs in mainstream research, and increase the production of socially and economically disadvantaged students with advanced degrees in NASA-related fields.

NASA-CORE presently consists of three departments of the College of Engineering at North Carolina A&T State University (NCA&TSU): Mechanical Engineering, Electrical Engineering and Industrial Engineering. NASA-CORE is the first of its kind at North Carolina A&T State University in terms of the scope of research and the number of participating faculty and students. The mission of NASA-CORE is to educate and train socially and economically disadvantaged students in Aerospace Engineering and Research, and also enhance opportunities for socially and economically disadvantaged faculty who are US citizens in Aerospace Engineering and Technologies profession.

The Center for Aerospace Research conducts interdisciplinary research to develop techniques, and analytical and design tools that will aid in the design of next generation supersonic aircraft, hypersonic vehicles and spacecraft. Five interdisciplinary research groups: Aerospace Structures, Controls and Guidance, Computational Fluid Dynamics (CFD), Human-Machine Interactions and Propulsion conduct innovative research for the development of new technologies in spacecraft and high speed aircraft design. The goals of NASA-CORE are to establish a strong research capability to advance the state-of-the-art in aerospace research and development, establish a quality aerospace education program through curriculum and laboratory development, and to develop a robust systems engineering tool with a strong interdisciplinary research focus in support of NASA's mission.

The Proceedings of the technical papers and presentations of the students of the participating fourteen institutions at the First National Student Conference manifest the scope and quality of the research being conducted at these institutions through the auspices of NASA. The support from NASA, therefore, gives students and faculty at HBCUs and OMIs the opportunity to participate and contribute to the aerospace research and development base of the nation, and fosters new science and technology concepts and "fills the pipeline" by increasing the production of socially and economically disadvantaged students with advanced degrees in NASA-related fields. It is a unique opportunity for HBCUs and OMIs, NASA and the nation.

Endwell O. Daso

Message from the Organizing Committee Chairman

The First National Student Conference was born out of a need to provide a forum in order to discuss and exchange ideas and provide opportunities for the students who are participating in the research activities of the fourteen member institutions of NANURC. This Conference, therefore, has provided the participating students the opportunity to present the findings of their NASA funded research to their peers and the professional community, thereby preparing them for rewarding professional careers in aerospace related fields.

Every successful conference requires a considerable amount of planning in terms of organization and logistics. As such, many dedicated people have worked diligently to ensure the success of the First National Student Conference and I am indeed indebted to them for their tireless efforts and support. May I first acknowledge the support and advice of the Directors of the member Centers of NANURC. My sincere gratitude to Dr. Joseph A. Johnson, III, for his unwavering support and encouragement, and for his leadership in the planning of the Conference, as the Interim Chairman of NANURC. I also thank Dr. Doyle Temple for his support and advice.

At North Carolina A&T State University, let me first express my sincere gratitude to all the students and research faculty and staff of NASA-CORE for their diligent efforts and support. To the Conference Program Committee: Drs. Suresh Chandra, Abdollah Homaifar, Celestine Ntuen, P. Frank Pai, Meldon Human and William J. Craft I thank you very much for your contributions and guidance. In addition, I would also make a particular mention of the Conference Advisory Committee: Drs. Suresh Chandra, Abdollah Homaifar and Celestine Ntuen for their critical review of the daily progress of the planning of the Conference activities. The support staff: Ms Peggy Brooks, Mrs. Linda Dillard, Ms. Jocelyn Foy and our undergraduate student assistant, Ms. Stacie Mebane worked tirelessly to ensure the daily progress of the planning of the Conference activities. I am very grateful. My sincere thanks to Ms. Mebane for the many hours she devoted to the preparation of the Proceedings. Other offices in the University also provided considerable amount of time and resources to the conference and I thank everyone for their support and/or participation. My special thanks to Mr. Dennis Selby for his selfless dedication to this effort.

I am also very grateful to Dr. Edward B. Fort, Chancellor, North Carolina A&T State University. Your support and active participation underscored the interest of the University and accentuated the significance of the Conference. Thank you very much. My personal thanks to Dr. Harold L. Martin, Vice Chancellor, Academic Affairs, for his personal interest and steadfast support. To Dr. Earnestine Psalmonds, Vice Chancellor for Research and Dr. John C. Kelly, Jr., Associate Dean of the College of Engineering, my sincere thanks for your participation.

I would also like to acknowledge the support and critical role of NASA Headquarters and Field Installations. The Conference was funded by NASA Headquarters and the support was very much appreciated. My sincere thanks to Ms. Bettie White, Dr. Philip Sakimoto, Dr. Isaiah M. Blankson and Dr. Woodrow Whitlow, Jr. for their support and encouragement. To the many other people from NASA, DoD (Wright Laboratories) and the aerospace industry, thank you very much for your support and participation. To the Session Chairs I am very grateful.

Lastly, I am particularly grateful to all the invited guests and speakers, especially the keynote speaker, Dr. Yvonne Blanchard-Freeman of Clark Atlanta University and the closing luncheon speaker, Dr. Julian M. Earls of NASA Lewis. Your speeches were invigorating and challenging, with reassuring message of hope and the prospects of a better tomorrow for our students. Thank you very much for taking the time. I commend the students for their excellent presentations, which translates into the quality of their research. As resonated throughout the exciting speeches, the future is bright if you continue to worked diligently.

Endwell O. Daso

Organizing Committee

NANURC Student Conference Committee
Dr. Joseph J. Johnson, III, Center for Nonlinear and Nonequilibrium Aerosciences
Florida A&M University
Dr. Doyle Temple, Research Center for Optical Physics, Hampton University
Dr. Endwell O. Daso, Center for Aerospace Research, North Carolina A&T State University

Program Committee, North Carolina A&T State University
Dr. Endwell O. Daso, Chairman
Dr. Suresh Chandra, Mechanical Engineering
Dr. Abdollah Homaifar, Electrical Engineering
Dr. Celestine Ntuen, Industrial Engineering
Dr. P. Frank Pai, Mechanical Engineering
Dr. Meldon Human, Mechanical Engineering
Dr. William J. Craft, Mechanical Engineering

Advisory Committee, North Carolina A&T State University
Dr. Suresh Chandra, Mechanical Engineering
Dr. Celestine Ntuen, Industrial Engineering
Dr. Abdollah Homaifar, Electrical Engineering

North Carolina A&T State University
National Alliance
of University Research Centers

Keynote Address
Dr. Yvonne Blanchard-Freeman
Provost and Vice President for Academic Affairs
Clark Atlanta University

April 1, 1996

Discernment, Discipline, and Disciplineship: Visionary Leadership... Why We Can't Wait

Good Evening, I want to thank you for inviting me to be with you this evening, to visit with friends, to exchange witticisms, and to look upon the future of this great, but woefully misguided nation. The future of this country is indeed our children. I pray everyday in every way that they grow strong, mentally, physically, psychologically and spiritually so that they may lead with harmony and effectiveness.

To the young people in the audience, you must choose the proper venues to express yourselves and choose and use every opportunity to do so responsibly. I say responsibly because when I listen to some of the music to which you are exposed daily I think of the grand scheme to discredit you and all that so many before you have worked so hard to create and build so that your future would be better than that which my parents and grandparents and your parents and grandparents inherited.

Some of this music you listen to and are stimulated by defiles our history, our collective legacies, the infrastructure in our very existence. These legacies of which I speak are... the very legacies which give us reason to keep going... reason to keep fighting for equality... reasons to transfer to you young people knowledge, discretion, discipline and discipleship that will help you to turn this country right side up, to improve the quality of life for all Americans, and improve communication across communities, vast bodies of water and across continents.

I offer if there is hope, if there is any hope for us, for this country, it is vested in you our children. It is invested in the minds and souls of our youth. It is invested in the leadership in this room, who with or without resources continue to provide powerful, new and enriching opportunities for students.

Unfortunately, students often are not raised to know and appreciate the special effort it takes to make these experiences possible....If you do nothing else for the rest of the evening, I want you to think about how many ways and how many people to whom you may need to convey a thank you.

Those of us who work with young people know how utterly rewarding, ...and challenging this experience can be. Having done this for so long, it becomes a ministry.

I have had to develop special prayers as the challenges heighten, and... they heighten with each child, each young person, for each brings a number of special challenges with them. Some they recognize and some are invisible to them, some haunt them like ghost or demons, trying to distract, dissuade, dismay and discourage them from doing the right thing.

"Don't go there and don't allow any one else take you there. Do the right thing! Your gifts are too great and you are too wise." Don't mortgage your dreams. Don't allow anyone to steal your dreams. Just don't go there.

After many years of commitment to others I awakened one day and realized, in my zeal to share opportunities in math, science and engineering with countless others, my children were almost scientifically illiterate. So I thought... it was time for a reality check. So... I prayed, "God, help me not to be so busy helping other people's children that I neglect my own."

My prayers were answered. There is balance in my life, my children were not scarred, in fact, I found they had inculcated the teachings and the counsel that they observed and heard me sharing with others.

Pleasantly I learned, they had learned to manage adversity, to persevere, to not become discouraged but rather become resourceful, to not suffer silently but to seek help, to not be jealous in the sharing of their mother, but rather, be proud that there was enough of her to share.

In this process they have developed hundreds of red, black, brown, yellow and white brothers and sisters who still call our house home whenever they are, near, tired, troubled, broke or just in need of financial support or encouragement. What a marvelous legacy.

Pride is abundant. As abundant as each student, parent and faculty and

government and corporate official who is here today and who are identified with and who dignify themselves as a part of the core process or the enabling process which made this conference possible.

By now we have all learned that no branch bears fruit by itself, and at NASA, one of your benefactors, it takes a team to make a dream.

Let me say something about North Carolina A&T State University, your host institution. I began working with this Institution in 1980 as the Senior Manager and Director of Agriculture Research and Education Programs thru the U.S. Department of Agriculture.

I was instrumental in helping to develop legislation to assist NCA&T and 16 other Black Land Grant Colleges through the 1890 Land Grant Institution Facilities legislation introduced by then Congresswoman Shirley Chisholm and Congressman Kiki de La Garza in building the Alfreda Webb Animal Laboratory Science Building. Can you imagine the honor to work with two congressional giants and the 1st Black woman to run for President of the United States.

Every time I look at that building I think of Dr. Burliegh Webb a visionary educator and a man whom I love and respect enormously.

NCA&T exemplifies the visionary leadership this Institution attracts and claims. It translates that leadership into building its laboratories, faculty, curriculum, leadership team and its students. I trust that his appreciation for his "people treasures" will continue. And, if I know Dr. Ed Forte, it will continue because he recognizes a "Sacred Trust." Because I know he knows the world is watching him and his staff, leadership team and faculty and would in a moments notice invite their leadership to this campus.

Historically Black Colleges and Universities are undeterred in their mission to ensure that no child who would be educated, who desires to be educated is turned away.

Minority Universities have taught Discernment, Discipline, and Discipleship, all are qualities of Visionary leadership sorely needed for the 21st Century. I know because I am a product of a Historically Black College.

The importance of Black Colleges and Universities, Hispanic Serving Institutions and Tribal Colleges is affirmed daily as we observe that America is failing at

fairness. This failure is evidenced by so many things, chief among which are efforts to repeal Affirmative Action legislation and selective inclusion as evidenced in California with Proposition 187 which adversely impacts Hispanics and removal of any semblance of affirmative action from the laws of that state.

Institutions of higher education in America have failed to meet the aspirations and expectations of African American, Hispanic, Native American and other minority students. The total number of minority Ph.D. graduates from U.S. institutions of higher education is embarrassing and disgraceful. Historically Black Colleges and Universities, and Hispanic Serving Institutions, chiefly continue to be the primary source of African American, Hispanic and Native American graduates. These two facts underscore the importance of this conference.

If I may focus on HBCU's for a moment, the Historically Black Colleges and Universities, "have produced so much with so little that America thinks we can do almost anything with nothing."

Although HBCU's make up only 3 percent of all U.S. institutions of higher learning, they graduate approximately one-third of all Blacks with bachelors' degrees. In fact, more than one half of all Black professionals are HBCU graduates.

HBCUs generally have produced:
75% of all blacks with Ph.Ds
46% of all black business executives
50% of all black engineers
80% of all black federal judges
85% of all black physicians
50% of all black attorneys
75% of all black military officers.

In addition, historically black health professions and professional schools specifically, have trained an estimated:

40% of all black dentists
40% of all black physicians
50% of all black pharmacists
75% of all black veterinarians

Of the top ten undergraduate schools of any kind in the nation that send black students to medical school, the top five are HBCU's; Xavier, Morehouse, Spelman, Howard, and Hampton. We certainly know how to make excellent use of available resources, however limited they may be.

But we cannot rest on our laurels. We cannot ride on old news clippings. The number of African American, Hispanic and Native American Ph.Ds in physical sciences is to still too few. We have to find new ways to increase our capabilities for the production of intellectual capital.

We need to expand the capabilities of our HBCUs, HIS and Tribal Colleges for Ph.D. production, we need to strengthen the academic pipeline that feeds these institutions. We need to track the graduates of these institutions so that they are not lost, strayed, stolen or sold out!

This conference is one of the educational efforts for developing our capability for nurturing talents. It is an incubator for the development of college graduates in greater numbers who are fluent in science, engineering and technology. By your very presence here, you have already established yourselves as the scientifically elite in our community. You are cerebrally endowed for the demanding challenges ahead.

The responsibility of the elite is to actualize and operationalize yourselves and bring along those who are in a less advantaged position. You are what W.E.B. Dubois called the "talented tenth." I celebrate you, but, I beg you, emplore you to use this conference as your launching pad for exploring the universe.

There will be many challenges to keep you busy. One such challenge will be the long arm of human ignorance. Galieo Galilee, the great Italian scientist of the 17th century, faced ignorance and arrogance in his life. More than 380 years ago, after his landmark discovery of four moons circling Jupiter, an astronomer named Francisco Sizzi wrote, "Jupiter moons are invisible to the naked eye and therefore can have no influence on the earth, and therefore would be useless and therefore do not exist."

What a thought process! But as seekers of knowledge and truth, we as humans have almost always been up against a great wall of ignorance. The bottom line is, we don't know what we don't know.

Many of the greatest discoveries and technological breakthroughs have been

blind stumbles in the dark. The genius of many of those advances, however was that people had faith in themselves and a profound willingness to learn, despite the unknowns.

Today we can view the earth from space, but we understand precious little about what mechanisms drive our environment. We don't know whether the increase in atmospheric carbon dioxide we have measured will be absorbed by the oceans. We don't know if there was ever a seasonal hole of any size in the ozone layer over the Antarctic before we had satellites up there to measure it.

We know there is life on earth. What we can't agree on is the quality of life on earth. We don't know whether life was ever developed on Mars. We have plunged into the dense atmosphere of Venus and burned up a Pioneer Venus Probe which gave us new evidence that there may have at one time been oceans on Venus.

Did life once form there? We don't even know if there is life on planets outside our solar system, or even if there are planets outside those which we claim to know. So far we have only informed speculation and indirect evidence to support some of our theories. [1]

Space research now enables us to study healthy bodies and its reaction to a strange new environment. With gravity as a variable, we can peer into some of the mechanisms that drive the human body and better understand things like changes in bone mass which has implications for Osteoporosis; sterility from high speed flight; changes in muscle mass from zero gravity. Maybe somewhere in this is a cure Multiple Sclerosis or Epilepsy, which I have.

Science offers a unique window for innovative, invigorated, motivated students and faculty to carry out this work. The key to getting the most bang for the buck or the most return on the investment is to make sure that we continue to identify the widest possible array of individuals interested in scientific and mathematical disciplines for inclusion in a well trained multicisciplinary, skilled pipeline. That's you. You are the pipeline.

So in a world in which computers and television overwhelm us with wave after wave of information, true knowledge about our place in the cosmos is as elusive as ever. But every day, in some small way, with every space mission, and with

---

[1]Golden, Daniel - Evolution of Cultural Diversity at NASA, presented before the National Association for Equal Opportunities in Higher Education Conference. 1994.

every student we develop and grow and with every Ph.D. we support and enable, we roll back the boundaries of our ignorance just a little more.

But what we do have and what we can speculate is that the people who will find some of these answers are here tonight, they are you... You are the builders of pyramids... The architects of the new world order... the antidote for the poisonous social scorpions which destroys the equilibrium in our society.

You hold the sacred trust and that of developing and using your awesome gifts and talents to rebuild America and the world. You are a light and hope. Don't hide under the bushel of apathy. This conference is to make sure that you are on the right path, that your feet are on higher ground, that you succeed in your life's mission.

In line with what I expect of you, budding scientists of this nation and emerging leaders of the world, I would like to engage you in some serious reflections on the theme "Discernment, Discipline, and Discipleship: Qualities of Visionary Leadership for the 21st Century."

It used to be said, "when you look at the evening sky, you could predict weather for next day." Now, we rely on satellite imagery and remote sensing to give us a more reliable prediction of the weather. Whether we use folk wisdom or technological tools, we bring the totality of our past experience to a situation and make sense of it from the perspective of our experience. This is the nature of prediction. The process of making sense of what is presented to us is what I call discernment -- to grasp and comprehend what is not readily apparent.

How well we grasp and comprehend what is not readily apparent is determined to a great extent by our physical and mental readiness.

Archimedes was able to exclaim "eureka!" because he was absorbed in his search for an answer. He had the discipline of the mind. He was "ever ready," always prepared. We can discern the pattern, make sense of the chaos, and accomplish the seemingly impossible when we have discipline.

To attain this state of preparedness requires a period of discipleship, to school ourselves in a way of thinking and doing. Our training is so rigorous that what we learn becomes our second nature. Indeed, the words discipline and discipleship have a common root, that of a pupil, a learner.

A disciple is someone who gives herself/himself to the discipline of a study or a trade, to be molded and defined by the subject. In our present context, she/he is someone molded, defined and consumed by science, engineering and technology.

You are called to live, think, sleep, and walk as scientists, engineers, and technologists. It is certainly true that you are more than a scientist, engineer or technologist, but being a scientist, engineer or technologist is your sacred trust. It is given to you by your ancestors, your parents, your community and your personal interests and commitment.

There will be days when you will be so frustrated in your academic pursuits and the burden of life is so crushingly heavy that you seem to be in the "wilderness." Those are the days when you will be tested, just as Jesus and Buddha were tested prior to their enlightenment. From the trial, you will know your own strength and have a clearer sense of your vocation and your mission in life. Use your wilderness experience to your advantage. Lean on the Higher Power to give you the strength to face the next hour.

Make no mistake about this. Not everyone is as cerebrally endowed as you are. Not everyone can become a scientist, mathematician, engineer or technologist. By traveling this far, you have demonstrated your aptitude and fortitude for science, engineering, and technology.

Remember: "to whom much is given, much will be required!" The hope of your community and the nation rests with you; for the future will be dominated by science and technology; we already see this, holograms are everywhere; those who are disciplined and prepared will be the leaders of tomorrow. Architects of the new world order.

And we expect leaders of tomorrow to have the qualities of discernment, discipline, and discipleship.

It is easy for you at this stage of your academic career to recognize that you are a disciple. None of you is a bench scientist or a master technologist yet, but you will reach the peak of your profession if you persevere and have discipline. You must.

The question is not whether you are a disciple now but what kind of a disciple you are. I recall the story in the Bible, about the disciples who fell asleep while Jesus was praying. How often do some of us fall asleep while we should be

praying? Are you with me?

The challenge ahead is really the survival of the fittest. To have discipline is to make ourselves the fittest. Again, I remind you, few are given the opportunities that you have been given. We have the sacred trust to guard the gift that is ours - we ourselves, and you.

Recently, a friend was asked to say grace at lunch, and he began with the words, "we thank you Lord for each other, Amen." That was the end of the blessing, all that needed to be said. Lets give thanks for each other and the achievement and success of each other. Simple words and beautiful words. They recognize that other people are God's gifts to us. We need to appreciate ourselves, appreciate each other. This is the first sign of discernment. To know who we are and in what circle we are.

In our quest for scientific knowledge, some of us run the risk of forgetting our spiritual roots. By discarding the superstitious elements of our culture, some of us smother the spiritual dimension of our culture.

I believe we (black people, brown, red, yellow, white, disabled people and women) are people who are first and foremost a spiritual people. The spirit gives us power. We are in tune, constantly, with a Higher Power -- whether you name that power Yahweh, Allah, Sophia, or Tao, is not my emphasis here -- but a Higher Power that we know from our own experience that gives us insights and energies to accomplish our mission in life. Yet we call and acknowledge only when in trouble. "Oh Lord help me," always at the11th hour.

Frederick Douglas stated clearly in 1866, "our destiny is largely in our own hands, if we succeed in the race of life, it must be by our own terms and energies, and our own expectations. Others may clear the road, but we must go forward or be left behind in the race of life. If we are wasteful of time and money, the economy of other men and women will only make our destination that much more disgraceful and hurtful."

What Frederick Douglas said is the perfect segue to my next point about what kind of disciples we are. We are disciples who have discipline, to shape our destiny by our own hands. Discipline to make every minute count.

As former President Dr. Benjamin Mays of Morehouse College once said:

> I have only just a minute

Only sixty seconds in it.
Forced upon me -- can't refuse it.
Didn't seek it, didn't choose it.
But it's up to me to use it.
I must suffer if I lose it.
Give account if I abuse it.
Just a tiny little minute.
But eternity is in it.

You must be clear about what you determine for yourselves. Your foreparents paid their price for their determination. You can do no less. Use every minute of your life to contribute to something meaningful.

Never compromise your aptitude.

To live our lives of beauty, to be beautiful people, we must all eliminate the negatives and excuses in life. I was in a workshop many years ago. The facilitator gave each participant a token which she called "a round to it." She said, "Now, all of you have 'a round to it.' I do not want any more excuses for your not living up to your obligations. I don't want to hear any more about 'I will do it when I get around to it."

I do not have "a round to it" to distribute at this conference, but I do want you to get rid of excuses. They have no place in your lives, our lives or our personal equation to optimize or to change the world.

Shirley Chisholm, former Congresswoman, said, "You don't make progress by standing on the sidelines, whimpering and complaining. You make progress by implementing ideas."

Get rid of your excuses. No more "If only I had...," or "If only I could...," or "If only I were...." The really "if only" you should say is, "If only I would stop making excuses..." Regrets over the past, wishful dreaming about tomorrow, fretting about the confusion of the moment are all time wasting.

As I close, I say to you students, you are heirs to the builders of pyramid. Build!! Build with structural integrity. Build for tomorrow. Build to withstand the test of time. Build a more perfect union. Build to reunite a divided house, one nation under God, indivisible.

Don't be afraid to stand tall and tell all who ask, that your proclaimed height is

such because you stand on the shoulders of your ancestors who created for you this opportunity. Pride is abundant and so is your wisdom and intellectual prowess. Remember, go forward and make a difference! Redefine the world! Smoke your exams, and remember, the road of success is always under construction.

You have what it takes to be leaders of the 21st Century. You have what it takes to be leading scientists, engineers, and technologists. Your gifts and your talents are a sacred trust given to you by your Creator. Develop them. Use them. For the glory of your Maker and the benefits of the human race. Set a plan for yourselves. Spend quality time daily developing and implementing your plan. Do not let others distract you or allow business to delude you. What you believe, you become. What you believe, you can do.

Believe in yourselves! Focus on your goals! Don't compromise your ideals. Fortify your ambition. Be undeterred. Be proud.

Finally, an artistic perspective on life:

A kindergartner was drawing a picture.
Her teacher asked her, "What are you drawing?"
She replied, "God."
"But nobody knows what God looks like," the teacher said.
The little girl said, "they will, when I finish my drawing."

You too have an opportunity to be involved in drawing. You are drawing a picture of supreme excellence. When you complete your drawing, the world will know the majesty and magnificence of execllence. The world will know you.

May all the power of the univese sustain and carry you forward as you develop, cultivate and nurture the gifts of discernment, discipline, and discipleship, to assume leadership in the 21st Century.

Go forward and break the long arm of human ignorance, go forward and make this world a better place only as you can.

I thank you for inviting my expressions and to interact in the midst of such majesty - you are the future.

Overview of NASA and the URC Program*

Dr. Philip J. Sakimoto
Team Leader, Minority University Research & Compliance
Minority University Research and Education Division
Office of Equal Opportunity Programs
NASA Headquarters
Washington, DC

NASA

NASA is an investment in America's future. As explorers, pioneers, and innovators we boldly expand frontiers in air and space to inspire and serve America and to benefit the quality of life on Earth.

Those words from NASA's Strategic Plan form the Vision for all of NASA's activities. Our Mission is threefold:

- *To advance and communicate scientific knowledge and understanding of the Earth, the solar system, and the universe and use the environment of space for research.*
- *To explore, use, and enable the development of space for human enterprise.*
- *To research, develop, verify, and transfer advanced aeronautics, space, and related technologies.*

The expected outcomes from our work are:

- *Economic Growth and Security*
- *Preservation of the Environment*
- *Educational Excellence*
- *Peaceful Exploration and Discovery*

The NASA Strategic Plan lists a number of Values by which we operate. The very first of these Values is:

***People**. Our greatest strength is our workforce. We aggressively build a team of highly qualified individuals that is representative, at all levels, of America's diversity.*

To carry out our Mission, we have defined five key Strategic Enterprises:

- *Mission to Planet Earth*
- *Aeronautics*
- *Human Exploration and Development of Space*
- *Space Science*
- *Space Access and Technology*

It within these Strategic Enterprises that you are carrying out the research activities on which you are reporting at this conference.

University Research Centers

The NASA University Research Centers (URC's) at Minority Institutions began five years ago with the program goals of achieving a broad-based competitive aerospace research capability among the Nation's HBCUs and OMUs which would:

- foster new science and technology concepts,
- expand the Nation's base for aerospace research and development,
- develop mechanisms for increased participation by faculty and students of HBCUs and OMUs in mainstream research, and
- increase the production of socially and economically disadvantaged students, who are U.S. citizens and who have historically been underrepresented, with advanced degrees in NASA-related fields.

The first competition for NASA HBCU Research Centers was held in FY 1991. It resulted in awards to the 7 universities who

---

* The statements about NASA's Vision, Mission, Goals, Values, and Strategic Enterprises are taken from the *NASA Strategic Plan*, February 1996 Edition. This document is available on the World Wide Web at http://www.hq.nasa.gov/officce/nsp/NSPTOC.html

were designated as NASA HBCU Research Centers. The program was expanded in FY 1995 when a second competition for new starts was held for awards as Minority University Research Centers. This competition, open to both HBCU's and Other Minority Universities (OMU's), resulted in new awards to 4 HBCU's and 3 OMU's. Shortly thereafter, the HBCU Research Centers and the Minority University Research Centers were formally combined into a single program, the NASA University Research Centers (URC's) at Minority Institutions, with the two sets of awardees designated as Group 1 and Group 2, respectively.

The URC's present for this student conference are:

Group 1:

- **Clark-Atlanta University**, *High Performance Polymers and Ceramics Research Center*. Dr. Eric Mintz, Director.
- **Fisk University**, *Center for Photonic Materials and Devices*. Dr. Enrique Silberman, Director.
- **Florida A&M University**, *Center for Nonlinear and Nonequilibrium Aeroscience (CeNNAs)*. Dr. Joseph Johnson III, Director.
- **Hampton University**, *Research Center for Optical Physics (RCOP)*. Dr. Doyle Temple, Director.
- **Howard University**, *Center for the Study of Terrestrial and Extraterrestrial Atmospheres*. Dr. Arthur Thorpe, Director.
- **North Carolina A&T State University**, *Center for Aerospace Research*. Dr. Endwell O. Daso, Director.
- **Tuskegee University**, *Center for Food Production, Processing and Waste Management for Controlled Ecological Life Support Systems*. Dr. Walter Hill, Director.

Group 2:

- **Alabama A&M University**, *Center for Hydrology, Soil Climatology and Remote Sensing*. Dr. Tommy L. Coleman, Director.
- **Morehouse School of Medicine**, *Space Medicine and Life Sciences Research Center*. Dr. Gary L. Sanford, Director.
- **Prairie View A&M University**, *Center for Applied Radiation Research*. Dr. Thomas. N. Fogarty, Director.
- **Tennessee State University**, *Center for Automated Space Science*. Dr. Michael R. Busby, Director.
- **The University of New Mexico**, *Center for Autonomous Control Engineering*. Dr. Mo Jamshidi, Director.
- **University of Puerto Rico Mayaguez**, *Tropical Center for Earth and Space Studies*. Dr. Rafael Fernández-Sein, Director.
- **University of Texas at El Paso**, *Pan American Center for Earth and Environmental Studies*. Dr. Scott A. Starks, Director.

Office of Aeronautics
National Aeronautics and Space Administration

NASA Aeronautics

- Preparing the Engineering Student for the 21st Century.

Dr. Isaiah M. Blankson

Aerospace Research Division, Code RT

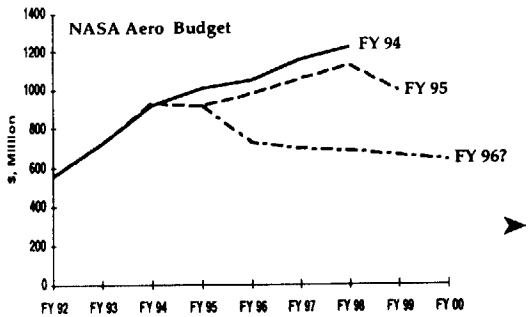
APRIL 2, 1996

OUTLINE OF PRESENTATION

- NASA AERONAUTICS -- quo vadis?
  - Restructuring for the future
  - some technical challenges and needs
- UNIVERSITY RESEARCH
  - New Paradigms (making innovative uses of invested $'s)
- ENGINEERING EDUCATION
  - Preparing the Engineering Student for the 21st Century

Aeronautics Enterprise Challenge

- Achieve the National Goals...



- ➤In a Severely Constrained Budget Environment

Approach to Meeting the Challenge

- Maintain Commitment to On-Going Focused Programs
  - ➤ High Speed Research, Advanced Subsonic Technology, High Performance Computing and Communications
- Re-Structure the R&T Base Program
  - ➤ Align with National Needs
  - ➤ Increase Program Responsiveness
  - ➤ Utilize "Aerocentric" Management Approach
  - ➤ Revitalize Flight Research as an integral element of the Program
- Re-engineer the Management of the NASA Aeronautics Enterprise to Increase Efficiency and Capability
  - ➤ Aerocentric Context
- Step-Up to Major Unmet Needs
  - ➤ Global Aviation System Technology Needs - FY96 →
  - ➤ Vehicle Technology Needs - FY99 →
- Integrate our Efforts in a National Alliance to Achieve the National Goals

NASA Aeronautics

In the Past...

Aero Centers

Ames Research Center

Evolved Center Centric Infrastructure

Dryden Flight Research Center

R & T Base Program perceived as level of effort

Langley Research Center

Open-ended Programs

Lewis Research Center

Relatively Autonomous

Loosely Coupled

Currently and in the future

Aero Enterprise

Response to National Challenges and Need for Change

Central Management of Major Facilities:

- Wind Tunnels
- S & E Computing
- Structural Test Facilities
- Propulsion
- Simulators
- Aircraft and Flight Facilities

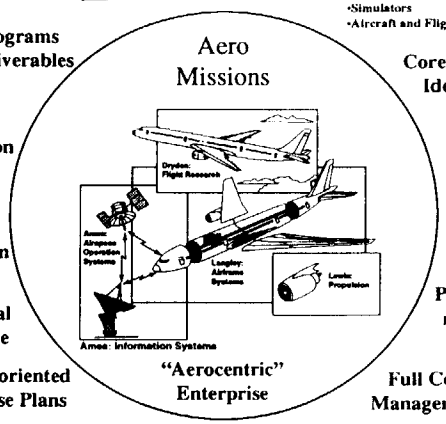
Focused programs with firm deliverables

Aeronautics Reorganization

Peer review/ peer evaluation

National Alliance

Objective oriented R & T Base Plans



Core Competency Identification

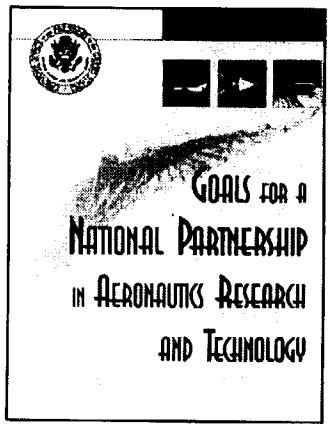
Minimum necessary infrastructure through Project Reliance

Principles - will maintain core competency

Full Cost Management

Strongly Coupled

Goals for a National Partnership



- Vision
  - World leadership in aircraft, engines, avionics, and air transportation system equipment for a sustainable, global aviation system
- Goals
  - Maintain the superiority of U.S. aircraft and engines
  - Improve the safety, efficiency, and cost effectiveness of the global air transportation system
  - Ensure the long-term environmental compatibility of the aviation system

Goal 1 / Superior U.S. Aircraft

- Major Issues
  - Competitiveness and Affordability
- NASA Role
  - Long-term investments in basic and applied research that enables commercial aeronautical innovation
    - Private disincentive to invest due to:
      - High-risk, long-term nature of investment
      - Difficult for a single firm to fully capture benefits
      - Subsidized foreign competition increases investment risk
  - Provision of National Facilities
    - Large-scale, 'generic' facilities in support of government and industry needs
    - Lowers barriers to entry and increases competition
- Current Major Programs
  - High-Speed Research, Phase 2
  - Advanced Subsonic Technology
- Major Unmet Needs
  - Affordable Design, Manufacturing, and Certification
- System Benefit Framework Elements
  - Performance, Efficiency and Affordability, and Survivability

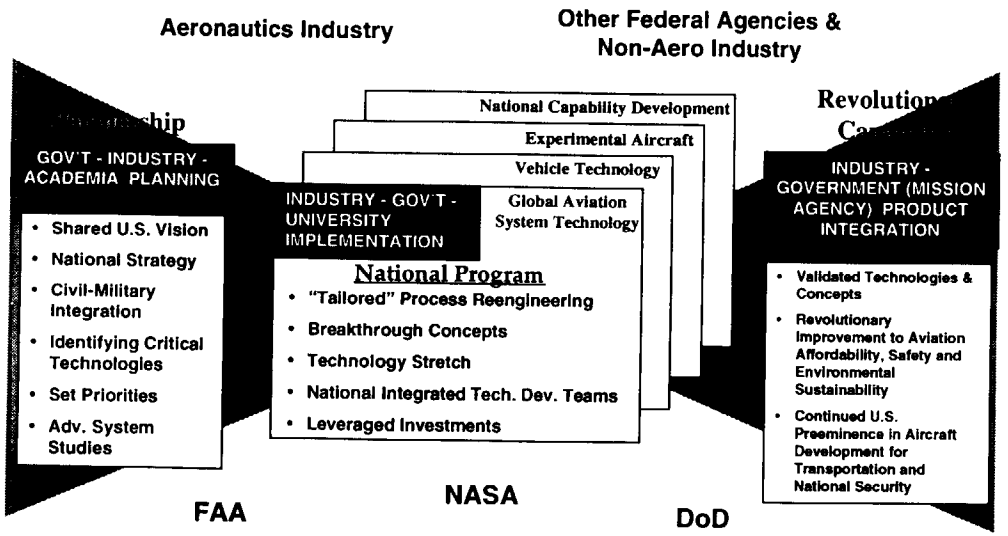
Goal 2—Global Air Transportation

- **Major Issues**
  - Airspace system capacity, user flexibility & predictable access to airspace
    - Technology opportunities enable innovative air traffic management approaches in emerging world regions
    - Investments in U.S. air traffic management infrastructure restricts introduction of new technology
  - Aviation Safety
- **NASA Role**
  - Long-term investments in basic and applied research to develop and demonstrate technologies and procedures for airspace system operations and aviation safety to enable commercial and government innovation
    - Requirement to operate & maintain U.S. air traffic management system forces incremental applied research by FAA
    - NASA maintains core competencies to respond to national needs in aviation safety (i.e., wind shear, aging aircraft)
- **Current Major Programs**
  - Aeronautics Base Research & Technology
  - Advanced Subsonic Technology
- **Major Unmet Needs**
  - Technologies to enable system transition and to support the DOT goal of zero accidents
- **System Benefit Framework Elements**
  - Capacity & Efficiency; Safety & Security

Goal 3 / Environmental Compatibility

- Major Issues
  - Community Noise and Sustainable Global Aviation, Market Limitations relative to Growth and New Vehicle Opportunities
- NASA Role
  - Long-term investments in basic and applied research that informs the environmental policy-making and regulatory processes and enables commercial aeronautical innovation
    - NASA is the key technical agency in a unified regulatory-R&D approach led by the FAA to assess environmental concerns, plan R&D, shape technical requirements, identify feasible technologies, and implement aircraft and engine certification regulations to mitigate the potential impacts. In this role, NASA:
      - Develops noise reduction technology, and, joint with FAA, noise abatement flight profiles and noise analysis/impact models
      - Performs research to define atmospheric impact of aircraft emissions and develops low emission engine technology
- Current Major Programs
  - High-Speed Research, Phase 1
  - Advanced Subsonic Technology
- Major Unmet Needs
  - Continuing needs across vehicle classes, Community acceptance of GA & Rotorcraft
- System Benefit Framework Elements
  - Environment

Context for a National Alliance

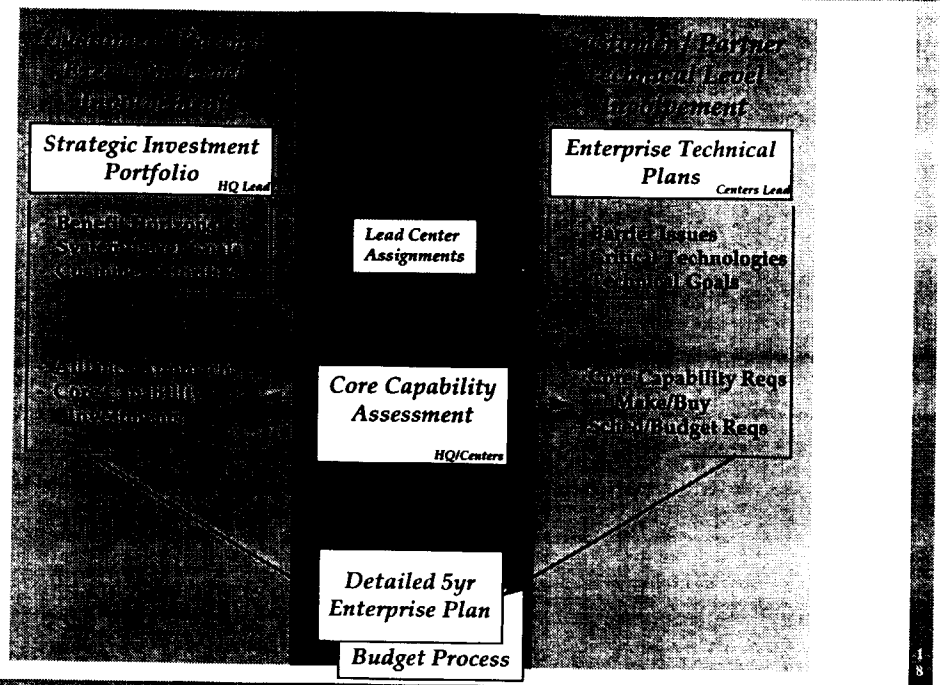


Strong Technology Base *High Leverage from Existing Programs*

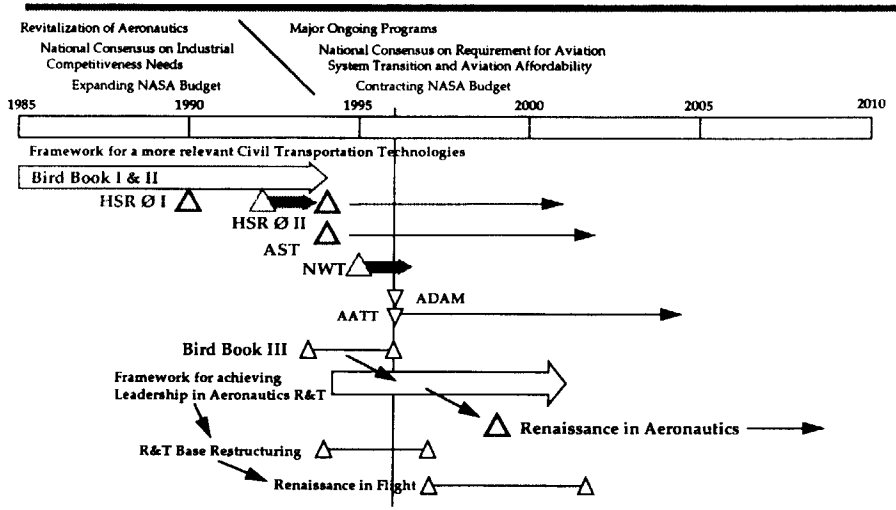
Missions, Centers of Excellence & Program Lead Center Assignments

Center:	ARC	DFRC	LaRC	LeRC	*Program Lead Center*
Mission:	Aviation Ops	Flt Rsrch	AF Sys / Aero / Atmos Science	Aeropropulsion	
COE:	Info Tech	Atmos Flt Ops	Structures & Materials	Turbomachinery	
Competency Group Lead:	Human Factors ARC	Experimental Aircraft DFRC	Airborne Systems LaRC	Subsonic LeRC	HSR / LaRC
	Air Transport Mgt ARC	Test Bed Research Prog DFRC	Structures & Materials LaRC	Supersonic LeRC	AST / LaRC
	Scientific Computing ARC	Flight Instr. & Test Tech DFRC	Aerodynamics LaRC	Hypersonic LaRC	HPCC / ARC
	Communications & Networking ARC		Mission / Sys Analysis LaRC	Propulsion Support Tech LeRC	NAS / ARC
	Intelligent Systems ARC		Crew Station Design & Integ LaRC		Airframe Sys LaRC
	Infrastructure Technologies ARC				Prop Sys LeRC
					Aviation Ops ARC
					Flt Rsrch DFRC
					Info Tech ARC
					Rotorcraft ARC

Integrated Strategic Planning Process



Transition in Strategy for Aeronautics R&T



- HSR and AST New Starts in FY94 were culmination of Aero Strategy initiated in early-mid 80's based on Administration policies set out in Bird Books I & II
- Proposal for New Wind Tunnels as a National Initiative to address productivity improvements for improved competitiveness in industry design processes
- Recognition that a new strategic framework is needed to guide future Aeronautics directions (Bird Book III) and embracement of reengineering principles

Summary

- NASA Aeronautics has developed and is implementing a strategic management approach
  - ➤ Provides framework for:
    - Future decision-making for meeting aeronautics needs now and into the future
    - Rational, integrated decision-making for budget reductions
- NASA Aeronautics is accelerating efforts to develop a forward-looking, aggressive strategic plan that will guide investments within the framework

DESIRED ATTRIBUTES OF AN ENGINEERING GRADUATE

- **A good understanding of engineering science fundamentals.**
  - **Mathematics (including statistics)**
  - **Physical and life sciences**
  - **Information technology (far more than "computer literacy")**
- **A good understanding of design and manufacturing processes (i.e., understands engineering)**
- **Possesses a multi-disciplinary, *systems* perspective.**
- **A basic understanding of the *context* in which engineering is practiced.**
  - **Economics**
  - **History**
  - **The environment**
  - **Customer and societal needs**
- **Good communication skills.**
  - **Written**
  - **Verbal**
  - **Graphic**
  - **Listening**

DESIRED ATTRIBUTES OF AN ENGINEERING GRADUATE

- High ethical standards
- An ability to think both critically and creatively - independently and cooperatively.
- Flexibility--An ability and the self-confidence to adapt to rapid/major change.
- Curiosity and a desire to learn - for life.
- A profound understanding of and commitment to team work.

Changes in Educational Needs Resulting from Industrial Shift

Pre - 1990

Classical Engineering Emphasis

- Aerodynamics
- Mechanical
- Electrical/Electronic
- Computer Science/Engrg
- Materials Sciences
- Industrial Engineering

Classical Business Management

Basic Communication and Dexterity Skills for Factory Workers

Post - 1990

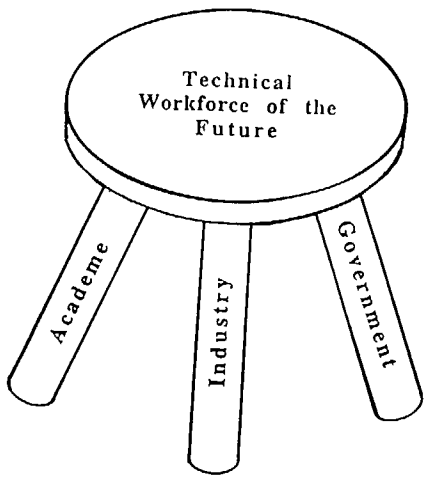
New Engineering Curriculum Emphasis

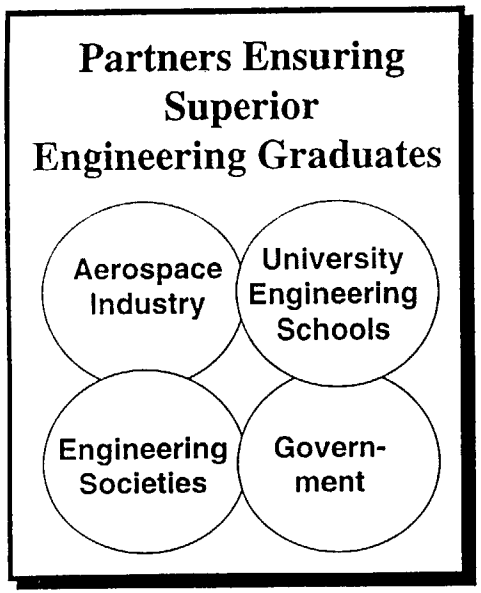
- Product Development Processes and Modeling
- Information Data Management
- Design for Manufacturability
- Manufacturing Processes with New Material
- Manufacturing Systems

New Management Processes for Highly Integrated Manufacturing Processes

Additional Skills in Math and Sciences, and Understanding of Computers

NEW PARTNERSHIPS FOR ENHANCING
ENGINEERING EDUCATION





Aerospace Education Partnership Workshop Results

10/11 Jan 1996

People Quality - A Key to Success

"There is one prediction we can make about tomorrow with certitude. Companies will be absolutely dependent on the quality of people. We will need well-educated, skilled, creative team members to guide us to innovative solutions. More than market share, more than reputation, more than superior resources, our STRENGTH will lie in our people. We must have the best and the brightest to succeed. We are doomed to fail if we settle for less."

Quote taken from the book "The World in 2020: Power, Culture and Prosperity" by Hamish McRae published by the Harvard Business School Press, 1994

University Strategy Team

Approach

- The University Strategy Team will lead the Headquarters effort and keep the task force on track and on schedule through internal Headquarters team meetings as well as interactions with the task force via telecon, e-mail, fax, and possible face-to-face meetings

- The University Strategy Team will review university research strategies employed by other Government agencies and NASA's industrial partners for comparison to NASA's Aeronautics Enterprise policies and practices.

- The University Strategy Team will develop a database of strategic university research (current, future, dual-use, spectra of applications, success measurement criteria, etc) in all disciplines and in a Mach number range pertinent to the interests of the Aeronautics Enterprise.

University Strategy

Special Needs: **RESEARCH AND EDUCATION**

The University Strategy Team will also explore new processes for:

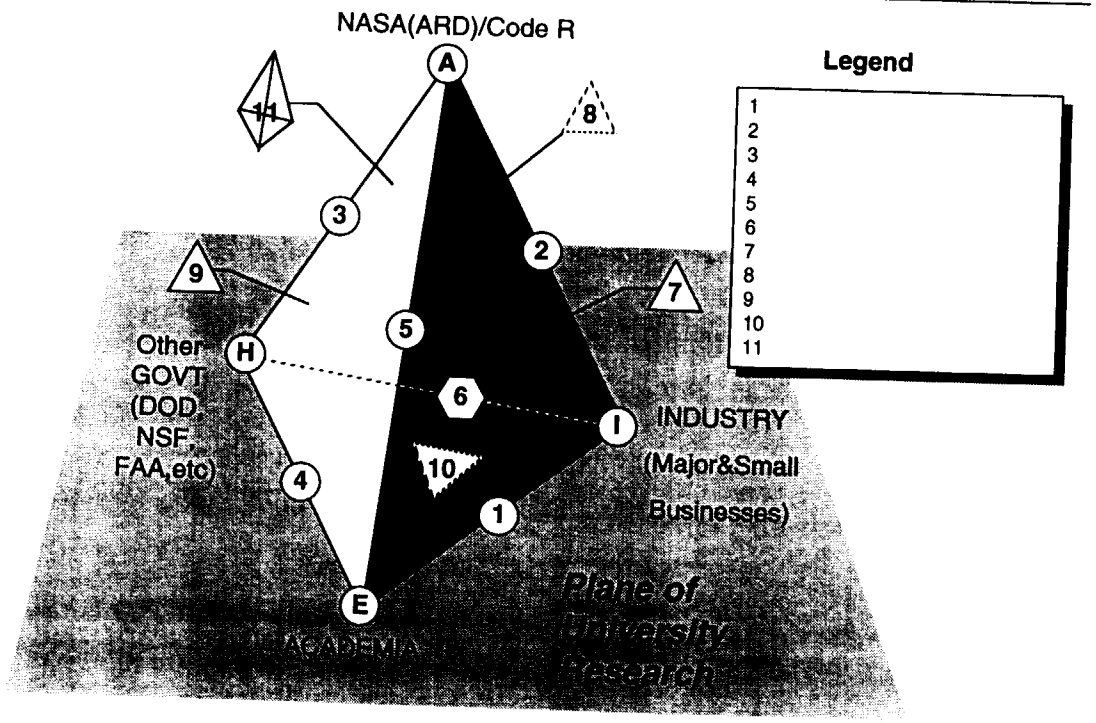
- **INNOVATIVE STRATEGIES FOR UNIVERSITY PARTICIPATION IN NASA AND INDUSTRY PROGRAMS**
- **ENGINEERING EDUCATION LIAISONS**
- **$ APPROACHES: Direct Funding, CRADAS(Standard, Non-standard), Special Cooperatives (with DoD/Industry, NSF, etc), SBIR's, etc.**
- **SPECIAL APPROACHES: Industry/NASA Mentor Programs for Universities, Leveraged IRAD's, etc**
- **SAFEGUARDING RESEARCH OPPORTUNITIES WHICH: Arise during middle of funding cycle (i.e. unscheduled); Lack "built-in advocacy" at NASA centers; No expertise exists at NASA centers, etc**
- **STRATEGIC ORGANIZATION FOR NEW MULTIDISCIPLINARY RESEARCH**

University Strategy - *Rationale*

Rationale

Three principal changes in the research environment have created the need for this review of NASA's university strategy. Firstly, due to differences in the administrative environment and the research programs at each Center, the implementation of NASA's broad agency policies vary. These differences have grown over time to the point that significiant differences exist and many university researchers have become confused as to how to interact with the NASA Aeronautics program. Secondly, the funding available to accomplish Aeronautical research, both in the Government and in industry, dictates that the funding that is available be applied in the most effective manner possible. Thirdly, the structure of the Aeronautics program has evolved to reflect a strategy for the Research and Technology (R&T) Base. Since University research grants are key to the success of this new strategy, it is essential that the policies and procedures for their application to research programs be reviewed and restructured.

NASA University Programs Strategy



NASA OFFICE OF HUMAN RESOURCES AND EDUCATION, EDUCATION DIVISION
PROGRAMS WITH RESEARCH CAPACITY BUILDING COMPONENTS

DR. ELIZABETH B. WARD
NASA SPACE GRANT FELLOW
NASA LANGLEY RESEARCH CENTER
HAMPTON, VA

The NASA Office of Human Resources has made a strong effort to increase the research capability of typically non-research intensive institutions. The three principal efforts have been through the National Space Grant College and Fellowship Program's Capability Enhancement Grants, the NASA Experimental Program to Stimulate Competitive Research (EPSCoR), and the NASA/University Joint Venture (JOVE) program.

With partners from academe, industry, and state and local government, NASA's capacity building programs reach wide segments of the population. These cornerstone programs provide a three dimensional approach for serving the American public. Collectively, the programs interface with all five of the NASA Strategic Enterprises: Aeronautics, Human Exploration and Development of Space, Mission to Planet Earth, Space Science and Space Technology. Separately, the three programs employ different strategies to stimulate research by faculty and students at colleges and universities that previously have not received substantial NASA funding.

Within a given state, the three programs reach a wide range of institutions, faculty, and students. The programs facilitate research building strategies such as the competitive awarding of faculty research grants, the involvement of students in the research process, and the publication of research findings in refereed journals. All three programs emphasize the involvement of underrepresented groups and all three have a matching funds requirement. The Space Grant and JOVE programs both have an outreach component, specifically to the precollege and general public sectors. NASA EPSCoR does not contain an education or outreach component.

THE NATIONAL SPACE GRANT PROGRAM

The National Space Grant Program contributes to the nation's science enterprise by funding research, education, and public service projects through a national network of 52 university-based Space Grant consortia. These consortia administer programs in all 50 states, the District of

Columbia, and Puerto Rico. The consortia have over 580 affiliates that include over 400 academic institutions. Other partners include business and industry, state and local government agencies, other federal agencies, and nonprofit organizations. In 1996, Space Grant administers 805 outreach programs, 790 higher education programs, and 417 research programs.

During the past six years, Space Grant awarded over 8000 US citizens with tuition assistance in science, engineering, and related fields of study. For 1995, individual student awards averaged $2500 for undergraduates and $8,000 for graduate students. Twenty two percent of these awards were to students from underrepresented groups; thirty nine percent were to women. The majority of Space Grant student awards include a mentored research experience with university faculty or NASA scientists. Space Grant funds curriculum enhancement and faculty development as well. Space Grant colleges and universities also administer precollege and public service education programs that help to meet the education needs of their states.

The Space Grant consortia have leveraged federal funds to more than double the Space Grant budget with matching contributions from state and local sources. Space Grant encourages collaboration among departments, across institutions, and with business and industry. All Space Grant programs emphasize the diversity of human resources, the participation of students in research, and the communication of the benefits of science and technology to the general public.

NASA EXPERIMENTAL PROGRAM TO STIMULATE COMPETITIVE RESEARCH (EPSCOR)

NASA initiated the NASA Experimental Program to Stimulate Competitive Research (EPSCoR) in 1994. This program targets states of traditionally weak research infrastructure with funds to develop a more competitive research base within their member academic institutions. The research projects are closely related to one of NASA's five strategic enterprises: Aeronautics, Human Exploration and Development of Space, Mission to Planet Earth, Space Science, or Space Technology. The EPSCoR project research clusters collaborate frequently with NASA's field installations and NASA Headquarters. NASA EPSCoR fosters cooperation among departments, across institutions, with state and local government, and with business and industry. In addition to strengthening the research base of the state's academic institutions, EPSCoR outcomes include the transfer of developed technology to industry, the increased economic development of the participating states, and the communication of the benefits of research to the public. The goal of NASA EPSCoR is to produce strong academic

research enterprises that are long-term, self-sustaining, and nationally competitive for non-EPSCoR research dollars.

After only two years as EPSCoR grantees, the six NASA states have collectively accomplished the publication of 225 papers in peer-reviewed media; have acquired $62 Million in additional research grants from federal and non-federal sources; have been granted one patent and applied for many more; and have secured the participation of 42 institutions and hundreds of students and faculty. In addition to the results listed above, the six states have collectively proposed for another $15 Million in 1996 and have produced another 274 publications in the form of student theses, dissertations, conference presentations, and so on. The six NASA EPSCoR states have clearly begun their ascent to full competitive success in the national R & D enterprise.

Though there are many partnerships that have resulted from NASA EPSCoR projects, here is one example that illustrates the potential economic benefit. Louisiana EPSCoR university scientists have teamed with Specialty Plastics Inc. of Baton Rouge and NASA's Marshall Space Flight Center on a three-year $1.8 million Advanced Technology Program (ATP) grant from the Department of Commerce's National Institute of Standards and Technology (NIST). The partnership will develop innovative joining and fitting technologies for lighter, corrosion-resistant advanced composite piping that could save the U.S. offshore oil and gas industry billions of dollars. Replacing a portion of the heavy metal piping used to build offshore platforms with lighter weight composite material could reduce the nearly $1.2 billion building costs for an oil-rig by as much as $150 million. Other NASA EPSCoR consortia are working closely with Industry to develop dual use technology.

NASA UNIVERSITY JOINT VENTURE (JOVE) PROGRAM

The NASA/University Joint Venture program began in 1989 to develop aerospace research capabilities and to promote science and engineering education by means of NASA/university partnerships. JOVE is a NASA-sponsored research program designed to build research capabilities at U.S. institutions of higher education that have had little or no involvement in the Nation's space program. To qualify for participation in the JOVE program, a college or university must have less than $100,000 of annual research support from NASA. The university commitment to the joint venture includes granting release time for each of its JOVE Faculty Research Associates to conduct research during the academic year. A minimum of 25 percent full-time equivalent per faculty member must be obligated to the project to permit adequate research progress. The summer stipends of the

faculty are shared, with 50% contributed by the institution and 50%, up to a maximum of $6,000 per summer, contributed by NASA.

The participating universities are expected to use their JOVE research connection as a basis for curriculum development, the enhancement of student research potential, and outreach programs, which should be designed to serve students at that institution, K through 12 students, as well as individuals and groups in the broader community served by the JOVE institution. NASA will make money available to support student participation in the JOVE research. The student support is shared equally by NASA and the university up to an annual maximum of $8,000 from NASA. In return for university support of faculty time for research, and matching support for both student scholarships and faculty summer stipends, NASA selects faculty participants to become members of its varied research teams, makes data available, provides "mentoring" by a researcher at one of the nine NASA field centers or at a NASA sponsored center of research, provides funds for a workstation, and facilitates the establishment of a computer network connection to the campus to electronically link the Faculty Research Associates to the NASA mission databases.

The JOVE experience begins in the summer with each of the Faculty Research Associates spending 10 weeks in residence at their respective research sites. This period of research is equivalent to the NASA Summer Faculty Fellowship Program experience. The joint research, curriculum development, and outreach will begin in the fall time frame, after the Faculty Research Associates have returned to their home institutions. The commitment to the JOVE program by both NASA and the institution of higher education is for three years.

Current JOVE participants include 89 academic institutions, 164faculty, and approximately 330 students, both graduate and undergraduate. In 1995, the JOVE program universities leveraged the NASA investment of $3.0 million by raising almost $13 million in additional funds.

Contact Information

Space Grant/EPSCoR Program

Dr. Elizabeth B. Ward
NASA Space Grant Fellow
Mail Stop 221
NASA Langley Research Center
Hampton, VA 23681-0001
Phone 757-864-3299

Fax 757-864-9083
Email: e.b.ward@larc.nasa.gov

JOVE Program

Ms. Ernestine Cothran
Acting University Affairs Officer
Mail Code ES01
Marshall Space Flight Center
Huntsville, AL 35812
Phone 205-544-0649
Fax 205-544-9243
Email: ernestine.cothran@msfc.nasa.gov

The Boeing Educational Initiative

A Presentation to the NANURC First National Student Conference, North Carolina A&T State University

April 1, 1996

by

Dr. Gerald Paynter

Associate Technical Fellow

The Boeing Company

Current Aerospace Market

-Defense

- •Reduced
- •Highly competitive
- •New emphasis on low cost

-Commercial

- •Highly competitive
- •New aircraft must compete against upgrades of current
- •Low cost, fast delivery increasingly important

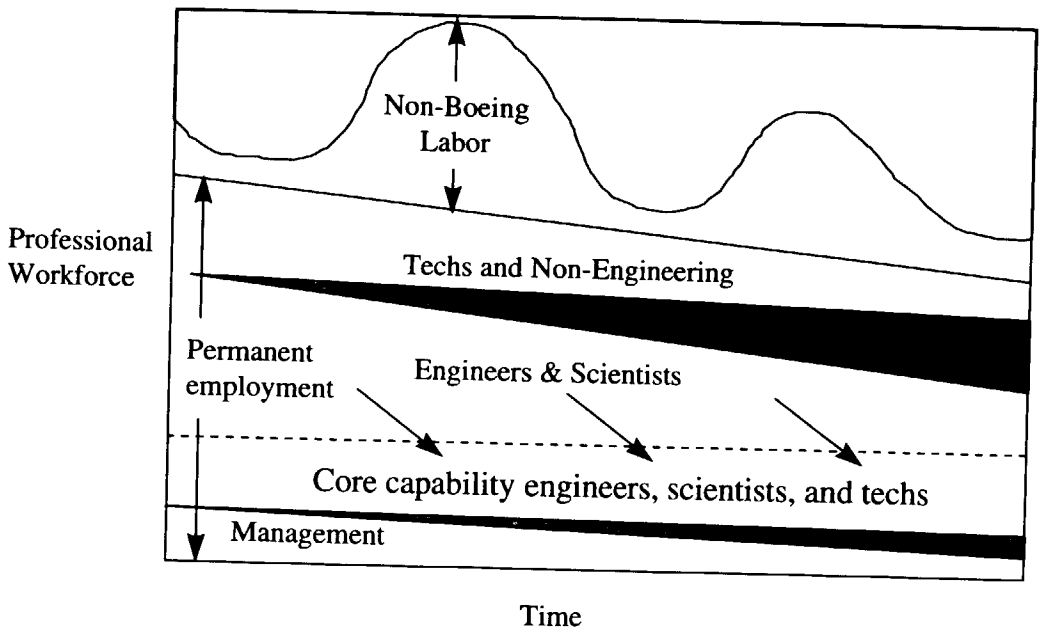
Boeing Goals

- Win increased share of the commercial & military markets
  - Reduce for a new aircraft
    - development time
    - cost
    - time from order-to-delivery
  - Maintain traditional technology strengths

Boeing Strategies

- DCAC/MRM
- New creative teaming arrangements
- Smaller, more skilled workforce
  - New attributes for future Boeing engineers
  - New emphasis on upgrading skills of current engineers at Boeing

Professional Workforce of the Future Model



BOING INITIATIVE ON ENHANCING ENGINEERING EDUCATION

- Traditional Boing-University Interactions
  - Research
  - Recruiting
  - Summer Intern/Coop Programs → BCAG Engineering Summer Intern program
  - Scholarships
  - Grants
- Cooperative Workforce Initiatives (1991)
  - Training
  - University Interactions and Recruiting → Eng./Ops. University Relations Coord. Committee
    - •Key School List
    - •Executive Focals
    - •.....
    - •New Initiatives
  - .....
- National Efforts to Reform/Enhance Engineering (& Math and Science) Education
  - ASEE
  - NRC
  - NSF
  - .....

→ Boeing-University Workshop Process (1994)

BOEING INITIATIVE ON ENHANCING ENGINEERING EDUCATION

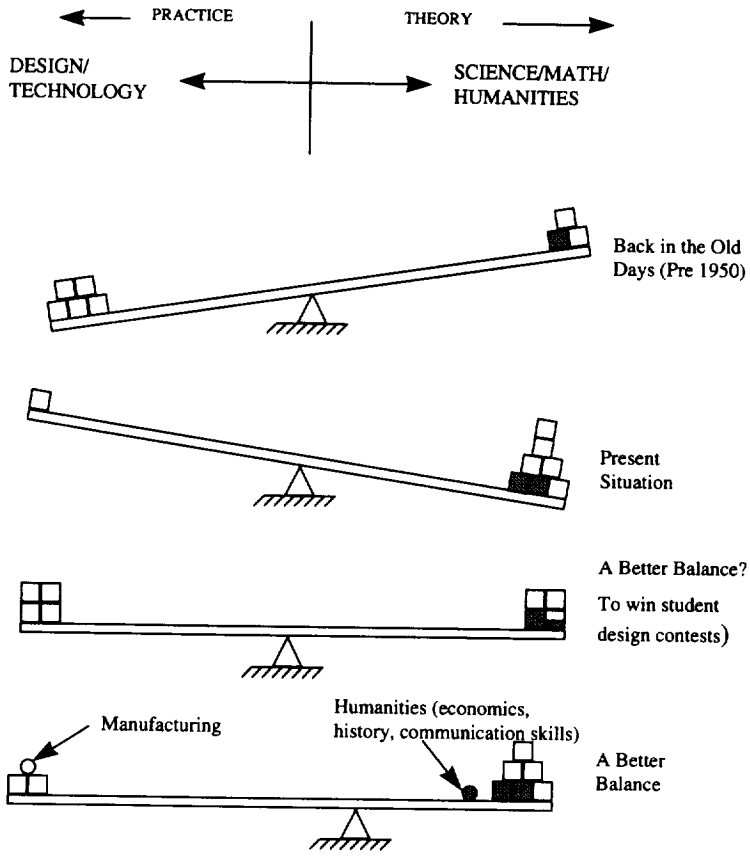
OBJECTIVES

- Maintain/Enhance Competence and Vitality of Boeing Technical Workforce
- Enhance and Sustain Durable Partnerships with Universities
- Catalyze Industry Support for National Engineering Education Reform Efforts

Background

- Engineering education "vocational" into 1950's.
- Swing towards science in the late 1950's.
- Now heavily dependent upon funded research.
  - Publish or perish.
  - Teaching brings in little money.
- Academe distrustful of industry's "mercantile mentality"
- Industry depends on the university but considers it "not in the real world."
- Somewhere engineering (design) education got lost.
- Inadequacies in current engineer education only part of a bigger problem.

Paradigms for Undergraduate Engineering Curricula



□ = one semester in a four year program

○ = additional semester in a five year program

■ = humanities

ATTRIBUTES OF AN ATTRACTIVE ENGINEER GRADUATE

(A Traditional Industry View)

- GRADUATE OF ACCREDITED PROGRAM AT A "RESPECTED" UNIVERSITY
- HIGH GRADE POINT AVERAGE (>3.0)
- TECHNICAL ABILITY
- WORKED WHILE IN COLLEGE
- EXPERIENCE (WORK, COOP, ETC.)
- EXTRACURRICULAR ACTIVITIES
- GOOD COMMUNICATION SKILLS
- ENTHUSIASM AND INTEREST
- GENERAL ATTITUDE AND NEATNESS

AN INDUSTRY [BOEING] ROLE IN ENHANCING ENGINEERING EDUCATION

- WIDE GULF BETWEEN ACADEMIA AND INDUSTRY
- SERIOUS IMBALANCE IN CURRENT ENGINEERING CURRICULA
  - Heavy Emphasis on 'Engineering Sciences" (Analysis)
  - Very Limited Exposure to Design (Creative Synthesis)
  - Heavy Emphasis on Competition
  - Little Emphasis on Teamwork and Cooperative Learning
  - Little Exposure to Manufacturing Issues or Practice
- MOST FACULTIES HAVE LITTLE KNOWLEDGE OF ENGINEERING PRACTICE
  - Rewords Come Via (Government) Funded Research
  - Too Much University Research of Limited Value of Industry
- CURRENT SYSTEM POSES A LONG-RANGE PROBLEM FOR BOEING AND AEROSPACE INDUSTRY
- WHAT CAN/SHOULD WE DO?

SKILLS AND ATTITUDE OF ENGINEERING GRADUATES

- To Promote Needed Changes In Engineering Education, Industry Must Deliver a Common Message to Academe.
- Industry Is a Major "Customer" for Academe.
  - *Should Define Desired Outcome of Education.*
  - *Should Not Tell Individual Schools How to Do Their Jobs (I.e. Develop Curricula).*
- The World Is Changing Rapidly - *Education* Must Concentrate on *Durable* Skills/Attributes.

WHAT WE NEED - SOME OPPORTUNITIES

- Recognition That *Engineering Science* Is Not *Engineering.*
- Recognition That Engineering (Design) and Manufacturing Are Inextricably Bound Together In Practice - And Should Be In The Education Of Any Future Engineer.
- Recognition That A University Education Can No Longer Be A One-Time Affair - Education Must Be A Life-Long Process.
- Industry and Academe Must became Cooperative Partners in The Education Of Engineers from The Beginning To Career End
- Universities Must Develop An Appropriate Hiring And Reward System For Engineering (Design) And Manufacturing *Educators.*
- Industry and Academe Must Come To New Understandings Regarding Provision Of Education And Training And Recognize The Future Needs of Both Large And Small Companies.

BASIC PREMISES

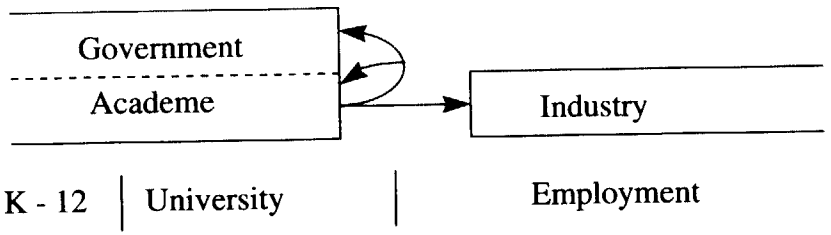
- Engineering Is About *Design*, Which In Turn Is About *Manufacturing* Things People Need and Want - *For Sale.*
- The U.S. Has The Best Graduate *Engineering Science* Education System In The World - *Engineering* Education Needs Help.
- You Can Not jam Everything A Student Needs To Know Into a 4-year (or 5-year or 10-year) Box - Education (and Training) Must Be A *Life-Long Process.*
- Universities and Industry (Aided By Government, Professional Societies, etc.) Must Form New partnerships - We Are *All* In This Together.

ROAD BLOCKS TO CHANGE

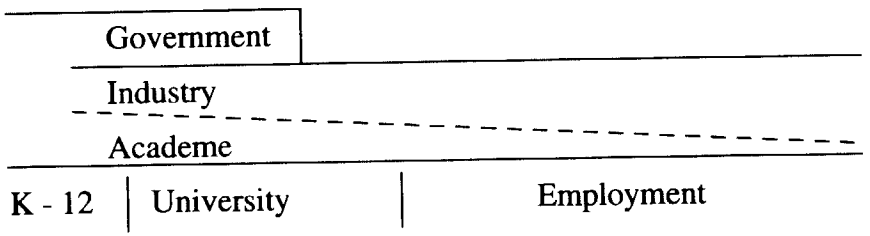
- Many Industry and Academe Now Agree That Change In Engineering Education Is Needed - But At What Rate? Academe and Industry Are Two Separate Cultures Operating On Two Separate Time Scales.
- Ignorance - Of Industry On The Part Of Many Faculty.
- Ignorance - Of Most In Industry Of How The Modern University Works (e.g. The faculty Reward System).
- Education Generates Very Limited Revenue And Few Prestigious Publications Compared To Research.

PARADIGM SHIFTS IN UNIVERSITY-INDUSTRY-GOVERNMENT RELATIONS

Current Paradigm



New Paradigm



WANTED

B.S. Engineering Graduate

"A Generalist With An Emerging Specialty
(or Two or Three)."

With The Desire and Ability to Learn for Life!

Desired Attributes of An Engineer

- A good understanding of engineering science fundamentals
  - Mathematics (including statistics)
  - Physical and life sciences
  - Information technology (far more than computer literacy)
- A good understanding of design and manufacturing processes (i.e., understands engineering)
- Possesses a multi-disciplinary, *system* prospective
- A basic understanding of the context in which engineering is practiced
  - Economics
  - History
  - The environment
  - Customer and social needs
- Good communicator
  - Written, Verbal, Graphic and Listening
- High ethical standards
- An ability to think both critically and creatively - independently and cooperatively
- Flexibility-an ability and self confidence to adapt to rapid/major change
- Curiosity and desire to learn for life
- A profound understanding of and commitment to team work

Note: This is a list of basic, durable *attributes* into which can be mapped specific skills reflecting the diversity of the overall engineering environment in which the professional practice operates. In specifying desired attributes (i.e.., desired outcomes of the *educational process*), we avoid specifying how a given university goes about meeting industry needs Curriculum development is viewed as a university task to be done in cooperation their "customers", and in recognition of their own local resources and constraints. Industry, as an important customer, must be an active partner in this process.

KEY BOEING PROGRAM STATUS

February 1996

- Boeing Student Summer Intern Program
- Boing Outstanding Educator Award
- A. D. Welliver Faculty Summer Fellowship Program
- [Boeing Fellows on Campus]

THE BOEING STUDENT SUMMER INTERN PROGRAM

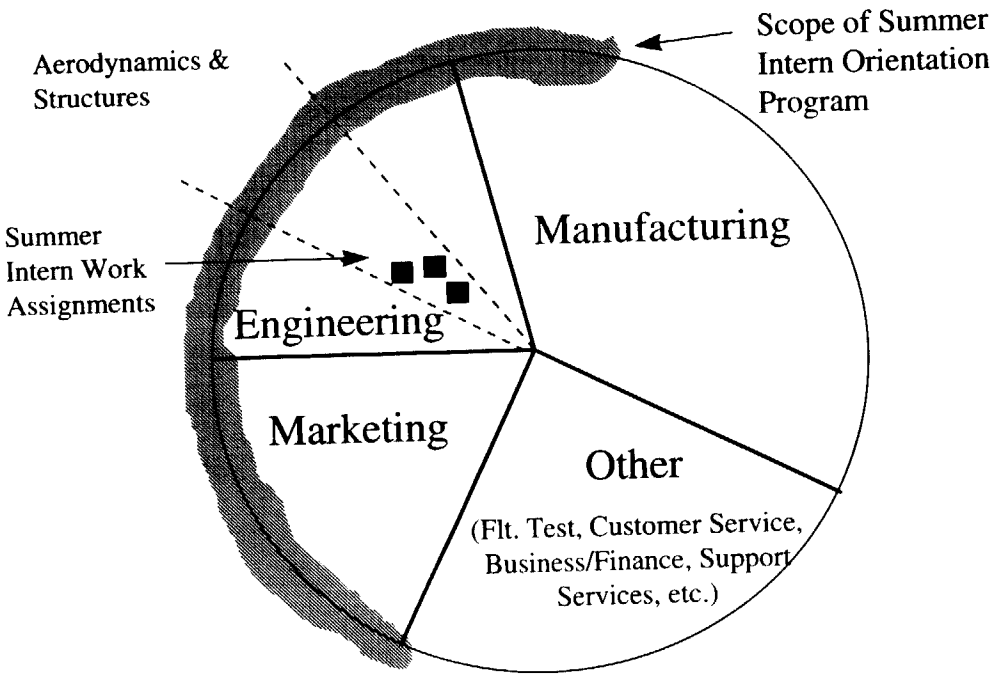
Program Objectives

- Demonstrate Boeing's commitment to diversified workforce
- Provide a comprehensive view of aerospace vehicle design, manufacturing and business processes at Boeing and within the overall aerospace industry
- Provide a broad overview of professional practice and career opportunities at Boeing
- Develop interns' skills in team-based problem solving
- Provide Boeing an opportunity to identify "stars" for possible future employment
- Develop interns' as "Boeing ambassadors" on campus

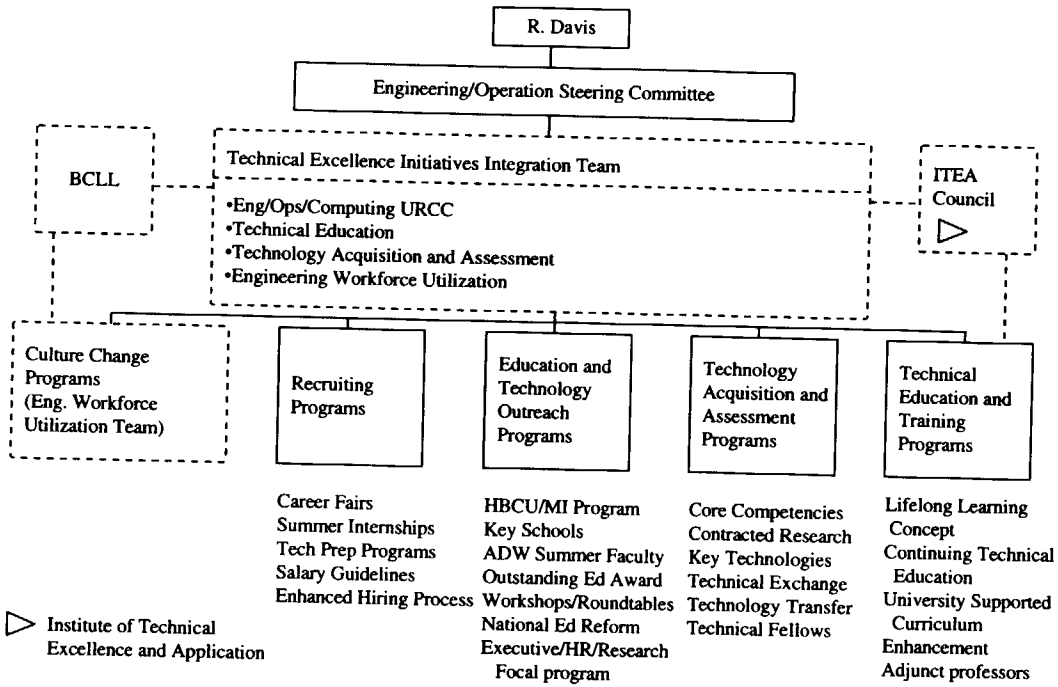
Summer Intern Program
(Seattle/Puget Sound, Washington)

- Recognized as one of top ten in 1994 and 1995 editions of America's Top 100 internships (a thousand programs were reviewed for this publication)
- Round trip air fair, relocation allowance and housing assistance service
- Summer intern classes
- Presentations, off-hour tours and other special events for interns
- Good benefits, including extensive off-hour training program

The Boeing Commercial Airplane Group

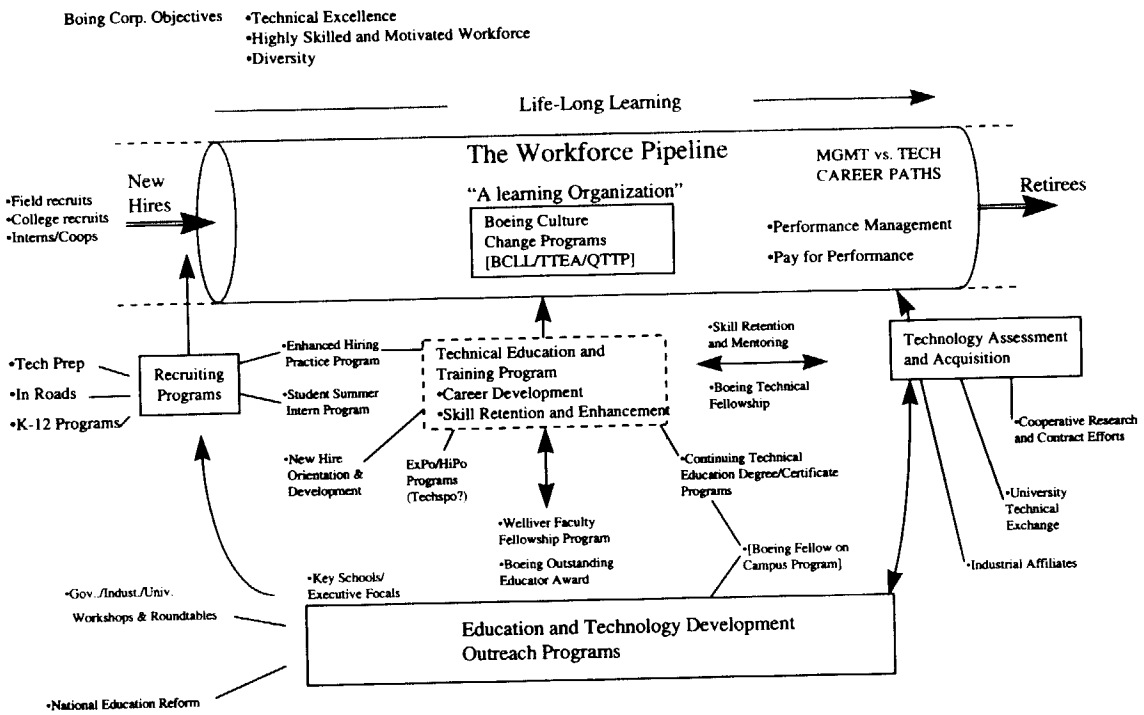


Proposed Working Relationship Infrastructure



John McMaster
February 1996

TECHNICAL EXCELLENCE BY DESIGN [=ENHANCING]



Survival Strategies that You May Not Learn In Engineering School

Austin A. Wolfrey
Principal Scientist
Raytheon Electronic Systems

- **Will your bachelor's or master's engineering degree guarantee success in the world of work?**
- **Will your degree guarantee a long career with a Fortune 500 company?**
- **Has your engineering degree availed you with the necessary skills that industry will demand?**

It is projected that less than 25% of you can expect to work for the same employer for more than 10 years.

WHAT TO TALK ABOUT

A Shopping List:

- Emerging technologies
- Multidisciplinary nature of work assignments
- Their expectations vs. reality (keep it positive!)
- The changing groundrules for career advancement
- Skills or *"what they don't teach you at Engineering School"*
- Total quality management principles
- Lessons learned in navigating the corporate world (again keep it positive!)

What are your Ideas?

UNIVERSITY INTERACTION WITH INDUSTRY

- In what ways does your College of Engineering involve industry in its:
  - overall program development?
  - faculty development?
  - research activities?
  - preparation of students for entering industry?
  - placement of graduates in industry?
  - continuous education of practicing engineers?
  - efforts to increase diversity in engineering?
- What does your College of Engineering consider its core research competencies to be?
- What approaches does your College of Engineering use for technology transfer?
- What is your College's vision as regards working with industry?
- Who is your College's key industry interface?

ATTRIBUTES OF AN ATTRACTIVE ENGINEERING GRADUATE

- A good grasp of engineering science fundamentals
  - Mathematics, including statistics
  - Physical and life sciences
  - Information technology
- A good understanding of the design process and manufacturing
- A basic understanding of the context in which engineering is practiced
  - Economics
  - History
  - Ethics and the environment
  - Customer and social needs
- Good communication skills
- Ability to think critically and creatively; independently and cooperatively
- Flexibility: ability and self-confidence to adapt to rapid/ major change
- Curiosity and a desire to learn - for life
- Profound understanding of the importance of teamwork

"ABET ENGINEERING CRITERIA 2000"

Engineering programs must demonstrate that their graduates have:

- an ability to apply knowledge of mathematics, science and engineering
- an ability to design and conduct experiments and analyze and interpret data
- an ability to design a system, component, or process to meet desired needs
- an ability to function on multi-disciplinary teams
- an ability to identify, formulate and solve engineering problems
- an understanding of professional and ethical responsibility
- an ability to effectively communicate
- the broad education necessary to understand the impact of engineering solutions in a global societal context
- a recognition of the need for and an ability to engage in life-long learning
- a knowledge of contemporary issues
- an ability to use the techniques, skills and modern engineering tools necessary for engineering practice

SOURCE: ABET Engineering Criteria 2000, Draft #4, Criterion 3: Program Outcomes and Assessment

Development of Plant Gene Vectors for Tissue-Specific Expression Using GFP as a Reporter Gene

Jacquelyn Jackson,* Marceline Egnin, Qihan Xue and C. S. Prakash
Tuskegee University NASA Center for Food Production, Processing, and Waste Management for Controlled Ecological Life Support Systems, Tuskegee, Al 36088

Abstract

Reporter genes are widely employed in plant molecular biology research to analyze gene expression and to identify promoters. Gus (*UidA*) is currently the most popular reporter gene but its detection requires a destructive assay. The use of jellyfish green fluorescent protein (GFP) gene from *Aequorea victoria* holds promise for non-invasive detection of *in vivo* gene expression. To study how various plant promoters are expressed in sweetpotato (*Ipomoea batatas*), we are transcriptionally fusing the intron-modified (mGFP) or synthetic (modified for codon-usage) GFP coding regions to these promoters: double cauliflower mosaic virus 35S (CaMV 35S) with AMV translational enhancer, *ubiquitin7*-intron-ubiquitin coding region (*ubi7*-intron-UQ) and sporaminA. A few of these vectors have been constructed and introduced into *E. coli* DH5α and *Agrobacterium tumefaciens* EHA105. Transient expression studies are underway using protoplast-electroporation and particle bombardment of leaf tissues.

Key Words GFP (Green Fluorescent Protein), GUS (β-glucuronidase), X-gal, IptG, SporaminA, pUC19, bluescript, cloning

Introduction

Reporter genes have been used as convenient markers to visualize gene expression and protein localization *in vivo* in a wide spectrum of prokaryotes and eukaryotes. Hence, in plant biotechnology, reporter genes have been the prerequisite for developing reliable transformation and regeneration techniques, in which gene expression and selection of transformants require the use of marker genes that respectively function as reporters of the temporal and spatial expression of the foreign gene, and permit the recovery of transformed cells, tissues, or organs. Each reporter has its own specific characters which determine its limitations and its usefulness. However, the detection of the commonly used reporters, such as the *Escherichia coli UidA* gene, commonly known as GUS (Jefferson, 1987; Jefferson *et al.*, 1987) requires either exogenous substrates/co-factors or antibodies. The

*Jacquelyn Jackson is a senior majoring in Plant and Soil Sciences at Tuskegee University. She plans to graduate in the spring of 1997 and later attend graduate school studying horticulture. She is currently working on developing plant gene vectors for studying promoters in sweetpotato. Ms. Jackson has spent two summers at Michigan State University cloning cold-regulated genes with Michael Thomashow and analyzing a putative glucosyltransferase in arabidopsis under the guidance of Kenneth Keegstra.

assays are toxic to tissues and cannot be used to follow gene expression *in planta*. Other commonly used reporters include: luciferase (LUC) (Gallie *et al.*, 1989; Millar *et al.*, 1992; Ow *et al.*, 1986), NptII (also used as a selectable marker in which transformed plant cells are resistant to antibiotics such as kanamycin, paromomycin, and geneticin) (GIBCO BRL, Gaithersburg, MD), β-galactosidase (LacZ) (Helmer *et al.*, 1984; Terri *et al.*, 1989), and chloramphenicol acetyltransferase (CAT) (Seed and Sheen, 1988; Sheen and Jang, unpublished).

Our interest was to develop plant gene vectors for tissue-specific expression (constitutive, storage root, and periderm) in sweetpotato using the green fluorescent protein (mGFP) as a reporter gene. Respectively, chimeric genes were constructed, in which the reporter gene, mGFP was either fused directly to the following promoters: double 35S-AMV translational enhancer and sporaminA, or introduced as a translational fusion to the ubiquitin coding region driven by the *ubi7* promoter with an intron (*ubi7*-intron-UQ) (Hattori *et al.*, 1985; Ohta *et al.*, 1991; Belknap *et al.*, 1992)

Materials and Methods

Plasmids. The plasmids used in this study include: pBIN 35S-mGFP4, a gift from Dr. Jim Haseloff; p213, containing the sporaminA promoter, was provided by Dr. K. Nakamura; pCGN1547, containing the *ubiquitin7* promoter, was provided by Dr. Belknap; pBI426, containing the double 35S CaMV promoter; pUC19, and bluescript.

Culture Conditions and Isolation of Plasmid DNA. Bacterial colonies containing the plasmids were grown in 30ml broth overnight at 37°C with constant shaking at 200rpm. Plasmid DNA was isolated according to the Terrific Broth Plasmid DNA Prep. protocol (Biotechniques, 1990 9, 676-679). The purified plasmid DNA was cut with the appropriate restriction enzyme and electrophoresied on an 0.8% agarose gel with TBE buffer. The GFP coding sequence, appropriate promoter fragments, and linearized plasmid fragments were eluted from the agarose gel based on standard methods (Sambrook et al., 1989).

Vector Construction

mGFP. pBIN 35S-mGFP4 was cut with HindIII/EcoRI, and the GFP cassette (approx. 2.2kb) was subcloned in pUC19. This new plasmid was restricted with PstI/XbaI or HindII/XbaI, to drop the 35S promoter (approx. 900bp). The remaining pUC19 promoterless mGFP was used in ligation with sporaminA. In addition, the pBIN 35S-mGFP4 plasmid was cut with HindIII/XbaI to generate a promoterless mGFP in pBIN, which was used in ligation with the *ubi7* promoter.

SporaminA. The p213 (bluescript) plasmid, carrying the sporaminA promoter, was cut with SalI/PstI. The restricted fragment was subcloned into pUC19. The resulting pUC19/sporaminA plasmid was restricted with PstI and XbaI to generate a fragment containing the sporaminA promoter with the start codon towards the XbaI restriction site. This PstI/XbaI fragment was ligated into the linearized final plasmid, pUC19/pr-mGFP obtained above. The resulting plasmid being, pUC19 containing the chimeric sporaminA promoter fused to mGFP with the nos-terminator (approx. 2.3kb).

Ubiquitin7. The *ubi7* promoter, from pCGN1547, was dropped with BamHI and ligated into bluescript, which was cut with the same enzyme. Because of the orientation problem with the *ubi7*-intron-UQ fragment, plasmid DNA was isolated from

twenty positively selected colonies and digested with HindIII and XbaI to drop again, the *ubi7*-intron-UQ fragment. The twenty eluted fragments were ligated into the pBIN promoterless mGFP. The resulting ligation was plated on an Lb km plate with X-gal and IptG. DNA was isolated from positive colonies and restricted with HindIII/EcoRI to check for the presence of the *ubi7*-intron-UQ/mGFP fragment. Plasmid DNA from confirmed colonies were mobilized into *Agrobacterium tumefaciens* EHA105, which will be used in transformation studies to properly determine the orientation of the promoter.

d35S with enhancer: pBI426 was cut with BamHI and HindIII, which drops the double 35S promoter with the AMV enhancer (approx. 650bp). This promoter fragment was ligated into pBIN/pr-mGFP (HindIII/BamHI cut) and transformed into *E. coli*. Plasmid DNA from positively selected colonies was mobilized into *Agrobacterium tumefaciens* EHA105 for transformation purposes.

Results and Discussion

To determine whether GFP could be used as a vital marker or reporter in sweetpotato tissue, we generated chimeric GFP constructs that were transcriptionally fused to a storage root specific promoter (sporaminA), a periderm-specific promoter (*ubi7*-intron-UQ), and a constitutive promoter (double 35S promoter with AMV translational enhancer). Three independent lines of clones were generated for this purpose. In the first line, we restricted our plasmids with the appropriate enzymes, resulting in pBIN promoterless mGFP, and pUC19 promoterless mGFP. In the second line, we were able to generate promoter inserts such as, the *ubi7*-intron-UQ, sporaminA, and double 35S promoters, by using appropriate enzymes (Fig.1). In the third line, we ligated the promoter fragments into the plasmid vectors (Fig. 2). For example, the promoterless mGFP in pBIN was ligated to the double 35S promoter and the sporaminA (SpA) promoter, to obtain two different constructs: pBINSpA-mGFP4-NosT and pBINd35SAMV-mGFP4-NosT (Fig. 2), which we mobilized into *Agrobacterium tumefaciens* EHA105. The *ubi7*-intron-UQ fragment will be cut and subcloned into bluescript. The fragment will then be dropped out of bluescript and fused to mGFP in the pBIN binary vector. DNA from positively selected clones, at this stage, will be mobilized into *Agrobacterium tumefaciens* EHA105.

These newly developed chimeric constructs (Fig. 2), will be used in transformation studies on sweetpotato, to determine whether GFP could be used to detect changes in gene expression within living tissues. Work done on plants using GFP as a reporter gene include: Niedz *et al.*, 1995- Sweet orange protoplasts; Baulcombe *et al.*, 1995- PVX virus; Sheen *et al.*, 1995- C4PPDK enhancher-Arabidopsis; and Heinlein *et al.*, 1995- Tobamovirus-Tobacco.

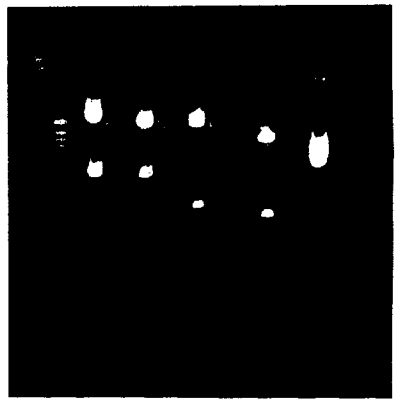
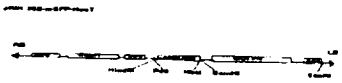
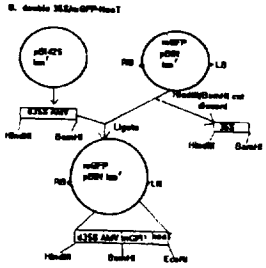
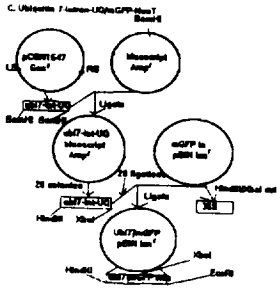


Figure 1. Gel electrophoresis of plasmid DNA fragments on an 0.8% agarose gel. Lane 1: weight IV marker, Lane 2: Ubi7 promoter fragment, Lane 3: mGFP (H3/RI cut), Lane 4: mGFP (BamHI/HindIII cut), Lane 5: pBI426 (d35S promoter; BamHI/HindIII cut).

Vector Construction Maps







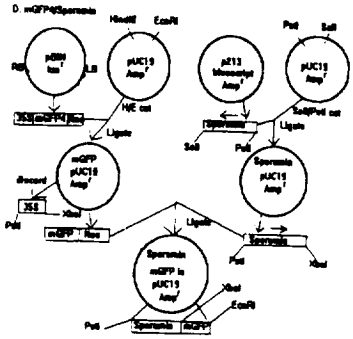


Fig. 2. Vector construction maps of (a.) the pBIN 35S-mGFP4-NosT construct, (b.) the double CaMV35S with AMV translational enhancer/mGFP construct, (c.) the *ubiquitin7*/mGFP construct, and (d.) the Sporamin/mGFP construct.

Acknowledgments

Special thanks to Dr. Marceline Egnin, C.S. Prakash, and Qihan Xue. This research is supported by NASA (Grant No. NAGW-2940).

References

Baulcombe, D. C., Chapman, S., and Cruz S. S. (1995) Jellyfish green fluorescent protein as a reporter for virus infections. *Plant J* 7(6), 1045-1053.

Galbraith, D. W., Lambert, G. M., Grebenok, R. J., and Sheen, J. (1995) Flow cytometric analysis of transgene expression in higher plants: green fluorescent protein. *Methods in Cell Biology* 50, 3-14.

Gallie, D. R., Lucas, W. J. and Walbot, V. (1989) Visualizing mRNA expression in plant protoplast: factors influencing efficient mRNA uptake and translation. *Plant Cell*, 1, 301-311.

Heinlein, M., Epel, B. L., Padgett, H. S., Beachy, R. N. (1995) Interaction of Tobamovirus Movement Proteins with the Plant Cytoskeleton. *Science* 270, 1983-1985.

Helmer, G., Casadaban, M., Bevan, M., Kayes, L. and Chilton, M. D. (1984) A new chimeric gene as a marker for plant transformation: the expression of *Escherichia coli* β-galactosidase

in sunflower and tobacco cells. *Bio/Technology*, 2, 520-527.

Niedz, R. P., Sussman, M. R., and Satterlee, J.S. (1995) Green fluorescent protein: an *in vivo* reporter of plant gene expression. *Plant Cell* 14, 403-406.

Sambrook, J., Fritsch, E. F., Maniatis, T., (1989) Molecular Cloning, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Seed, B. and Sheen, J. Y., (1988) A simple phase-extraction assay for chloramphenicol acetyltransferase activity. *Gene*, 67, 271-277.

Sheen, J., Hwang, S., Niwa, Y., Kobayashi, H., and Galbraith, D. W. (1995) Green fluorescent protein as a new vital marker in plant cells. *Plant J.* 8(5), 777-784.

STUDIES ON SOMATIC EMBRYOGENESIS IN SWEETPOTATO

J. Rasheed Bennett* and C. S. Prakash
Tuskegee University NASA Center for Food Production, Processing and Waste Management for Controlled Ecological Life Support Systems, Tuskegee, AL 36088

Abstract

The purpose of this study was to improve the somatic embryo (SE) system for plant production of sweetpotato *Ipomoea batatas* L.(Lam)l. Explants isolated from SE-derived sweetpotato plants were compared with control (non SE-derived) plants for their competency for SE production. Leaf explants were cultured on Murashige-Skoog (MS) medium with 2,4-dichlorophenoxy acetic acid (0.2 mg/L) and 6-benzylaminopurine (2.5 mg/L) for 2 weeks in darkness and transferred to MS medium with abscisic acid (2.5 mg/L). Explants isolated from those plants developed through somatic embryogenesis produced new somatic embryos rapidly and in higher frequency than those isolated from control plants. They also appeared to grow faster in tissue culture than the control plants. Current studies in the laboratory are examining whether plants derived from a cyclical embryogenesis system (five cycles) would have any further positive impact on the rapidity and frequency of somatic embryo development. More detailed studies using electron microscopy are expected to show the point of origin of the embryos and to allow determination of their quality throughout the cyclical process. This study may facilitate improved plant micropropagation, gene transfer and germplasm conservation in sweetpotato.

Abbreviations: ABA = abscisic acid, BAP = 6-benzylaminopurine, 2,4-D = 2,4-dichlorophenoxyacetic acid, GA3 = gibberellic acid, MS = medium of Murashige and Skoog

Keywords: synthetic seed, somatic embryo, somatic embryo derived, competency, regeneration, gene transfer.

Introduction

Regeneration of plants in vitro using somatic embryogenesis (SE) has some distinct features such as single cell origin, the consequent low frequency of chimeras and the production of a high number of regenerates (Haccing, 1978; Ammirato, 1983; Sato et al., 1993). Thus, the somatic embryo system is being increasingly employed to produce transgenic plants and to develop synthetic seeds using encapsulation strategies (Ritchie and Hodges, 1993; Red-enbaugh et al., 1986). Somatic embryos have been produced in sweetpotato (Tsay and Tseng, 1979; Lui and Cantliffe, 1884; Jarret et al., 1984). A reliable and reproducible SE system that is rapid and repetitive was developed in our laboratory using the -3 genotype (Zeng, 1994). Direct and callus-mediated embryogenesis has been used in sweetpotato micropropagation, germplasm conservation, and stable genetic transformation studies. Even though this system has been successful, further applications and developments are needed to improve the yield and regeneration rate.

Here we report observations and preliminary results of studies aimed at expanding and improving the current system of somatic embryogenesis in sweetpotato, a crop of considerable importance as a source of protein and calories in the developing world.

*John R. Bennett is a senior at Tuskegee University majoring in Plant and Soil Science. Originally from Buffalo, NY, he now resides in Auburn, AL with his wife. He has been conducting research in the Plant Cellular and Molecular Genetics Lab at Tuskegee University for the past two years. His overall research interests include developing and improving tissue culture methods used for the transformation of sweetpotato. His current research aim is to improve the yield and regeneration of sweetpotato somatic embryos using a novel cyclic regeneration technique. John R. Bennett says he came to Tuskegee University because his grandmother taught him about the accomplishments of George Washington Carver as a child. He is expected to graduate in August 1996. After graduation he plans to pursue a Ph.D. in Plant Molecular Biology.

Materials and Methods

Sweetpotato genotypes were obtained from the USDA Plant Introduction Center, Griffin, GA, as in vitro cultured shoot tips. Explants were obtained from plants grown in vitro on multiplication medium consisting of MS inorganic salts and vitamins (Murashige and Skoog, 1962), myo-inositol 100 mg/L, gibberellic acid (GA3) 5 mg/L, sucrose 30 g/L, and Phytagel 3 g/l. The basal medium consisted of MS inorganic salts, myo-inositol 100 mg/L, thiamine-HCl 1.7 mg/L, nicotinic acid 1.2 mg/L, pyridoxine-HCl 1.0 mg/L, sucrose 30 g/L and Phytagel 3 g/L.

The stage I medium consisted of basal medium supplemented with potassium chloride (2.235 g/L) (Chee et al., 1992), while the stage II medium consisted of basal medium with ammonium nitrate reduced to half-strength (Chee et al. 1992). The medium pH was adjusted to 5.8 prior by addition of Phytagel (Sigma), and autoclaved at 121 degrees C, 105 Kpa for 20 min.

The stage I and II media were further modified. The stage I medium had additions of 2,4-D at 2.5 mg/L and BAP at 0.25 mg/L, while the stage II medium contained additional ABA (+99% cis-trans; Sigma) at 2.5 mg/L. A stage III medium consisting of only basal medium was used for embryo germination and conversion of somatic embryos to plants.

Cut pieces of lamina (10 x 5 nm) were isolated from sweetpotato plants. Genotype PI 318846-3, grown in vitro, was used to provide explants. The explants were placed in petri dishes (100 x 15 mm) containing 30 mL of stage I medium. After a 2 week incubation in darkness at 25°C, explants were transferred to the stage II medium until mature SE formation occurred. Mature SE's were separated and geminated on the stage III medium. Upon completion of regeneration and conversion, these embryo-derived plants comprised the E1 generation and were isolated and used as explants in the production of new embryos; after completion of regeneration and conversion, these new embryos established the E2 generation. We refer to the method of obtaining SEs from SE derived plants as cyclic regeneration.

Discussion

To improve the efficiency of somatic embryogenesis in sweetpotato, we sought to test whether plants derived from somatic embryos provide competent explants that have high regenerable capacity. We established a population of somatic embryo-derived plants that had their origin through in vitro somatic embryogenesis. Cyclic regeneration was used to establish E1 to E4 generations. The generation number refers to the number of times that generation has completed the somatic embryogenesis system. Observations of preliminary results obtained through the development of SE derived plants seem to indicate that plants that have completed the SE system produced better quality embryos with fewer abnormalities, and more embryos than those from control plants. Ongoing studies are testing SE-derived plants which have undergone cyclical regeneration (over five SE cycles) compared to control plants to test whether the enhanced regeneration observed in SE plants is additive in nature over several SE cycles.

Acknowledgment

This research was supported by NASA through the Tuskegee University Center for Food Production, Processing and Waste Management for Controlled Ecological Life Support Systems (NASA Grant No. NAGW-2940).

References

Ammirato PV. (1983) In: Evans DA, Sharp WR, Ammirato PV and Yamada Y (eds) Handbook of Plant Cell Culture, Vol 1. Macmillan, New York, pp82-123

Chee RP, Leskovar DI and Cantliffe DJ. (1992) J Amer Soc Hort Sci 117:663-667

Haccing B. (1978) Phytomorphology 28:74-81

Jarret RL, Salazer S and Fernandez R. (1984) HortScience 19:397-398

Lui JR and Cantliffe DJ. (1984) Plant Cell Reports 3:112-115

Murashige T and Skoog F. (1962) Physiol Plant 15: 473-497

Redenbaugh K, Paasch BD, Nichol JW, Kossler ME, Viss PR and Walker KA. (1986) Bio/Technology 4: 797-801

Sato S, Newell C, Kolacz K, Tredo L, Finer J and Hinchee M. (1993) Plant Cell Reports 12:408-413

Tsay HS and Tseng MT. (1979) Bot Bull Acd Sin 20:117-122

290344

AUTOMATED LIQUID-LEVEL CONTROL OF A NUTRIENT RESERVOIR FOR A HYDROPONIC SYSTEM

[1]Boris Smith, Johnson A. Asumadu and Numan S. Dogan

Tuskegee University NASA Center for Food Production, Processing and Waste Management for Controlled Ecological Life Support Systems
Tuskegee, AL 36088
Tel: (334) 727 - 8992 Fax: (334) 724 - 4806 E-MAIL: bsmith25@acd.tusk.edu

ABSTRACT

A microprocessor-based system for control of the liquid level of a nutrient reservoir for a plant hydroponic growing system has been developed. The system uses an ultrasonic transducer to sense the liquid level or height. A National Instruments' Multifunction Analog and Digital Input/Output PC Kit includes NI-DAQ DOS/Windows driver software for an IBM 486 personal computer. A Labview Full Development system for Windows is the graphical programming system being used. The system allows liquid level control to within 0.1 cm for all levels tried between 8 and 36 cm in the hydroponic system application. The detailed alogorithms have been developed and a fully automated microprocessor-based nutrient replenishment system has been described for this hydroponic system.

Keywords: *Hydroponics, sweetpotatoes, data-acquisition, ultrasonic-sensor, virtual instrument, liquid-level.*

1.0 INTRODUCTION

Using manual support adjustments, sweetpotatoes and peanuts have been grown consistently in experimental hydroponic systems designed for future long-term NASA space missions involving humans at the Tuskegee University NASA Center for Controlled Ecological Life Support Systems (CELSS) [1]. The hydroponic systems used for growing sweetpotatoes employ the nutrient film technique (NFT) method. Such systems require accurate control of nutrient solution volume, pH, temperature, and solution flow rate to improve nutrient and O_2 delivery to the root system.

The use of digital techniques in instrumentation are well established [2] - [4]. The National Instruments' LabVIEW is a programming system for data acquisition and control, data analysis, and data presentation using a Graphical Unit Interface (GUI) platform. LabVIEW offers an innovative programming technique in which software objects can be assembled graphically to form a virtual representation of the actual control system. With LabVIEW, block diagrams are assembled to represent the functionalities expressed in the algorithms without worrying about the many syntactical details of conventional programming. The objects in the block diagrams are connected by wires to pass data from one block to the next depending on the functionalities expressed in the algorithms - from simple arithmetic functions to complex data acquisition and complex analysis routines. The LabVIEW platform development is low cost, provides enhanced accuracy and added functionality, and saves time.

An ultrasonic sensor is interfaced to a PC through the National Instruments' plug-in data acquisition card. The LabVIEW provides functions to acquire and output data from the plug-in card.

The purpose of this work is to present in detail algorithms for a virtual instrument (VI) that acquires data through a conventional plug-in data acquisition card to provide multifunction instrumentation for hydroponic crop production systems.

2.0 ALGORITHMS

Fig. 1 shows the functional diagram used to measure the height (volume) of the nutrient solution in a hydroponic system.

The data acquisition plug-in (DAQ) card is mounted in the host PC which receives the height information from an ultrasonic sensor mounted on top of the tank containing the nutrient solution. The hydroponic system consists of a growing channel (for sweetpotatoes or peanuts), a nutrient solution reservoir and a pump to supply the nutrient solution to the growing channel. The nutrient solution must be increasingly replenished as the plant 'moves through the growth cycle. A submersible pump is engaged to replenish the nutrient solution if the percent error change in the nutrient solution height is larger than a predetermined value.

[1]Boris Smith received the B.S. degree in electrical engineering from Tuskegee University, Alabama in December, 1995. He is currently working towards an M.S. degree in electrical engineering. His research interest include control systems and power systems.

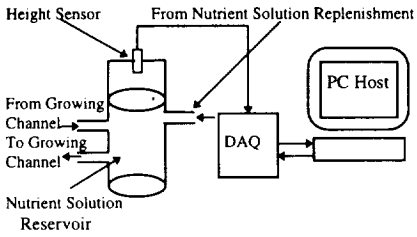


Fig. 1: Nutrient Solution Height Control

2.1 Algorithms

Fig. 2 shows the basic principle involved in the use of an ultrasonic pulse for a distance measurement. The ultrasonic sensor pulse is transmitted to the surface of the nutrient solution. The pulse, reflected from the surface of the nutrient solution, returns to the sensor. The distance from the sensor to the desired level of the nutrient solution is given by H.

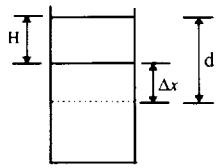


Fig. 2. Height Control System Using an Ultrasonic Sensor

The distance measured from the sensor to the surface of the nutrient solution at any time is represented by d. However, d can also be found from (1)

$$d = vt/2 \qquad (1)$$

where v is the velocity of sound in air, and t is the delay time of the ultrasonic pulse. In [3] the velocity of sound is given by

$$v = 331.5 + 0.61\ T\ (m/s) \qquad (2)$$

where T is the temperature (^{O}C) of air. The difference in height Δx is given by (3)

$$\Delta x = d - H = vt/2 - H \qquad (3)$$

Using (3), it can be shown that the percent error is given by

$$\text{Error} = \frac{vt - 2H}{2H} * 100\% \qquad (4)$$

The resolution of the phase detector and the directivity of the sensor seriously affect the accuracy of the height measurement because the orientation from the sensor to an object is determined by the sensitivity.

3.0 COMPLETE INSTRUMENTATION SYSTEM

3.1 Instrument Overview

In LabVIEW, graphical front panels, block diagrams and icons/connectors are used to represent the functions expressed in the algorithms. The front panels serve as the user interface for supplying inputs, placing controls and observing outputs of the actual instrument system. The front panel objects are chosen from a ***Controls*** menu - numeric displays, meters, gauges, thermometers, tanks, LED's, graphs, etc. The front panel will be a "virtual" representation of the actual instrument system and can therefore be used to control the actual control system.

A block diagram contains I/O operations, computational components, advanced acquisition and analysis routines, network and other components. The block diagrams are represented by icons that are interconnected by wires or lines to direct flow of data. The block diagram icons represent the source code, but these icons are free from text-based conventional programming details. A block diagram offers a complete programming structure such as loops, and conditional branching operations. The icons are selected from a ***Functions*** menu; the icons are referred to as nodes. A node operates only when data are available at all of its inputs.

LabVIEW I/O drivers communicate with data acquisition cards for acquiring the height of the nutrient solution from the appropriate sensor and responding depending on the sensor value.

3.2 Graphical Programming

Fig. 3 characterizes the front panel and the block diagram of the complete virtual instrumentation system used to control the volume of the nutrient solution in the hydroponic systems.

3.2.1 Top Level Front Panel

Fig. 3(a) represents the front panel of the virtual instrument. By using the objects from the ***Controls***

menu the front panel components are assembled as shown in Fig. 3(a). The front panel consists of LED maximum alarms, setpoints, meters to indicate

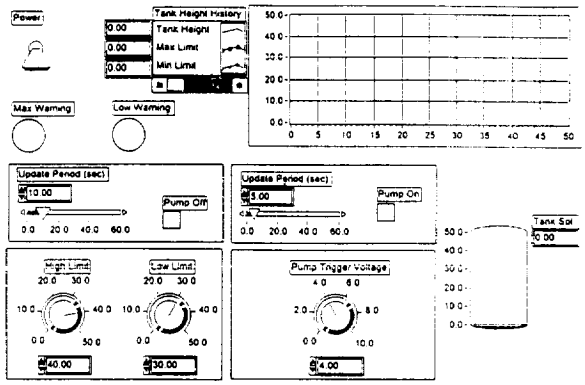


Fig. 3(a) Top Level Front Panel Diagram

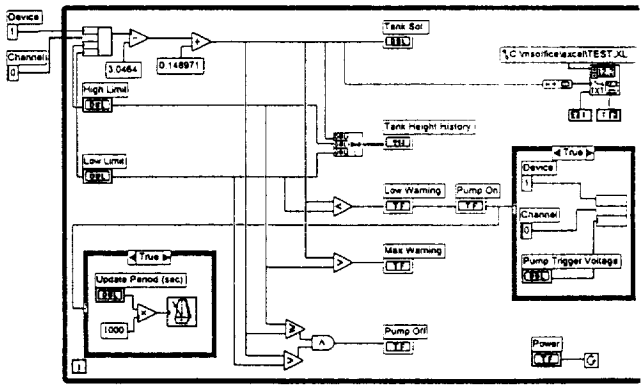


Fig. 3(b) Top Level Block Diagram

and minimum limits, knobs, slides, strip chart (showing the height of the nutrient solution in reservoir), pump indicators, and a replica of the nutrient reservoir containing the nutrient solution. The front panel is used to control the system.

3.2.2 Top Level Block Diagram

Fig. 3(b) represents the block diagram of the control system. The block diagram is the source code of the front panel. By using the objects from the ***Functions*** menu, the icons and interconnectors of the block diagram are assembled as shown in Fig. 3(b). Firstly, the objects of the block are selected, manipulated and positioned using the ***Positioning Tool***. The height of the nutrient solution is acquired by a PC-based plug-in DAQ board (AT-MIO-16E-2 by National Instruments). The data has a minimum differential input voltage of 100 mV and a maximum differential input of 10 V. Channel 0 is used to acquire the height of the nutrient solution. Device 0 is used to turn on a pump to replenish the nutrient solution. Second, the objects are wired by interconnecting the icons to represent the functionalities expressed in the algorithms.

4.0 RESULTS

Fig. 4(a) gives the plot for the height reading obtained from the virtual instrument (VI) versus the height of the actual nutrient solution in the reservoir. Fig. 4(b) gives the percent error for the height obtained with the LabVIEW with respect to the height of the nutrient solution reservoir. Fig. 4(b) shows that, for the measurements within the bandwidth of the ultrasonic sensor, the percent errors are less than $\pm$ 1.0% or less than 0.1 cm in height. It should be noted, however, that the ultrasonic sensor should not be located too close to (within 8.0 cm) or too far from (more than 36.0 cm) the nutrient solution level or it will provide erroneous readings, as displayed in Fig. 4(b).

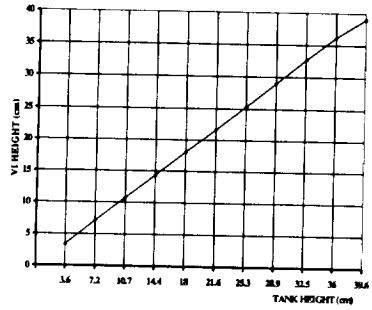


Fig. 4(a) VI Height Reading Versus Actual Reservoir Height

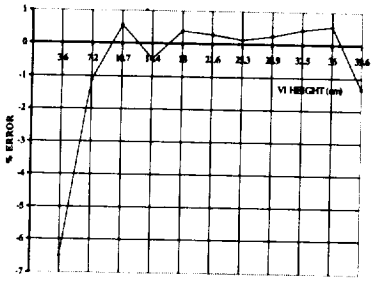


Fig. 4(b) Percent Error of VI Reading With Respect to Reservoir Height

5.0 CONCLUSIONS

This paper shows how efficient algorithms can be implemented using National Instruments' LabVIEW virtual instrument GUI platform. It describes a fully automated microprocessor-based nutrient replenishment system for hydroponic crop production systems. The system measures and maintains a constant height (volume control). An ultrasonic sensor is used to measure the height. The algorithms are very efficient and require minimum hardware to implement. It is a low-cost system intended to be used in hydroponics systems but could have many other applications as well.

ACKNOWLEDGMENT

This work is supported by NASA through the Tuskegee University NASA Center for Food Production, Processing and Waste Management for CELSS (Grant No. NAGW - 2940).

REFERENCES

[1] C. Bonsi, W. A. Hill, D. G. Mortley, P. A. Loretan, C. E. Morris and E. R. Calisle, "Growing Sweetpotatoes for Space Missions Using NFT," In: W. A. Hill, C. K. Bonsi and P. A. Loretan (editors) *Sweetpotato Technology for the 21st Century*, Tuskegee University, Tuskegee, Alabama, 1992, pp. 110 - 119.

[2] H. Nomako and T. Da-te, "Ultrasonic Position Measurement and Its Applications to Human Interface," *IEEE Transactions on Instruction and Measurement*, Vol. 44, No. 3, June 1995, pp. 771 - 774.

[3] T. Tsujimura, T. Yabuta and T. Morimitsu, "Shape-Reconstruction System for Three-Dimensional Objects Using an Ultrasonic Distance Sensor Mounted on a Manipulator," *Transactions of ASME*, Vol. 111, June 1989, pp. 180 - 186.

[4] C. L. Karr and E. J. Gentry, "Fuzzy Control of pH Using Genetic Algorithms," *IEEE Trans. Fuzzy Syst.*, Vol. 1, 1993, pp. 46 - 53.

FOOD SYSTEM DEVELOPMENT FOR CONTROLLED ECOLOGICAL LIFE SUPPORT SYSTEMS

CARLA D. WILSON*
TUSKEGEE UNIVERSITY NASA CENTER FOR CELSS
TUSKEGEE, ALABAMA

ABSTRACT
The development of a food system for the Human Rated Test Facility (HRTF) based on Controlled Ecological Life Support Systems (CELSS) crops is a critical component of Advanced Life Support (ALS) for extended duration space missions and exploration. The Biomass Production System (BPS) will provide for growth of higher plants for the purpose of supplying food to the CELSS crew and will also account for the planting, harvesting, processing and storage of the biomass prior to preparation for consumption. The human physiological and psychological requirements for food on extended duration missions should be measurable in terms of nutritional requirements and food acceptability. Current objectives at Tuskegee include the development of a vegetarian menu plan utilizing CELSS crop products, with a special emphasis on the sweetpotato; recipe modification and sensory evaluation of recipes, and the measurement of nutritional adequacy of the menu. These nutritional requirements will be unique to the CELSS crew. It is vital that this food system meet the nutrient requirements and aesthetic standards of the crew members to maintain good health to endure the challenges of CELSS. Nutrient requirements are based on physiological changes of crew members occurring in space. These changes have already been documented (1).

Keywords: vegetarian, food system, nutrient bioavailability, RDA, CELSS

A critical objective for an advanced life support system is the research and development of a functional space food system designed to provide crew members with appetizing, safe and nutritious food that can easily be prepared by crew members during long duration space missions. This food system consists of the food, its inedible biomass, and all of the equipment needed to preserve, prepare, and serve the food for consumption. It also requires precise limits for safety, acceptability, nutritional content, weight, volume, and storage stability.

This paper will provide an overview of the current work completed in response to outlined objectives researched at Tuskegee University, which concentrate on vegetarian menu development, recipe modification and evaluation, and nutritional adequacy.

*Carla D. Wilson is a graduate student at Tuskegee University pursuing a Master of Science degree in the area of Nutritional Science. She has recently earned two undergraduate degrees at Tuskegee in the areas of Hospitality Management and in General Dietetics. Her most recent internship experience took place at NASA Johnson Space Center, Houston, Texas where she was involved in food system development for CELSS. Her current research too involves food system development with a special emphasis on the nutritional adequacy and acceptability of a CELSS crop based vegetarian menu plan. Carla is expected to graduate in Spring of 1997.

MENU DEVELOPMENT

Crew members will be consuming mainly vegetarian meals due to the abundance of edible biomass and to the limited supply of stored meat products. Development of acceptable, appetizing menu plans that make available a variety of foods and food combinations is necessary to prevent food boredom and aversions.

Knowledge of proper nutrition and a great deal of imagination are needed to formulate nutritionally adequate vegetarian menus. These menus will be made up of food products from the CELSS crops; some of the major raw products to be used frequently will be wheat, rice, potato, sweet potato and soybean. Generally, these crops are staples for people, of various cultures.

So far, a three week vegetarian menu plan has been modified from a preexisting vegetarian menu plan designed for testing of zero gravity processing of a select number of crops. The modifications to the menus include a greater variety of foods with addition of more CELSS crop products. Menu items such as waffles, biscuits, noodles, cookies and bread are some of the products that will include sweetpotato as a major ingredient, which is intended to increase variety and the use of the sweetpotato and the amount of vitamin C.

Further research and development in this area should extend the menu plan beyond three weeks to avoid redundancy. Food frequencies and serving amounts will depend on the quantity of raw products available. A rough estimation of the raw products needed was made based on the three week vegetarian menu plan to support five crew members.

RECIPE DEVELOPMENT

Vegetarian recipes, or recipes that are mostly made up of CELSS crop products, will be necessary for meal preparation and nutritional analysis for CELSS. It is important that these recipes be easy to prepare and acceptable for consumption.

Several recipes were collected that correspond with the menu plan, and some were selected to be evaluated for acceptability. Some modification had to be done to the recipes; for example, soy milk was substituted for cow's milk and silken tofu was substituted for eggs. Recipes might be also modified to overcome specific limits in CELSS. Products evaluated were tofu lasagna, scrambled tofu, Southern Sweetpotato Pie, creamy carrot-potato soup, sweet potato rolls, sweetpotato waffle, hydroponic wheat waffle, rice waffle, sweetpotato biscuit, and hydroponic wheat biscuit. These products were evaluated on appearance, color, odor, flavor, texture and overall acceptability. They were rated on a Hedonic scale of nine to one, nine being liked extremely and one being disliked extremely. All but one of the products had average scores of 6.1 and higher meaning that they were acceptable. Scrambled tofu received average scores of 5.3 to 6.0. This was unacceptable.

NUTRITIONAL ADEQUACY

The recommended dietary allowances (RDA) have been used as a standard for planning the energy and nutrient content of current and past food systems for NASA space travel (2) and is also used for food system development for CELSS. The RDA for a CELSS crew member is modified to adapt to the environment and activity levels of future space flights (Table 1).

While a vegetarian diet can be rich in minerals and vitamins, there are some nutrients that are naturally present in animal products that would have to be supplemented(3). A prototype menu was developed and analyzed by the American Institute of Biological Science Biomass Processing Technical Panel at Kennedy Space Center who based this menu on eight selected crops: wheat, soybeans, rice, white potatoes, sweetpotatoes, greenbeans, leaf lettuce and sugar beets (1). The average nutrient content assessed from a ten day prototype CELSS menu (Table 2) showed high levels of most of the nutrients required daily for CELSS.

Table 1. A Comparison of the RDA for CELSS Crew members to the Recommended Dietary Allowance (RDA).

Nutrient and Unit	RDA for Adult Male	RDA for CELSS
Calories, kcal	2700	3500
Protein, g	56	140
Vitamin A, mcg RE	1000	1000
Vitamin D, mcg	5	5
Vitamin E, mg alpha-TE	10	10
Vitamin C, mg	60	75
Thiamin, mg	1.4	1.8
Riboflavin, mg	1.6	2.1
Niacin, mg NE	18	23.1
Vitamin B-6, mg	2.2	2.8
Folacin, mcg	400	400
Vitamin B-12, mcg	3.0	3.0
Calcium, mg	800	800
Phosphorus, mg	800	800
Magnesium, mg	350	350
Iron, mg	10	20--33.3
Zinc, mg	15	15
Iodine, mcg	150	150

Table 2. Nutritional Assessment of the Ten Day Prototype CELSS Menu

Nutrient and Units	Content of Average Prototype CELSS Menu
Calories, kcal	3515
Protein, g	135
Vitamin A, mcg RE	2101
Vitamin D, mcg	0
Vitamin E, mg alpha-TE	95.5
Vitamin C, mg	68.5
Thiamin, mg	3.64
Riboflavin, mg	2.05
Niacin, mg NE	31.6
Vitamin B-6, mg	3.9
Folacin, mcg	1028
Vitamin B-12, mcg	0
Calcium, mg	953
Phosphorus, mg	2845
Magnesium, mg	1208
Iron, mg	35.5
Zinc, mg	20.4
Iodine, mcg	0

However, it also showed slight deficits or absences of other essential nutrients such as protein, vitamin D, vitamin C, vitamin B-12, and iodine. These deficiencies can cause problems like bone demineralization, pernicious anemia, and goiter. However, appropriate supplementation of these nutrients to the crew member's diet should minimize the risk of these nutrient deficiencies (1,3).

Continued research on vegetarian menu development and inclusion of other recommended crops for CELSS is expected to alter the nutritional content of the menu plan, hopefully with desirable nutrient amounts.

RECOMMENDATIONS

Development of a food system intended for CELSS is a large project considering objectives already defined and others yet to be made. Further research and information on the selected crops to be grown in CELSS such as nutrient composition is required to develop a nutritional database unique to CELSS that will aid in accurately assessing the nutritional adequacy of the menu plan. Software is available, such as, Nutritionist III and Nutritionist IV to assess the vegetarian menu.

Research in the area of easy and efficient raw food processing is recommended for the production of foodstuffs that would be acceptable for consumption. For example, the whole wheat flour, used as an ingredient in most of the bread recipes, made the finished products dense and dry. This flour needs to be processed into flour with the consistency of all-purpose and bread flours.

Recipe and menu development are areas in need of expansion, improvement and sensory evaluation, especially when the recipes are to be modified to include more sweetpotato. Modifications are intended to maintain good nutrient amounts and increase the amounts that are slightly deficit while also improving acceptability and preparation methods. Equipment evaluations are also necessary for the identification of the most efficient and most easily maintained and operated

equipment to be used. And research of effective storage methods may be necessary for those raw food products or finished products that would be produced in bulk for practical reasons.

There are many more things to consider for a fully functional and sufficient food subsystem, especially when integrated with the other subsystems of the CELSS. Those issues will be addressed as research continues.

REFERENCES

1. AIBS-KSC Biomass Processing Technical Panel. Final Report: Biomass processing and human nutrition for controlled ecological life support systems (CELSS). NASA Biomedical Operations and Research Office, Kennedy Space Center, Florida. July 11, 1988.

2. Altman, P.L., Talbot, J.M. Nutrition and metabolism in spaceflight. Journal of Nutrition. 1987; 117:421-27.

3. Salisbury, F.B., Clark, M.A.Z. Suggestions for crops grown in controlled ecological life-support systems, based on attractive vegetarian diets. Advanced Space Research. 1996; 18:(4/5)33-39.

CHARACTERIZATION OF PROTEINS IN FILTRATE FROM BIODEGRADATION OF CROP RESIDUE

Wileatha Horton and A. A. Trotman
Tuskegee University Center for Food Production, Processing and Waste Management for CELSS
Tuskegee, AL

Abstract

Biodegradation of plant biomass is a feasible path for transformation of crop residue and recycling of nutrients for crop growth. The need to model the effects of factors associated with recycling of plant biomass resulting from hydroponic sweetpotato production has led to investigation of natural soil isolates with the capacity for starch hydrolysis. This study sought to use nondenaturing gel electrophoresis to characterize the proteins present in filtered effluent from bioreactors seeded with starch hydrolyzing bacterial culture used in the biodegradation of senesced sweetpotato biomass. The study determined the relative molecular weight of proteins in sampled effluent and the protein banding pattern was characterized. The protein profiles of effluent were similar for samples taken from independent runs under similar conditions of starch hydrolysis. The method can be used as a quality control tool for confirmation of starch hydrolysis of crop biomass. In addition, this method will allow monitoring for presence of contaminants within the system—protein profiles indicative of new enzymes in the bioreactors.

(Keywords: biodegradation; electrophoresis; protein analysis; sweetpotato; nondenaturing PAGE)

Introduction

Waste management has an integral role in Controlled Ecological Life Support Systems (CELSS). The basic elements of CELSS are the crew, biomass production, food and waste processing (Wieland, 1994). At the Tuskegee University NASA Center for CELSS (TUNACC), hydroponically grown sweetpotato [(*Ipomoea batatas*) L. Walp] is researched as a candidate crop for future long-term space missions involving humans. At TUNACC, sweetpotato biomass not used for food is being biodegraded and the filtered effluent recycled at a 20% rate (v/v) into the nutrient solution. Biodegradation transforms the inedible sweetpotato biomass into material that may be used as nutrients for plant growth (Eckart, 1994). A method for monitoring and evaluating biological waste management and plant growth systems must be included in CELSS to avoid deleterious impacts on functions that are essential to the bioregenerative nature of a CELSS. Microorganisms excrete complex proteins called enzymes that catalyze specific biochemical reactions. The enzymes associated with microbial degradation of plant biomass are exuded into the medium during hydrolysis. The presence of these biomolecules associated with biological transformation may be characterized in an effort to rapidly determine

---

Wileatha Horton, a senior majoring in Chemical Engineering, has been working with TUNACC for the past two years. In 1995 she participated in the Summer Intern Program at NASA Johnson Space Center. She was in the Advanced Life Support Division where she worked on physicochemical methods for waste management. Presently, Wileatha works in the TUNACC Waste Management Group. Wileatha graduates in May, 1996 and intends to study Environmental Engineering at the graduate level.

rate and to monitor the process of biological degradation. Also, characterization of the process will insure adherence to quality control criteria and allow rapid recognition of contamination in the system. Proteins associated with biodegradation of crop residue may be characterized by relative molecular weight using nondenaturing electrophoresis. Nondenaturing electrophoresis separates proteins based on their relative size and charge. The acrylamide gel pore size serves to sieve molecules of different sizes and allows the proteins that are more highly charged to exhibit greater mobility (Bollag and Edelstein, 1991). In this study nondenaturing electrophoresis was examined as a rapid, reliable and efficient method to characterize the filtered effluent from biological degradation of starch in sweetpotato biomass. The objective of this study was to determine if nondenaturing electrophoresis of filtered effluent from aerobic stirred tank bioreactors would allow a rapid and reproducible analysis of the quality of the effluent.

Experimental Materials and Procedure

General Experimental Plan. The filtered effluent was produced by sampling of aerobic stirred tank bioreactors. In a replicated experiment, sweetpotato biomass was hydrolyzed by WDSt 3A, a naturally isolated starch hydrolyzing bacterial culture isolated from the sweetpotato fields at the Tuskegee University research farm. The liquid fraction in the bioreactors was either a basal salts medium or modified half Hoagland solution. The starch hydrolysis in each bioreactor was conducted at 30°C for 14 days. Effluent used in this study was sampled at the time of inoculation of each bioreactor and at 2, 4, 7, 9, 11 and 14 days post-inoculation. All samples were stored at 4°C. Each sample was dispensed as 500 µL aliquots into microcentrifuge tubes and centrifuged at 12,000 rpm for 4 minutes. The supernatant was transferred to 1.5 mL centrifuge tubes and stored 4°C until protein analysis was completed. Protein concentration was measured in microtiter plates for all samples using the Bio-Rad (CA, USA) Protein Microassay protocol with bovine gamma globulin as a standard. The absorbance (600 nm) of effluent samples was measured with a microtiter plate reader and concentration determined using the measured values for the bovine gamma globulin standard. All samples were then used in nondenaturing gel electrophoresis.

Nondenaturing Gel Electrophoresis. Polyacrylamide gel electrophoresis (PAGE) of the effluent samples was conducted in a chromatography chamber at 4°C to prevent protein denaturation. For all PAGE analysis, a Bio-Rad Mini-PROTEAN II Cell and Ready Gels (Bio-Rad, CA) were used. Electrophoretic analyses of effluent samples were completed using:

Method 1. A 4-20% Tris-HCl Mini-PROTEAN II Gradient Ready Gel was used to determine the relative molecular weights of the proteins present in the effluent. Molecular weight determinations were done with incorporation of Bio-Rad SDS-PAGE Molecular Weight Broad Range Standards in the end wells of electrophoretic runs. Gels were run for 45 minutes under a constant voltage of 200 volts and 120 milliamps for a two gel system.

Method 2. A 5% and a 10% Tris, Borate, EDTA (TBE) Mini-PROTEAN II Continuous Ready Gel were used to determine the presence of specific starch hydrolyzing enzymes—alpha amylase or amyloglucosidase—in the filtered effluent samples. The end wells of each gel contained purified alpha amylase or amyloglucosidase (Sigma, MO) to allow a comparison between the relative migration of the bands in the effluent samples and in the commercially supplied enzymes. Gels were run for 45 minutes under a constant voltage of 100 volts and 33 milliamps for a two gel assembly.

Buffers used were as described in the instruction manuals accompanying the Ready Gel system. All gels were stored in a 40% methanol/10% acetic acid solution until stained.

Gel Staining and Visualization. Gels were stained using the Bio-Rad Silver Staining protocol for Ready Gels. Bands were visualized with a scanning densitometer and IBM compatible software that recorded band patterns and displayed images on the computer monitor.

Results and Discussion

Electrophoretic analysis with 4-20% Tris-HCl Mini-PROTEAN II Gradient Ready Gel showed that there was no difference, due to medium in which hydrolysis occurred, in the protein banding patterns obtained for all samples at the 2-day sampling. But effluent samples from the 3-liter bioreactors had an additional protein fraction (~66,000 daltons) not seen in samples obtained from the 1-liter bioreactors. The larger bioreactor probably allowed expression of additional enzymes due to the higher substrate loading rate. Effluent samples obtained for the 4-day check showed no difference in protein expression regardless of medium or bioreactor size. There were also four bands, the maximum number, visualized for the bioreactors and the pattern was unaltered when the 7-day samples were electrophoresed. These four bands represented proteins of molecular weights approximating 97,400, 66,200, 45,000 and 14,400 daltons, respectively. These findings indicate that by 96 hours the biological degradation occurring attained steady state and may be the sampling interval that should be used in quality control monitoring of effluent to be used as a supplemental crop nutrient solution.

Electrophoretic analysis with 5% Tris, Borate, EDTA (TBE) Mini-PROTEAN II Continuous Ready Gel resulted in only bands being visualized for alpha amylase. This may be the result of the low acrylamide percentage in the gel and the duration of the electrophoretic run allowing the protein fractions in the effluent samples to move with such rapidity so that they were off the gel. However in electrophoretic analysis with 10% Tris, Borate, EDTA (TBE) Mini-PROTEAN II Contiuous Ready Gel bands were visualized only for the 14-day samples. The bands corresponded to the amylase band profile. Also, the medium in which biodegradation occurred impacted the protein profile for effluent samples. This was not seen with the 4-5% Tris-HCl gradient gels.

Electrophoretic analyses of bioreactor effluent using nondenaturing, gradient polyalcrylamide in a Tris-HCl buffer system can be a rapid and reliable method for monitoring the biological degradation of crop residue when the effluent of a bioreactor is incorporated into the nutrient solution. In the 4 - 20% Tris-HCl Mini-PROTEAN II Gradient Ready Gel system bands were obtained for all samples. The medium in which starch hydrolysis was conducted did not influence the protein pattern when the Tris-HCl buffer system was used. It was also evident that the experimental procedure needs to be refined with respect to sampling interval. Further study will examine a 96-h sampling interval for starch hydrolysis so that a recommendation can be included in the TUNACC biological waste management protocol for CELSS.

Acknowledgments

The research was funded under a NASA grant (No. NAGW-2940). The authors thank Philip Loretan and Robyn Drew for assistance in manuscript preparation.

References

Bollag, D.M. and S. J. Edelstein. 1991. Protein Methods. Wiley-Liss Publishers, New York.

Eckart, P. 1994. Life Support and Biospherics Fundamental Technologies Applications. Herbert Utz Publishers, Munich, Germany.

Wieland, P. O. 1994. Designing for Human Presence in Space: An Introduction to Environmental Control and Life Support Systems. NASA Reference Publication 1324. Marshall Space Flight Center, Alabama.

STUDIES ON SOMATIC EMBRYOGENESIS IN SWEETPOTATO

J. Rasheed Bennett and C. S. Prakash, Tuskegee University NASA Center for Food Production, Processing and Waste Management for Controlled Ecological Life Support Systems, Tuskegee, AL 36088

P3

Abstract

The purpose of this study was to improve the somatic embryo (SE) system for plant production of sweetpotato [*Ipomoea batatas* L.(Lam)]. Explants isolated from SE-derived sweetpotato plants were compared with control (non SE-derived) plants for their competency for SE production. Leaf explants were cultured on Murashige-Skoog (MS) medium with 2,4-dichlorophenoxy acetic acid (0.2 mg/L) and 6-benzylaminopurine (2.5 mg/L) for 2 weeks in darkness and transferred to MS medium with abscisic acid (2.5 mg/L). Explants isolated from those plants developed through somatic embryogenesis produced new somatic embryos rapidly and in higher frequency than those isolated from control plants. They also appeared to grow faster in tissue culture than the control plants. Current studies in the laboratory are examining whether plants derived from a cyclical embryogenesis system (five cycles) would have any further positive impact on the rapidity and frequency of somatic embryo development. More detailed studies using electron microscopy are expected to show the point of origin of the embryos and to allow determination of their quality throughout the cyclical process. This study may facilitate improved plant micropropagation, gene transfer and germplasm conservation in sweetpotato.

Biography

John R. Bennett is a senior at Tuskegge University majoring in Pant and Soil Science. He is originally from Buffalo, NY and now resides in Auburn, AL with his wife. He has actively been conducting research in the Plant Cellular and Molecular Genetics Lab at Tuskegee University for the past two years. His overall research interests include developing and improving tissue culture methods used for the transformation of sweetpotato. His current research aim is to improve the yield and regeneration of sweetpotato somatic embryos using a novel cyclic regeneration technique. John R. Bennett says he came to Tuskegee University because his grandmother taught him about the accomplishments George Washington Carver as a child. He is expected to graduate in August 1996. After graduation he plans to pursue a Ph.D. in Plant Molecular Biology.

Abbreviations: ABA = abscisic acid, BAP = 6-benzylaminopurine, 2,4-D = 2,4-dichlorophenoxyacetic acid, GA3 = gibberellic acid, MS = medium of Murashige and Skoog (1962).

Keywords: synthetic seed, somatic embryo, somatic embryo derived, competency, regeneration, gene transfer.

Introduction

Regeneration of plants in vitro using somatic embryogenesis (SE) has some distinct features such as single-cell origin, the consequent low frequency of chimeras and the production of a high number of regenerates (Haccing 1978; Ammirato 1983; Sato et al. 1993). Thus, the somatic embryo system is being increasingly employed to produce

transgenic plants and to develop synthetic seeds using encapsulation strategies (Ritchie and Hodges, 1993; Redenbaugh et al., 1986). Somatic embryos have been produced in sweetpotato (Tsay and Tseng, 1979; Lui and Cantliffe, 1884; Jarret et al., 1984). A reliable and reproducible SE system that is rapid and repetitive was developed in our laboratory using the -3 genotype (Zeng, 1994). Direct and callus-mediated embryogenesis has been used in sweetpotato micrpropagation germplasm conservation, and stable genetic transformation studies. Even though this system has been successful, further applications and developments are needed to improve the yield and regeneration rate.

Here, we report observations and preliminary results of studies aimed at expanding and improving the current system of somatic embryogenesis in sweetpotato, a crop of considerable importance as a source of protein and calories in the developing world.

Materials and Methods

Sweetpotato genotypes were obtained from the USDA Plant Introduction Center, Griffin, Georgia, as in vitro cultured shoot tips. Explants were obtained from plants grown in vitro on multiplication medium consisting of MS inorganic salts and vitamins (Murashige and Skoog,1962), myo-inositol 100 mg/L, gibberellic acid (GA3) 5 mg/L, sucrose 30 g/L, and Phytagel 3 g/l. The basal medium consisted of MS inorganic salts, myo-inositol 100 mg/L, thiamine-HCl 1.7 mg/L, nicotinic acid 1.2 mg/L, pyridoxine-HCl 1.0 mg/L, sucrose 30 g/L and Phytagel 3 g/L. The stage I medium consisted of basal medium supplemented with potassium chloride (2.235 g/L) (Chee et al., 1992), while the stage II medium consisted of basal medium with ammonium nitrate reduced to half-strength (Chee et al. 1992). The medium pH was adjusted to 5.8 prior to addition of Phytagel (Sigma), and autoclaved at 121 degrees C, 105 Kpa for 20 min. The stage I and II media were further modified. The stage I medium had additions of 2,4-D at 2.5 mg/L and BAP at 0.25 mg/L, while the stage II medium contained additional ABA (±99% cis-trans; Sigma) at 2.5 mg/L. A stage III medium consisting of only basal medium was used for embryo germination and conversion of somatic embryos to plants.

Cut pieces of lamina (10 x 5 nm) were isolated from sweetpotato plants. Genotype PI 318846-3, grown in vitro, was used to provide explants. The explants were placed in petri dishes (100 x 15 mm) containing 30 mL of stage I medium. After a 2 week incubation in darkness at 25 degrees C, explants were transferred to the stage II medium until mature SE formation occurred. Mature SE's were separated and geminated on the stage III medium. Upon completion of regeneration and conversion, these embryo-derived plants comprised the E1 generation and were isolated and used as explants in the production of new embryos; after completion of regeneration and conversion, these new embryos established the E2 generation. We refer to the method of obtaining SE's from SE derived plants as cyclic regeneration.

Discussion

To improve the efficiency of somatic embryogenesis in sweetpotato, we sought to test whether plants derived from somatic embryos provide competent explants that have high regenerable capacity. We established a population of somatic embryo-derived plants that had their origin through in vitro somatic embryogenesis. Cyclic regeneration was used to establish E1 to E4 generations. The generation number refers to the number of times that generation has completed the somatic embryogenesis system. Observations of preliminary results obtained through the development of SE derived plants seem to indicate that plants that have completed the SE system produced better quality embryos with fewer

abnormalities, and more embryos than those from control plants. Ongoing studies are testing SE-derived plants which have undergone cyclicle regeneration (over five SE cycles) compared to control plants to test whether the enhanced regeneration observed in SE plants is additive in nature over several SE cycles.

Acknowledgment

This research was supported by NASA through the Tuskegee University Center for Food Production, Processing and Waste Management for Controlled Ecological Life Support Systems (NASA Grant No. NAGW-2940).

References

Ammirato PV. (1983) In: Evans DA, Sharp WR, Ammirato PV and Yamada Y (eds) Handbook of Plant Cell Culture, vol 1. Macmillan, New York, pp82-123

Chee RP, Leskovar DI and Cantliffe DJ. (1992) J Amer Soc Hort Sci 117:663-667

Haccing B. (1978) Phytomorphology 28:74-81

Jarret RL, Salazer S and Fernandez R. (1984) HortScience 19:397-398

Lui JR and Cantliffe DJ. (1984) Plant Cell Reoorts 3:112-115

Murashige T and Skoog F. (1962) Physiol Plant 15: 473-497

Redenbaugh K, Paasch BD, Nichol JW, Kossler ME, Viss PR and Walker KA. (1986) Bio/Technology 4: 797-801

Sato S, Newell C, Kolacz K, Tredo L, Finer J and Hinchee M. (1993) Plant Cell Reports 12:408-413

Tsay HS and Tseng MT. (1979) Bot Bull Acd Sin 20: 117-122

LIGHT-INDUCED ALTERATIONS IN BASAL GANGLIA KYNURENIC ACID LEVELS.

Angela E. Sroufe, J.A. Whittaker, J.W. Patrickson and M.C. Orr. Department of Anatomy & Neuroscience Institute, Morehouse School of Medicine, 720 Westview Drive, Atlanta, GA 30310.

Abstract: The metabolic synthesis, release and breakdown of several known CNS neurotransmitters have been shown to follow a circadian pattern entrained to the environmental light/dark cycle. The levels of excitatory amino acid (EAA) transmitters such as glutamate, have been shown to vary with environmental lighting conditions. Kynurenic Acid (KA), an endogenous tryptophan metabolite and glutamate receptor antagonist, has been reported to have neuroprotective effects against EAA-induced excitotoxic cell damage. Changes in KA's activity within the mammalian basal ganglia has been proposed as being contibutory to neurotoxicity in Huntington's Disease. It is not known whether CNS KA levels follow a circadian pattern or exhibit light-induced fluctuations. However, because the symptoms of certain degenerative motor disorders seem to fluctuate with daily 24 hour rhythm, we initiated studies to determine if basal ganglia KA levels were influenced by the daily light/dark cycle and could influence motor function. Therefore, in this study, HPLC-EC was utilized to determine basal ganglia KA levels in tissue extracts from adult male Long-Evans rats (200-250g) entrained to 24 and 48 hours constant light and dark conditions, respectively. Samples were taken one hour before the onset of the subjective day and one hour prior to the onset of the subjective night in order to detect possible phase differences in KA levels and to allow for accumulation of factors expressed in association with the light or dark phase. Data analysis revealed that KA levels in the basal ganglia vary with environmental lighting conditions; being elevated generally during the dark. Circadian phase differences in KA levels were also evident during the subjective night and subjective day, respectively. Results from these studies are discussed with respect to potential cyclic changes in neuronal susceptibility to excitotoxic damage during the daily 24 hour cycle and its possible relevance to future therapeutic approaches in treating neurodegenerative disorders.

Key Words: kynurenic acid, excitatory amino acids, basal ganglia, neurotoxicity, circadian-rhythm

Introduction

Numerous physiological functions, including alterations in the levels of several known neurotransmitters within the central nervous system, are modulated by the circadian oscillator localized in the hypothalamic suprachiasmatic nucleus (SCN).[1-3] The output of the SCN oscillator is entrained to the environmental 24-hour light-dark cycle and is tightly coupled to the release of melatonin, a tryptophan metabolite released from the pineal gland.[4-5] In addition to melatonin, tryptophan metabolism generates 5-HT as well as components of the kynurenine pathway. Both melatonin and CNS levels of the neurotransmitter serotonin (5-HT), follow a circadian pattern. These have been implicated in modulating SCN output functions as well as the photic entrainment of sleep/wake states and locomotor activities.[4-8] By-products of the kynurenine pathway, on the other hand, are recognized as endogenous ligands for excitatory amino acid (EAA) receptors. Kynurenic acid (KA), the only known endogenous antagonist, is a product of this pathway and has been shown to interact with glutamate binding sites on the N-methyl-D-aspartate (NMDA), a-amino-3-hydroxy-5-methyl-isoxazolepropionate (AMPA) and kainate receptors.[9-12] It is uncertain whether the generation of endogenous, metabolically-derived EAA receptor ligands such as KA, is influenced by environmental lighting or whether its levels within the CNS follow a circadian pattern.

Within the mammalian basal ganglia, an area of the brain known to be involved in behavioral and locomotor activities, Beal et al. have shown that the levels of kynurenic acid (KA) are reduced in many neurological diseases involving neurodegeneration and motor system dysfunction.[13] This suggests that an increase in susceptibility to EAA-induced excitotoxicity may be due to a loss of KA's neuroprotective effects. In fact, excessive activation of EAA receptors has been implicated in the pathologic sequelae of stroke and epilepsy, as well as in neurodegenerative disorders such as Huntington's and Parkinson's disease.[14-20] Although the timing of behavioral tasks and basal ganglia-mediated motor expressions follow a circadian rhythm,[21] it is not known whether KA levels in the basal ganglia exhibit a circadian pattern. Therefore, we hypothesized that circadian or light-induced alterations in KA could potentially confer variable susceptibility to excitotoxicity during the typical 24 hour light-dark cycle, possibly mediated by the presence of glutamate and its analogues. Based on this, we examined the circadian and light-induced alterations in KA levels in the rat striatum and substantia nigra.

Materials and Methods

Animals

Adult male Long-Evans Hooded rats (200-250 g) were maintained on a 12 hour light/dark (12L:12D) cycle (lights on at 0600 hours; lights off at 1800 hours) in laboratory quarters with free access to food and water. After at least 3 days of 12L:12D synchronization, rats were exposed to constant dark or constant light conditions for 24 and 48 hours, respectively. Control animals remained in the 12L:12D cycles for the same time duration as the experimental groups. Following entrainment, animals were decapitated either one hour prior to the onset of day light (subjective night) or one hour prior to the onset of night (subjective day). These times to sacrifice the animals were chosen to allow for the possible accumulation of the metabolic changes over the entire light or dark period, respectively. The brains were rapidly removed, submerged in ice-cold 0.9% saline and dissected on ice. Prior to KA extraction and analysis procedures, wet tissue weights were determined. Tissue samples were frozen at -80°C until ready for use.

Tissue Extraction and Kynurenic Acid Analysis

Frozen sample aliquots from rat striatum and substantia nigra were sonicated on ice in 0.1 M $HClO_4$ (10µL/mg wet tissue weight) for three 10-second intervals. The homogenate was centrifuged at 12,000 rpm for 30 minutes at 4°C. 20 µL of the supernatant was injected into a high-performance liquid chromatography (HPLC) system with 8 electrochemical sensors (CoulArray Electrode Array System model 5500, ESA, Inc., Bedford, MA). KA levels were measured using an isocratic method. The mobile phase consisted of 50 mM Na_2HPO_4 in 5% methanol at pH 6.26 (adjusted with H_3PO_4). A 15 cm x 4.6 mm 5µm C_{18} reverse-phase column (NBS, ESA, Inc.) with a flow rate of 1 mL/min was used. The electrochemical sensors were set at 150, 200, 300, 600, 750, 1000, 1040 and 1090 mV. KA was detected at 1040mV. The average retention time was 14.5 minutes.

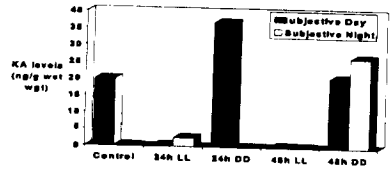


Figure 1. KA levels in striatum of rats exposed to varying LL and DD photoperiods.

Results and Discussion

Findings from these studies indicate that there are major changes in KA levels occur within the basal ganglia of the rat in response to alterations in the environmental photoperiod.

Striatum

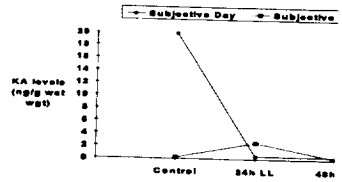


Figure 2. KA levels in striatum follow exposure to contstant light.

The overall profile for KA levels in the striatum is shown in Figure 1. Each point on the graph represents data averaged from three animals. This graph includes data points taken during the subjective day and subjective night. In control animals, those kept in 12L:12D cycles, KA levels were approximately 20 times higher during the subjective day than subjective night. The same data is represented in Figure 2, separated according to photoperiods. Here, the constant light data has been separated from the constant dark data to better illustrate the changes that are taking place relative to controls. Note that exposure to constant light conditions suppressed KA levels during both the subjective day and subjective night. During the subjective night, KA levels did not vary significantly. However, major changes occurred during the subjective day in which KA levels decreased significantly within the first 24 hours. These light-induced levels remained low as long as continuous light conditions were maintained. In contrast, an increase in striatal KA levels was observed within 24 hours of constant dark conditions (Figure 3). This increase was transient however, in that there is a decrease to initial control values within 48 hours.

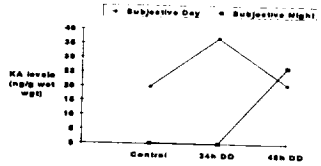


Figure 3. KA levels in striatum following exposure to constant dark.

Subjective night-time levels did not vary in the first 24 hours of constant dark, however, these levels rose rapidly to reflect daytime levels within 48 hours. Clearly, at the end of 48 hours, basal ganglia KA levels are low during constant light and are elevated to approximately day-time control levels in constant dark. This suggest a loss of KA's neuroprotective antagonistic activity at EAA receptors and a potential increase in EAA excitotoxic susceptibility during light photoperiods.

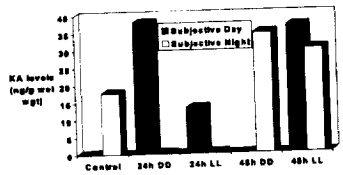


Figure 4. KA levels in substantia nigra of rats exposed to varying LL and DD photoperiods.

Substantia Nigra

The profile for KA levels in the substantia nigra is shown in Figure 4. Note that control values are higher during the subjective night than during the subjective day in contrast to 12L:12D values

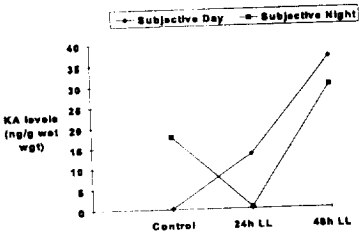


Figure 5. KA levels in substantia nigra following exposure to constant light.

observed in the striatum. Like the striatum, major changes in KA levels occured within 48 hours of photoperiod alteration. Analysis of constant light data (Figure 5) revealed that KA levels were increased at the end of 48 hours of constant light in both the subjective day and subjective night. There was an apparent linear relationship between subjective day KA levels and light exposure. Subjective night levels showed a decrease between 12 and 24 hours, but after 24 hours showed a steep rise. In other words, there was an inverse relationship between subjective day and night-time KA levels within the first 24 hours. Within 48 hours however, these became synchronized to rise in a linear fashion above initial night-time control levels. In constant dark (Figure 6), night-time levels followed the same pattern as observed in constant light conditions. Subjective day values were

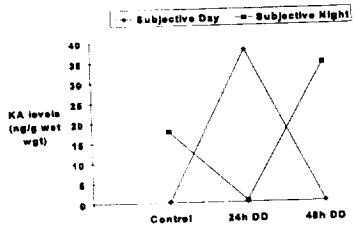


Figure 6. KA levels in substantia nigra following exposure to constant dark.

completely out of phase with the night-time photoperiod becoming elevated at 24 hours and returning to control values within 48 hours.

Conclusions

Together, these data demonstrate that there are obvious changes in KA levels occuring in the basal ganglia in response to variations in environmental lighting conditions. These changes could result from 24 hour cyclic changes in tryptophan metabolism, reflected as variations in endogenous KA release and cellular transport into the basal ganglia neuropil. The local neurochemical environment within the basal ganglia structures may facilitate differential cellular responses to these receptor alterations. In addition, susceptibility to EAA-associated excitotoxicity may increase in the striatum during extended light exposure due to local reduction in KA levels. Potential light-induced reductions in tryptophan breakdown could account for possible loss of KA's endogenous antagonistic actions at EAA receptors, thereby reducing its neuroprotective effects. The converse appears true for the substantia nigra. These observations may be of importance in a wide number of clinical areas including basal ganglia-mediated mechanisms underlying locomotor functions and behavior. Findings from this study could have future impacts on the therapeutic management of neurological diseases whose symptoms show 24 hour

variations in their responses to drug therapy.

Address Corespondence to Dr. J.A. Whittaker, Department of Anatomy, Morehouse School of Medicine, 720 Westview Drive, Atlanta, GA, 30311

Acknowledgement: This work was supported by NIH Grants GM08248, RR03034, NS34194 and NASA Grant NCCW-0083.

References

1. Ding J.M., Chen D., Weber E.T., Faiman L.E., Rea M.A. and Gillette M.U. *Science* 266:1713-1717, 1994.
2. Smith A.D., Olson R.J. and Justice Jr. J.B. *J. Neuroscience Methods* 44:33-41, 1992.
3. Lundberg G.P., Liebson P.R. and Calvin J.E. *J Crit Illness* 10:693-707, 1995.
4. Margraf R.R. and Lynch G.R. *Am. J. Physiol.* 264:R615-R621, 1993.
5. Hastings M.H., Mead S.M., Vindlacheruvu R.R., Ebling F.J.P., Maywood E.S., and Grosse J. *Brain Res.* 591:20-26, 1992.
6. Koumenis C. And Eskin A. *Chronobiol. Int.* 9:201-221, 1993.
7. Shibata S., Tsuneyoshi A., Hamada T., Tominaga K. And Watanabe S. *Brain Res.*582:353-356, 1992.
8. Trachsel L., Edgar D.M., Seidel W.F., Heller H.C. and Dement W.C. *Brain Res.* 589:253-261, 1992.
9. Danysz W., Fadda E., Wroblewski J.T. and Costa E. *Neurosci. Lett.* 96:340-344, 1989.
10. Stone T.W. and Connick J.H. *Neuroscience* 15:597-617, 1985.
11. Birch P., Grossman C.J. and Hayes A.G. *Eur. J. Pharmacol.* 151:313-315, 1988.
12. Birch P., Grossman C.J. and Hayes A.G. *Eur J. Pharmacol.* 154:85-75, 1988.
13. Beal M.F., Matson W.R., Swartz K.J., Gamache P.H. and Bird E.D. *J. Neurochem.* 55:1327-1339, 1990.
14. Schwarcz R., Foster A.C., French E.D., Whetsell W.O. and Kohler C. *Life Sci.* 35:19-32, 1984.
15. Foster A.C., Vezzani A., French E.D. and Schwarcz R. *Neurosci. Lett.* 48:273-278, 1984.
16. Simon R.P., Swan J.H., Griffith T. and Meldrum B.S. *Science*226:850-852, 1984.
17. Wieloch T. *Science* 230:681-683, 1985.
18. Coyle J.T. and Schwarcz R. *Nature(London)* 263:244-246, 1976.
19. McGeer E.G. and McGeer P.L. *Nature (London)*263:517-519, 1976.
20. Beal M.F., Kowall N.W., Ellison D.W., Marzurek M.F., Swartz K.J. and Martin J.B. *Nature(London)* 321:168-171.
21. Morell V. *Science* 271:905-906, 1995.

ENDOTHELIAL CELL MORPHOLOGY AND MIGRATION ARE ALTERED BY CHANGES IN GRAVITATIONAL FIELDS.

Caroline Melhado, Gary Sanford and Sandra Harris-Hooker

Departments of Biochemistry and Medicine
Morehouse School of Medicine
Atlanta, GA 30310.

ABSTRACT: Endothelial cell migration is important to vascular wall regeneration following injury or stress. However, the mechanism(s) governing this response is not well understood. The microgravity environment of space may complicate the response of these cells to injury. To date, there are no reports in this area. We examined how bovine aortic (BAEC) and pulmonary (BPEC) endothelial cells respond to denudation injury under hypergravity (HGrav) and simulated microgravity (MGrav), using image analysis. In 10% FBS, the migration of confluent BAEC and BPEC into the denuded area was not affected by HGrav or MGrav. However, in low FBS (0.5%), signficantly retarded migration under MGrav, and increased migration under HGrav was found. MGrav also decreased the migration of postconfluent BPEC while HGrav showed no difference. Both MGrav and HGrav strongly decreased the migration of postconfluent BAEC. Also, both cell lines showed significant morphological changes by scanning electron microscopy. These studies indicate that endothelial cell function is affected by changes in gravity.

INTRODUCTION: Many of the physiological changes seen in humans and animals during space flight may originate from the dysfunction of basic biological mechanisms caused by microgravity. Basic cellular functions such as electrolyte concentration, cell growth rate, glucose utilization, bone formation, response to growth stimulation, and exocytosis are modified by microgravity (1-4). Few studies have addressed the effect of microgravity on vascular cell function and metabolism. However, intact animal studies on spaceflights have shown lipid accumulations within the lumen of the vasculature and degeneration of the vascular wall (5). Capillary alterations with extensive endothelial invaginations were also seen (6). Hemodynamic studies have shown that there is a redistribution of blood from the lower extremities to the upper part of the body, this will alter vascular permeability, resulting in leakage into surrounding tissues (7). These studies indicate that changes in gravitational fields will alter the morphology and function of the vascular system. However, few studies have been directed at assessing the effect of altered gravitational fields on vascular cell function and metabolism. It is possible that cell:cell and cell:matrix interactions, as well as, migration and adherence may depend greatly on gravity. We examined the following questions: 1) Does the morphological phenotype of vascular cells change under different gravitational simulations?, and 2) Are there also alterations in cell motility?

EXPERIMENTAL METHODS: Bovine aorta (BAEC) and pulmonary artery (BPEC) endothelial cells were obtained from the NIGMS/Coriell Cell Repository. Cells were adapted to and maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum and 1X antibiotics. We assessed the influence of horizontal clinostat rotation at 30 rpm (MGrav) and centrifugation at 5-6G (HGrav) on BAEC and BPEC morphology and migration.

Morphological Studies: Confluent and 4-day postconfluent (PC) cultures were subjected to denuding an area of the culture, and then centrifuged at 6G (HGrav) or clinostat rotated at 30 rpm (MGrav) for 24 hr. Photomicrographs were taken 4, 8, 12 and 24 hr after injury to assess the morphology of the cells as they moved into the denuded area. Cultures were also rinsed with Hanks Buffered Salt Solution and fixed with 10% buffered formalin/2.5% glutaraldehyde for examination by scanning electron microscopy.

Migration and Motility Studies: Cultures were grown to confluency or to 1-day and 4-days PC, then subjected to HGrav and MGrav treatment for 24, 48, and 72 hr. At each time, an area of the culture was denuded to assess the migration of cells; cultures were kept under HGrav or MGrav treatment. Cultures placed in 0.5% FBS were incubated an additional 24 hr before being subjected to denudation. The migration of cells into the denuded area was monitored by video recording, under phase contrast microscopy, of random fields along the border of the denuded area. The cultures were removed from the altered gravity simulations 4, 8, 12, 16 and 24 hr after denudation, recorded then returned to HGrav or MGrav conditions in the shortest time possible (usually 5-10 min). Video image analysis was used to quantitate cell migration and motility (changes in cellular area and shape). Control cultures were treated similarly, except

they were maintained under standard cell culture conditions and normal gravity.

Statistical Analysis: Values are given as the means ± s.e.m. Tests of significance were performed using paired Student's t-test.

RESULTS AND DISCUSSION:

Morphological Assessment: Confluent BAEC subjected to hypergravity and examined by scanning electron microscopy had less surface area, fewer membrane-bound vesicles, smaller and more flattened nuclei, and membrane ruffling around the edges when compared to control cells. MGrav treated BAEC cultures were less elongated and had a more cobblestone appearance than controls. Examination of BPEC by scanning electron microscopy showed fewer surface vesicles but no other significant changes in the appearance of these cells under MGrav compared to controls.

Endothelial Cell Migration: Confluent BAEC maintained in DMEM containing 10% FBS did not show any appreciable differences in total distance migrated under control conditions, HGrav or MGrav (Fig.1). Similarly, 1-day PC BAEC did not show any significant differences under either altered gravity conditions compared to controls, when cultured in the presence of 10% FBS (Fig.1). The migration of BAEC was affected by changes in the gravity environment when cultured in low FBS (0.5%). The migration of confluent and 1-day PC BAEC was retarded by MGrav treatment but stimulated by HGrav when compared to controls (Fig. 2). The migration of 4-day PC BAEC was retarded by both MGrav and HGrav, with HGrav producing the greatest decrease in total migration.

The total distance migrated by confluent BPEC, cultured in DMEM contianing 10% FBS was only slightly retarded by MGrav conditions, but was not affected by HGrav (Fig. 3). However, as found with BAEC, BPEC cultured in low FBS was found to have significantly retarded BPEC migration under MGrav conditions, compared to controls (Fig. 3); HGrav did not produce a significant effect. The total migration of 1-day PC BPEC in low FBS medium, was initially higher (by 2 hr after denudation) under HGrav but not MGrav compared to controls (Fig. 3). However, migration for these cultures were significantly lower after 4 hr for both HGrav and MGrav conditions, with the greatest inhibition seen under MGrav.

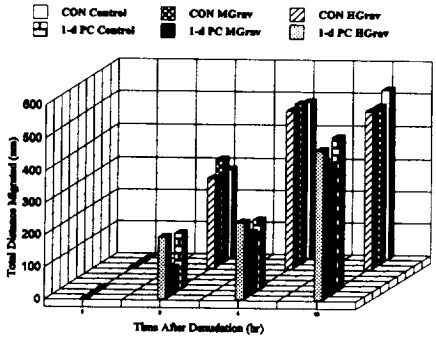


Figure 1. Alterations in BAEC migration induced by culture under HGrav and Mgrav conditons and 10% FBS: Dependency of BAEC migration on the degree of confluency

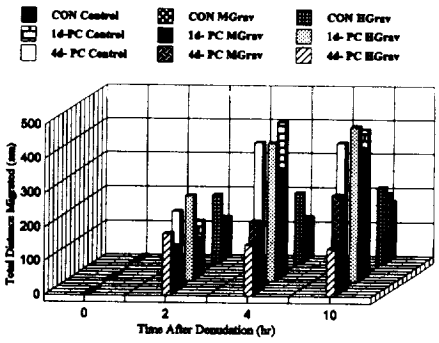


Figure 2. Alterations in BAEC migration induced by culture under HGrav and Mgrav conditons and 0.5% FBS: Dependency of BAEC migration on the degree of confluency.

Changes in Endothelial Cell Area: The results of the image analysis of endothelial cell motility are shown in Fig. 4. The 4-day PC BAEC have a 50% smaller mean cell area than confluent cell; this is seen for cell in 10% or 0.5% FBS. Increasing the serum from 0.5% to 10% resulted in a three-fold increase in cell area for confluent cell; PC BAEC cell area also increased but only by 40%.

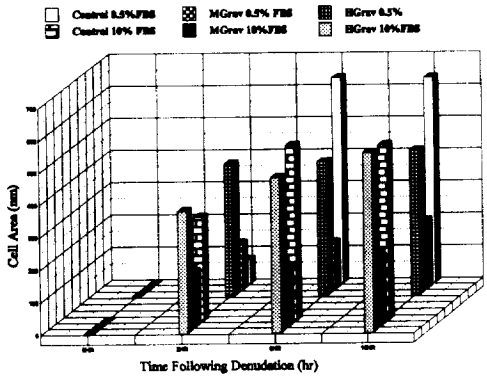


FIGURE 3. Effect of Hgrav and Mgrav on BPEC migration in the presence of 0.5% and 10% FBS.

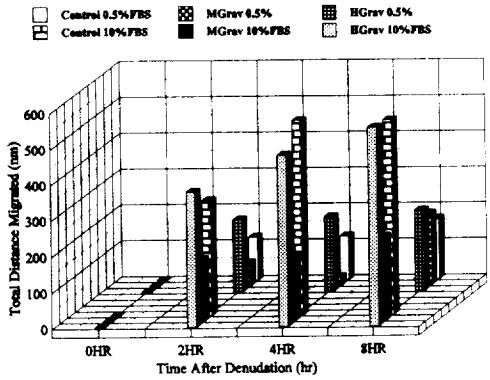


Figure 4. Effect of Hgrav and Mgrav on BAEC migration in the prsence of 0.5% and 10% FBS.

Under HGrav conditions, 4-day PC BAEC are nonresponsive to serum-induced increase cell area (or spreading). There a dramatic decrease in responsiness to serum-induced increase in cell area for confluent BAEC; cell area increased only 75% compared to 300% for control cells. MGrav inhibited the responsiness of PC BAEC to serum-induced increase in cell area but resulted in an actual reversal of the serum effect with confluent BAEC. In the latter case, cells area decreased by 40% when serum levels increased from 0.5% to 10%.

There were no significant differences found for cell shape (circularity or rectangularity) measured with either BAEC or BPEC under any of the conditions studied.

These studies show that alterations in the migration of confluent BAEC and BPEC under HGrav and MGrav conditions was to a large extent masked by culture in 10% FBS. Serum contains a number of growth factors (e.g., bFGF, PDGF and TGF_β) and cytokines that have been shown to affect endothelial cell growth and migration (8,9). The effect of these strong growth and migration modulators seems to far outweigh the effects of changes in gravitational fields. In the presence of low serum (0.5% FBS), cultures just at confluency were found to react differently to changes in gravity. However, once cultures have gone past the point of confluency, wider and significant differences can be seen. Postconfluent cultures have greater cell:cell contacts (10), have deposited more of an underlying extracellular matrix (11) and have formed greater cell:matrix contacts. The interactions of cells with extracellular matrices is thought to modulate a number of cellular response, including altering the proliferative response to several growth mitogens or the chemotactic response to cytokines (12). Our findings suggest that cell:cell and cell:matrix interactions are of major importance in the response of endothelial cells to a denudation injury, as well as, to changes in gravity.

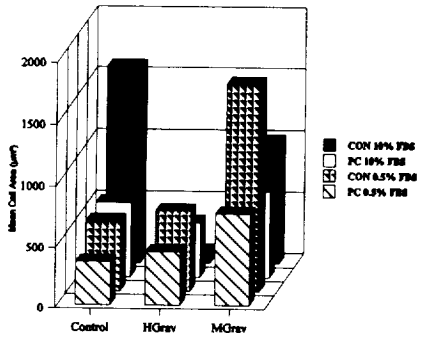


Figure 5. Changes in migrating BAEC area following culture under Hgrav and Mgrav conditions: dependency on the degree of confluency and serum levels.

ACKNOWLEDGEMENT: These studies were supported by grants NASA NAG 9-644 and NIH/RCMI RR03034

REFERENCES

1. Rijken PJ, et al. (1991). Aviat Space Environ Med 62:32-36.
2. Kumei Y, et al. (1989). J Cell Sci 93:221-226.
3. Cogoli A, Tschogg A and Fuchs-Bislin P (1984). Science 225:228-230.
4. Gruener R and Hoeger G (1990). Am. J Physiol 258:C489-C494.
5. Doty S, E Holton, G Durnova and A Kaplansky. (1990) FASEB J 4:16.
6. Philpott D, I Popova, K Kato, et al. (1990) FASEB J 4:73.
7. Nixon J, R Murray, C Byrant, et al. (1979). J Appl Physiol 46:541.
8. Nilsson J, Volk-Jovinge S, Svensson J, Landou C, DeFaire U and Hamsten A (1992) J Intern Med. 232:397-404.
9. Rosen EM, Meromsky L, Setter E, Vinter DW and Goldberg ID (1990). Exp Cell Res 186:22-31.
10. Finbow ME and Pitts JD (1981). Exp Cell Res 131:1-3
11. Pratt BM, Form D and Madri JA (1985). Annal NY Acad Sci 460:274-288.
12. Madri JA, Bell L, Marx M, Merwin JR, Basson C and Prinz C (1991). J Cell Biochem 45:123-130.

Galaptin Mediates the Effect of Hypergravity on Vascular Smooth Muscle Cell (SMC) Adhesion to Laminin Containing Matrices.

Fatisha T. Enahora, Francis N. Bosah, Sandra Harris-Hooker and Gary L. Sanford
Department of Biological Sciences, Clark Atlanta University and
The Departments of Biochemistry and Medicine, Morehouse School of Medicine
Atlanta, Georgia

ABSTRACT: Galaptin, an endogenous ß-galactoside specific lectin, has been reported to bind to laminin and subsequently decrease the binding of SMC. Cellular function depend on cell:matrix interactions. Hypergravity (HGrav) affect a number of cellular functions, yet little is known about its affect on cell adhesion. We examined the possibility that galaptin mediates the effects of hypergravity on SMC adherence. Confluent primate aorta SMC cultures were subjected to Hgrav (centrifuged at 6G) for 24 and 48 hr. Cells were non-enzymatically dispersed, pretreated with antisense (AS-oligo) or control sense (SS-oligo) oligonucleotides to galaptin mRNA (0.01 µg/ml), then seeded in uncoated or ECL-matrix coated plates. Adhesion of cells were monitored after 6 hr. HGrav increased adhesion by 100-300% compared to controls. AS-oligo decreased adhesion for both HGrav and control cells. SS-oligo did not affect adhesion for either HGrav or control cells. These studies show that HGrav affects cell adhesion and that galaptin expression is required for this effect.

INTRODUCTION: Little has been done to explore the cellular effects of gravity with vascular smooth muscle cells. However, with the development of the international space station, there is an increasing need for insight and understanding of the cellular and physiological roles of gravity (both microgravity and hypergravity). Decriptive studies have found that microgravity depresses the stimulation of lymphocytes by con A while hypergravity enhances con A stimulation of lymphocytes (1); microgravity diminishes while hypergravity stimulated the proliferative capacity of HeLa cells (2); simulated microgravity enhances EGF-induced cell round of A431 epidermal carcinoma cells, whereas hypergravity had little effect (3); synapse formation is sensitive to gravity (4); Hela cells, fibroblasts and lymphocytes cultured in a hypergravitational field were shown to have a 20-30% increase in proliferation rate but decreased glucose consumption and movement (5). These studies indicate the wide range of effects of gravity. More mechanistic studies are in the embryonic stage of research as indicated by the reports on the effects of microgravity on c-fos and c-jun expression (6); microgravity decreased EGF and TPA-induced expression of c-fos and c-jun but not A23187 or forskolin-induced expression of these genes; We have found that vascular SMC growth and movement are affected by gravity, suggesting that cell:cell and cell:matrix interactions may depend greatly on gravity. The studies reported here examines how SMC adhesion to extracellular matrices containing laminin, respond to increasing periods of exposure to hypergravity. We also examined the role of galaptin in mediating the effects of hypergravity, using an antisense oligonucleotide approach. Galaptin is a 14.5 kD ß-galactoside binding lectin which will bind laminin and is produced by a number of cell types including SMC (7-9).

EXPERIMENTAL METHODS: Primate aorta SMC were obtained from the laboratory of Dr. Sandra Harris-Hooker, Department of Medicine, Morehouse School of Medicine. Cells were maintained in DMEM containing 10% fetal bovine serum and 1X antibiotics. The oligonucleotides used were synthesized by Clark Atlanta University Molecular Biology Research Center. The following oligos were prepared based on the published nucleotide sequence for rat galaptin mRNA (10). The antisense oligonucleotide was complimentary to the first twenty bases of the galaptin cDNA. The random strand was randomly selected and it was not complimentary to any known sequence.

Cell Adhesion Assay The adhesion of vascular SMC to extracellular matrix proteins coated plates were assessed using a modification of previously reported methods (11, 12). Stock solutions of laminin and ECL-matrix proteins (1 mg/ml PBS, pH 7.4) were mixed 2% gelatin. Each well of 6-well or 12-well plates was coated with either laminin or ECL-matrix in 2% gelatin (1:9 v/v). The plates were washed with Hanks Buffered Salt Solution and used to assess the binding of cells. To evaluate the effect of antisense oligo on cell adhesion, vascular SMC were incubated with antisense, sense and random oligos (1 µg/ml each for the non-HGrav study and 0.01 µg/ml for the HGrav study). Cultures were treated with oligos in separate 25 cm^2 flasks for 24 hr under incubator

ANTISENSE: 5'-AACTTCTAATTCACACACCG-3'
SENSE: 5'-TCTCTTCGCTTCAATCATGG-3'

conditions or placed under 6 G (HGrav) for 24, 48 or 72 hr. Next, treated cells ($1x10^5$ or $1x10^4$, depending on the experiment) were added to each well of coated plates and incubated at 37°C in a humidified 5% CO_2 atmosphere for 2, 4 and 6 hr; HGrav treated cultures were incubated for a fixed time of 6 hr. Each well was washed twice with PBS, pH 7.4 to remove unattached cells and adhered cells in each well were detached, with either 0.25% trypsin-EDTA or non-enzymatic cell dissociation solution (Sigma Chemicals) and counted on the hemocytometer.

Statistical Analysis: Comparisons were made between the antisense and control (untreated, sense and random oligos) cultured cells. Data are represented as mean ± standard error of the mean (SEM). Statistical differences between the antisense oligos treated and control cultured cells were determined by a two-way analysis of variance (ANOVA).

RESULTS AND DISCUSSION:

Non-HGrav Studies: The ability of the untreated, antisense, sense and random oligos treated cells to adhere to ECM coated surfaces was examined. The data showed that the attachment of vascular SMC to the ECM proteins were significantly inhibited when cultures were treated with antisense oligos. But, the attachment of SMC treated with sense and random oligos to ECM proteins were similar to untreated control cells. The adherence of SMC to ECL-matrix after 2 and 4 hr of incubation with the antisense oligos was significantly decreased compared to the untreated, sense and random oligos treated cells. We found a 30% and 34% inhibition cells adherance to ECL-matrices at 2 and 4 hr, following treatment with antisense oligos (Fig. 1); this increased to 55% by 6 hr of incubation. The number of SMC adhering to either ECL or laminin after 6 hr was only $2.82\pm0.44x10^4$ and $2.25\pm0.38x10^4$ of the antisense oligos treated cells adhered to the ECL and laminin coated plates, compared to $5.25\text{-}5.85x10^4$ sense oligos, random oligos treated or untreated cells (Figure 2). This was statistically significant at $p < 0.01$.

HGrav Studies: In contrast to the finding for non-HGrav cultures, SMC treated with oligos and subjected to HGrav did not produce similar results. Table 1 shows the adherence of treated cells to ECL coated plates. Note that control cells (0 time under HGrav) has a 30% increase in adhering cells for antisense oligo (low oligo concentration) treated cultures; the sense oligo treated cells had a 20% decrease in adherence conpared to untreated control cells. The adherence, to ECL matrices, of cells placed under HGrav for 24 or 48 hr appear to be unresponsive to either antisense or sense oligos. Similarly, the adherence of treated cells to uncoated

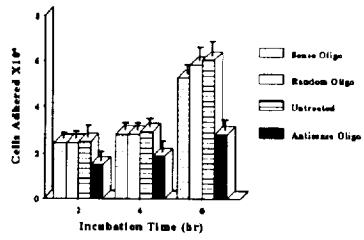


Figure 2. The effect of oligos treament on SMC adhesion to ECL coated plates. The antisense treated cell adhesion was signifcantly lower than untreated cells at $p > 0.01$.

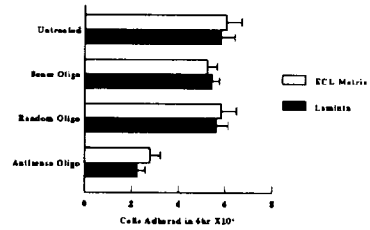


Figure 1. The effect of different oligos on the SMC adherence to ECL or laminin coated plates. Antisense treated cell adhesion was significantly decreased compared to untreated cells, $p > 0.01$.

plates shows a similar pattern (Table 2). Control cells treated with low concentration of antisense oligos have a 50% increase in adherence compared to untreated cells. Cells treated with antisense oligo under HGrav show no change in adherence to uncoated plates. The adherence to uncoated plates of cells treated with the sense oligo under HGrav is similar to that found for adherence to ECL matrices.

These results indicate that oligos may have different

Table I. The effect of Hgrav treatment on oligo induced changes in SMC adherence to ECL coated plates (Cell number $X10^3$).

Time at 6 G	Sense Oligo	Antisense Oligo	Untreated
0	7.70±0.49	13.4±1.23	9.60±0.42
24	5.62±0.77	4.20±0.74	3.75±0.67
48	5.42±0.57	4.10±0.63	4.76±0.44

Table II. The effect of Hgrav treament on oligo induced changes in SMC Adherence to Uncoated Plates (Cell Number $X10^3$).

Time at 6 G	Sense Oligo	Antisense Oligo	Untreated
0	10.6±0.91	12.1±0.74	7.95±0.49
24	5.67±0.67	4.86±0.47	5.00±0.84
48	6.87±0.79	4.17±0.54	5.27±-.49

effects on SMC depending on concentration high concentrations of antisense oligos inhibited SMC adherence to both ECL and laminin matrices. Whereas low concentrations of antisense oligo actually resulted in higher adherence than untreated or sense treated SMC. It appears that HGrav significantly lowers the adherence of cells to laminin coating matrces (e.g., ECL) as shown in Table 1; it can be seen that there is a 60% decrease in adherence of untreated and antisense treated cells when SMC are cultured under HGrav.

These studies suggest that the cellular effect(s) of HGrav may be partially mediated through the expression of galaptin. However, the exact nature of this relationship remains to be elucidated.

ACKNOWLEDGEMENT: These studies are supported by a NASA FAR Grant NAG9-852 and NIH/NIGMS S066GM08248.

REFERENCES:

1. Cogoli A, Tschogg A and Fuchs-Bislin P (1984). Science 225:228-230.
2. Kumei Y, et al. (1989). J Cell Sci 93:221-226.
3. Rijken PJ, et al. (1991). Aviat Space Environ Med 62:32-36.
4. Gruener R and Hoeger G (1990). Am. J Physiol 258:C489-C494.
5. Tschopp A and Cognoli A (1983). Experientia 39:1323-1329.
6. deGroot RP, et al. (1990). Aviat Space Environ Med 62:32-40.
7. Bosah, F.N. and Sanford, G.L. (1994). FASEB J., 8:1553.
8. Wells V and Mallucci L (1991). Cell 64:91-97.
9. Cooper DNW, Massa SM and Barondes SH (1991). J Cell Biol 115:1437-1448.
10. Clerch, L. B., and et al. (1988). J. Biol. Chem., 262:11532-11538.
11. Klebe, R. J. (1974). Nature, 250:248-251.
12. Aumailley, M., Nurcombe, V., Edger, D., Pualsson, M. and Timpl, R. (1987). J. Biol. Chem., 262:11532-11538.

SIMULATED HYPERGRAVITY ALTERS VASCULAR SMOOTH MUSCLE CELL PROLIFERATION AND MOTILITY.

Shameka Hunt, Barika Bettis, Sandra Harris-Hooker, and Gary L. Sanford,
Department of Biology, Georgia State University and
The Departments of Medicine and Biochemistry, Morehouse School of Medicine,
Atlanta, Georgia.

ABSTRACT: The cellular effects of gravity are poorly understood due to its constancy and nonavailability of altered gravitational models. Such an understanding is crucial for prolonged space flights. In these studies, we assessed the influence of centrifugation at 6G (HGrav) on vascular smooth muscle (SMC) motility and proliferation. Cells were: a) plated at low density and subjected to HGrav for 24-72 hr for proliferation studies, or b) grown to confluency, subjected to HGrav, mechanically denuded and monitored for cell movement into the denuded area. Controls were maintained under normogravity. SMC showed a 50% inhibition of growth under HGrav and 10% serum; HGrav and low serum resulted in greater growth inhibition. The rate of movement of SMC into the denuded area was 2-3-fold higher under HGrav in low serum compared to controls, but similar in 10% serum. These studies show that HGrav has significant effects on SMC growth and motility, which are dependent on serum levels.

INTRODUCTION: The cellular effects of gravity are poorly understood; the responses of most cells are probably similar to their response to other environmental factors. However, because of its constancy and the nonavailability of altered gravitational environments, little has been done in this area. The need for prolonged space flights and the development of a space station have led to questions about the role of gravity in cellular function, development and maturation. Studies of whole animals subjected to spaceflights have shown the presence of lipid accumulations within the lumen of the vasculature and degeneration of the vascular wall (1). Capillary alterations with extensive endothelial invaginations were also seen (2). Hemodynamic studies have shown that there is a redistribution of blood from the lower extremities to the upper part of the body; this will alter vascular permeability, resulting in leakage into surrounding tissues (3). Short exposures to microgravity produced marked alterations in lung blood flow and ventilation distribution (4). These studies indicate that changes in gravitational fields will alter the morphology and function of the vascular system. However, few studies have been directed at assessing the effect of altered gravitational fields on vascular cell function and metabolism.

METHODS: Primate SMC were obtained from the Coriell Cell Repository. Cells were adapted to and maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum and 1X antibiotics. In these studies, we assessed the influence of centrifugation at 5-6G (HGrav) on SMC proliferation, migration and morphometry.

Proliferation Studies: For the proliferation studies, cells were initially plated at low density in medium containing 10% FBS and then synchronized by serum depletion for 48 hr. The medium was replenished with medium containing either 10% or 0.5% FBS and then subjected to HGrav for 24-72 hr. Following

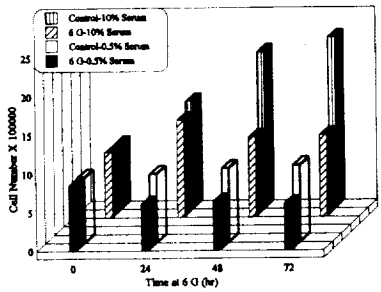


Figure 1. Effect of HGrav on SMC growth: modulation by serum.

centrifugation the cells were harvested and cell number determined using a Coulter counter.

Migration Studies: Cultures were grown to confluency and subjected to HGrav treatment for 24, 48 and 72 hr. Subsequently, an area of the cultures was mechanically denuded and the migration of cells into the denuded area measured by video image analysis. Control cultures were treated similarly, except maintained under normal gravity. The cultures

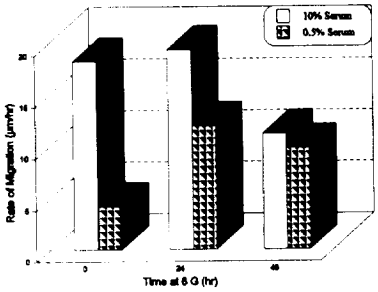


Figure 2. Serum modulation of HGrav effect on SMC migration following denudation.

were also analyzed for cell shape and area changes.

Statistical Analysis: Values are given as the means ± sem of n experiments. Within each experiment, cell number determination was made in triplicate and migration, cell shape and area measurements in septuplet. Tests of significance were performed using paired Student's t-test.

RESULTS AND DISCUSSION:

Cell Morphology: Sparse SMC exposed to HGrav were more enlarged with raised granulated nuclei, and had more cytoplasmic interactions between cells compared to controls. These effects were reversed after 72 hr under normogravity in fresh medium. Conditioned medium from HGrav-treated SMC induced similar morphological changes in control cultures. Confluent SMC subjected to HGrav had increased granularity and were more compact than controls.

Proliferation Studies: SMC growth is shown in Figure 1. Control cells, in 10% FBS, showed a 3-fold increase in cell number over a 72-hr period. Cells that were grown under HGrav in 10% FBS were significantly growth arrested over this same period, when

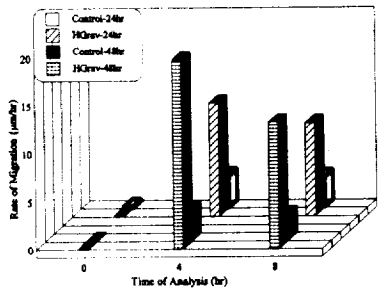


Figure 3. Random motility of sparse SMC under HGrav and control conditions.

compared to controls ($p > 0.01$). Control cells grown in 0.5% FBS were growth arrested over the study period; HGrav treated cells

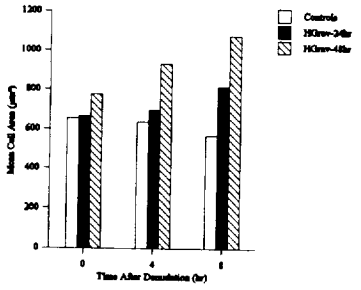


Figure 4. Changes in cell area of SMC, responding to denudation, following culture under HGrav conditions and 10% FBS.

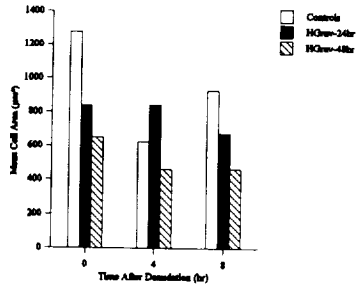


Figure 5. Changes in cell area of SMC, responding to denudation, following culture under HGrav conditions in 0.5% FBS.

under these conditions had a 25% loss in cell number. HGrav treatment resulted in the same degree of growth inhibition in 10% FBS as that found for control cells in 0.5% FBS. This suggest that HGrav treatment blocks the same signaling pathways that serum mitogens use to stimulate proliferation.

SMC Migration: We have investigated the factors involved in SMC migration using video image analysis (5,6). The rate of migration of SMC following denudation, in 10% FBS, was not altered by HGrav in 24 hr but showed a 50% inhibition by 48 hr (Figure 2). However, cultures in 0.5% FBS had a rate of migration that was 2-fold higher under HGrav conditions than for matched controls. This increased rate was seen throughout the 48 hr study period.

The migration of sparse SMC is stimulated by HGrav treatment as shown in Figure 3. This stimulatory effect increased with increasing time of HGrav treatment, with a maximal stimulation of 400%.

SMC Morphometry: Video image analysis of SMC has been used to quantitate several features which are useful for the characterization of cell behavior in response to several factors (7). Following HGrav treatment, SMC area was increased over the 8-hr image analysis period, compared to matched controls (Figure 4). Cells showed a 40% increase in cell area after 48 hr of HGrav treatment. Figure 5 shows similar studies carried out in 0.5% FBS. In this case, HGrav treatment resulted in a decreased cell area over 48 hr, compared to matched controls. These

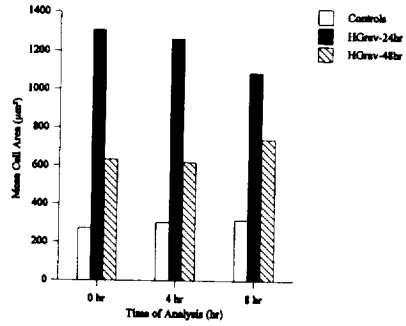


Figure 6. Changes in sparse SMC area following culture under HGrav conditions in 10% FBS.

studies suggest that control SMC, in 10% FBS, tend to move following denudation with little cell spreading. Cells maintained in 0.5% FBS are not very motile but exhibit more cell spreading, hence an increased cell area. HGrav treatment of cultures, maintained in 10% FBS, results in increased cell spreading, similar to controls under low serum conditions.

The effect of HGrav on cell area of sparse SMC is shown in Figure 6. This figure shows a dramatic increase in cell area after 24 hr of HGrav treatment, compared to matched controls. This influence on cell area decreased 30-50% by 48 hr in HGrav.

Other morphometric parameters, e.g., circularity and rectangularity are influenced to various degrees by HGrav.

ACKNOWLEDGEMENT: Supported by grants NASA NAG 9-64499 and NCCW-0083.

REFERENCES:

1. Doty S, E Holton, G Durnova and A Kaplansky. (1990) FASEB J 4:16.
2. Philpott D, I Popova, K Kato, et al. (1990) FASEB J 4:73.
3. Nixon J, R Murray, C Byrant, et al. (1979). J Appl Physiol 46:541.
4. West JB, Guy HB, Michels, DB (1982), The Physiologists 25, S21-S24.
5. Bosah, F. N, Sanford, G. L., Harris-Hooker, S. Galectin-1 Expression is required for Vascular Smooth Cell Migration during Vascular Remodeling. Manuscript in preparation
6. Matheravidathu SS, Marshall G, Sanford GL and Harris-Hooker S (1995), Protooncogene expression of vascular cells responding to mechanical injury in vitro. FASEB J 9 (Part II):1568.
7. Hooker-Harris S, Sanford GL, Montgomery V, Rivers, R. and Emmett N (1992). Cell Biol. Internat. Reports, 16:433-450

Chromium Diffusion Doping on ZnSe Crystals

Troy. D. Journigan, K.-T. Chen, H. Chen, A. Burger
NASA/Fisk Center for Photonic Materials and Devices
Department of Physics, Fisk University
Nashville, Tennessee 37208
and
K. Schaffers, R. H. Page and S. A. Payne
Advanced Solid State Laser Group
Lawrence Livermore National Laboratories
Livermore, CA 94550

Abstract

Chromium doped zinc selenide crystal have recently been demonstrated to be a promising material for near-IR room temperature tunable lasers which has a emission range of 2 - 3 μm. In this study a new diffusion doping process has been developed for incorporation of Cr^{+2} ion into ZnSe wafers. This process has been successfully performed under isothermal conditions, at temperatures above 800°C. Concentrations in excess of 10^{19} Cr^{+2} ions/cm^3, an order of magnitude larger than previously reported in melt grown ZnSe material, have been obtained by diffusion doping, as estimated from optical absorption measurements. The diffusivity was estimated to be about 10^{-8} cm^2/sec using a thin film diffusion model. Resistivity was derived from current-voltage measurements and in the of 10^{13} and 10^{16} Ω-cm. The emission spectra and temperature dependent lifetime data will also be presented and discussed.

Keyword: Tunable solid state Laser, Cr^{+2} doped ZnSe, Diffusion Doping, Current-Voltage Measurement, Lifetime.

Introduction

Room temperature solid state lasers capable of generating tunable emission in the mid-infrared (IR) region using optical parametric oscillators/ amplifiers (OPO/OPA) are currently employed with certain pulse format requirements. The tunability and wavelength of solid state laser materials operating in the mid-IR have been limited due to the competition of nonradioactive decay as the emission bandwidth increases and as the wavelength of interest become longer.

Recently a new class of transition metal-doped (Cr, Ni, Co) Zn chalcogenides (Se, S, Te) materials have been investigated as potential candidates as tunable solid state lasers with a spectral emission range of 1-4 μm [1,2]. The absorption and emission spectroscopic properties have been measured to evaluate their potential for room temperature lasing in the 1-4 μm. Notably, the emission cross section and temperature-dependent lifetimes have been measured and they indicated that chromium ion has a potentially high gain cross section and low non-radiactive decay losses resulting in a high emission quantum yield. The quantum yield efficiencies are generally approaching 100% for chromium, in contrast to nearly 20% for cobalt, while nickel exhibited no room temperature luminescence. Based on the previous results, chromium doped ZnSe was recognized to have spectroscopic properties with efficient lasing with diode pump laser and appears to have low level of the deleterious effect of nonradioactive decay. A pump volume of 10 × 6 ×3 mm^3 and a Cr^{+2} concentration of 1.5×10^{20} ion/cm^3 are needed for a 1.63 μm (40 W, InGaAsP diode laser) pump light to be absorbed within about 100 - 200 μm from surface [1,2]. ZnSe has been previously doped with Cu, Ag, Ga, In, Ni and Li for the achievement of suitable electrical properties. It has been reported that precipitation was observed for Ga [3], In [4] and Li [5] doped ZnSe. ZnSe has been doped with Cr^{+2} ion to investigate optical properties [6,7]. Based on the analysis of superhyperfine interactions with neighbors it has been reported that the chromium ion incorporated substitutionally into Zn lattice sites [8,9]. Up to date, there is little information about the usage of Cr doped ZnSe as solid state laser material and doping process which can

achieve Cr concentration as high as 10^{20} ion/cm^3. The objectives of this project are to develop a new combination of doping and diffusion process of Cr^{+2} ion into ZnSe, diffusion properties.

Experimental

The purification and growth ampoules were made from fused quartz. The cleaning processes of ampoules includes two steps: (a) the ampoules were cleaned with aqua regia (25%HNO_3 + 75% HCl) and rinsed with deionized water; (b) the ampoules were then heated at 100 °C above the diffusion temperature under 10^{-6} Torr vacuum. The Bridgman or seeded physical vapor transport (SPVT) grown ZnSe wafer, which were purchased from Eagle-Picher company , were loaded into pre-cleaned ampoule along with CrSe powder (4N grade from Alfa company). The loaded ampoule was then sealed under higher range of 10^{-7} Torr vacuum. The sealed ampoule (typical length of 10 cm and 1.2 cm ID) was inserted into the middle section of a 12 inches long heat pipe which provided a isotherm environment in the middle 10 inches region between 400 and 1100 °C, and the heat pipe was then placed in a two zone furnace. A schematic drawing of the experimental arrangement is shown in figure 1. Several diffusion experiments have performed under this doping process with temperatures from 800 to 1000 °C for either 2 or 5 days annealing times.

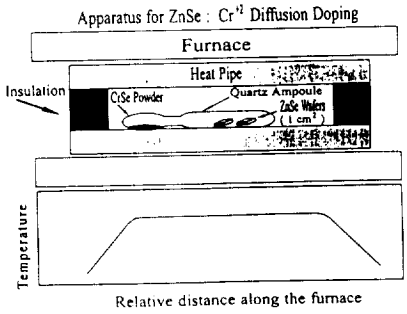


Figure 1. Experimental arrangement of diffusion doping process.

After annealing, the temperature was lowered by 5 °C/min to 400 °C, and then the furnace was shutdown and cooled to room temperature. The Cr^{+2} concentration, laser lifetime and other optical properties were measured using optical absorption, inductively coupled plasma-mass spectroscopy and a solid state laser testing by a solid state laser group in Lawrence Livermore National Laboratory [1,2]. The atomic force microscopy (AFM) and current-voltage (I-V) measurement were performed at Fisk. Silver conducting paint and copper wire were used as metal contact and lead wire for connection to the I-V electronic system. A typical increment of 1 V and 1 min waiting time was used.

Results and Discussion

Table 1 lists the doping conditions, Cr^{+2} concentration, laser lifetime and resistivity. First column indicate the crystal growth method and its structure format. The first seven samples were grown by Bridgman method, and they are in form of polycrystalline (Poly.) and single crystalline with twin (Twined).

Table 1. Summary of diffusion doping parameter, resulting Cr^{+2} concentration, resistivity and lifetime.

Sample	Annealing Temperature (°C)	Annealing Time (days)	Concentration Cr^{+2} (cm^{-3})×10^{20}	Lifetime (μsec)	Resistivity (Ω-cm)
#12413: Bridg. Poly.	800	2	0.05	6	-
#12413A: Bridg. Poly.	850	5	0.14	5	-
#12429: Bridg. Twinned-top	900	5	0.19	5	-
#12429A: Bridg. Twinned-bot.	900	5	0.25	-	2.5×10^{7}
#12428: Bridg. Poly.	900	5	0.18	3	6.0×10^{6}
#12427: Bridg. Poly-top	950	2	0.38	2.5	2.0×10^{7}
#12426: Bridg. Poly-bot.	950	2	0.33	3.1	8.3×10^{7}
#12430: SPVT, Single	950	2	0.17	6	9.9×10^{7}
SPVT, Single	1000	5.5	1.01	-	1.9×10^{6}

The last two sample were single crystalline grown from seeded physical vapor transport method. The annealing temperature and Cr^{+2} concentration have plotted in figure 2. For both 2 and 5 days annealed samples, the Cr^{+2} concentration increases with annealing temperatures. This trend indicates that for 1000 °C and 5 days annealing conditions we have not reached a saturation that would arise from the existence of a Cr^{+2} solubility limit in ZnSe. By annealing either at higher temperature and/or longer time one can achieve a higher Cr^{+2} concentration in doped ZnSe crystals. The linear least square fit lines (dashed line-2 days, solid line-5 days) seem to be fairly parallel to each

other. The result propose that the same Cr^{+2} concentration can be obtained by annealing at low temperature for longer time as annealing at higher temperature.

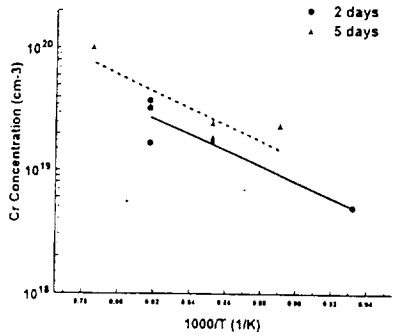


Figure 2. Plot of Cr^{+2} concentration under various annealing temperatures and times.

The possible disadvantages of high temperature annealing are the increase of the concentration of defects, such as vacancies and dislocations, and the deviation from stoichiometry of ZnSe. All samples have been polished after diffusion doping process, and the colors are from light reddish to darker reddish as the Cr^{+2} concentration increasing. Except for the sample annealed at 1000 °C, the color of this sample is the darkest of all sample, and it can not see through even with backlighting. Due to the fact that the crystal growth temperature of SPVT grown ZnSe is around 1100 °C, it is a reasonable assumption that 1000 °C is the close to the limit of annealing temperature for SPVT grown ZnSe wafer. The various mechanisms affecting the optical properties as solid state laser by means of lifetime. The lifetime of single and polycrystalline wafer does not show a clear trend. This result may be a good sign for industrial manufacture since that polycrystalline ZnSe has a much more lower production cost the single crystalline material. Figure 3. show a typical I-V measurement result on the sample annealed at 1000 °C and 5 days.

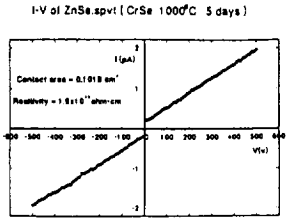


Figure 3. A typical I-V plot for sample annealed at 1000 °C and 5 days.

The resistance was derived from the slope of I-V curve, and the resistivity was calculated by taking account of the contact area and thickness of sample. The resistivity of Cr^{+2} doped ZnSe 5-8 times order of magnitude higher than undoped ZnSe. By assuming the equilibrium diffusion depth as 0.1 cm, which is deeper than the requirement to be used as solid state laser material, the diffusivity, D, can be estimated by a thin film diffusion model [10]. In this model, the equilibrium diffusion depth, x, is defined as $x = 2(Dt)^{\frac{1}{2}}$, where t is the diffusion time. By applying 0.1 cm for x and 5 days annealing time, the diffusivity was estimated to be in the 10^{-8} cm^2/sec which lower than Ag (~ 10^{-6}) [11] and Cu (~ 10^{-7}) [12]. Figure 4 show the AFM images on freshly cleaved Cr^{+2} doped ZnSe for 2 days at 950 °C

Summary

A new Cr^{+2} diffusion doped ZnSe process has been developed and demonstrated to yield Cr^{+2} concentrations as high as 10^{20} ions/cm^3 which is required for tunable solid state laser material use. The crystallinity of ZnSe wafers seem to be have significant influence on the Cr^{+2} concentration in this doping process. The Cr^{+2} concentration has not reached its solubility limit in ZnSe solid solution, and it could in the future be increased by annealing at temperature above 1000 °C or longer time at lower temperature. The resistivity of doped ZnSe increases dramatically by 5-7 orders of magnitude, in the range of 10^{13} - 10^{16} Ω-cm compared with undoped materials. The diffusivity of Cr^{+2} in ZnSe has been estimated to be in the 10^{-8} cm^2/sec range.

Acknowledgments

Funding for this project have been supported by NASA through NASA/Fisk Center for Photonic Materials and Devices and by the DOE grant No DE-F603-94 SF20368. The authors at Fisk University would like to acknowledge Dr. Kennedy J. Reed of Lawrence Livermore National Laboratory and Research Collaboration Program (RCP) for HBCU's through the LLNL Education Programs Offices.

References

[1] L. D. DeLoach, R. H. Page, G. D. Wilke, S. A. Payne, and W. F. Krupke, *CALIOPE Anual Meeting Technical Review Proceeding*, **2**, July (1994).

[2] L. D. DeLoach, P. H. Page, G. D. Wike, S. A. Payne, and W. F. Krupke, *Advanced Solid State Lasers Conference*, Menphis, Tennessee, January (1995).

[3] A. K. Ray and F. A. Kroger, *J. Electrochem. Soc.*, **125**, 1355 (1978).

[4] G. Jones and J. Woods, *J. Phys. D.*, **9**, 799 (1976).

[5] G. F. Neumark and S. P. Herko, *J. Cryst. Growth*, **59**, 189 (1982).

[6] J. T. Vallin and G. D. Watkins, *Physical review B*, **9 (3)**, 2051 (1974).

[7] G. Goetz, H. Zimmermann, and H.-J. Schulz, *Z. Phys. B*, **91**, 429 (1993).

[8] M. DeWit, A. R. Reinberg, W. C. Holton, and T. L. Estle, *Bull. Am. Phys. Soc.*, **10**, 329 (1965).

[9] T. L. Estle and W. C. Holton, *Phys. Rev.*, **150**, 159 (1966).

[10] Paul G. Shewmon, *"Diffusion in Solids"*, McGraw-Hill Book Company, London (1985).

[11] T. Lukaszewicz and J. Zmija, *J. Phys. Status Solidi A*, **62**, 696 (1980).

[12] M. Aven and R. E. Halstead, *Phys. Rev.*, **137**, A228 (1965).

Biography

Troy. D. Journigan is a graduate student in Master program in Chemistry, and mainly involves tunable solid state laser materials research project in NASA Center. He involve in this project since summer of 1995 when he was in summer research program, Advanced Solid State Laser group, Lawrence Livermore National Laboratory. He is interest in similar research and solid state Physics, and will continue to pursue his Ph.D. degree after graduate at Fisk University.

Momentum Impact Studies on Solar Cells

James Derrick Conner[1] and W. Eugene Collins
NASA/Fisk Center for Photonic Materials and Devices
Fisk University
Nashville, TN 37208-3051
and
S. Ray Bullock, NASA LeRC, Cleveland, Ohio

Introduction:

One efficient way of propelling a spacecraft in space is by means of a rocket (thruster). It is important to understand the process which governs the action. The propulsion system of all rockets consists of a fuel storage system of which there are many different types, due to the different varieties of fuels. There must also be a fuel transport system. The fuel is then expanded in an expansion region, triggered by some form of combustion, and is then accelerated to the nozzle region where it is expelled through the nozzle. The expelled propellent from the rocket comes out and creates a plume. The plume's interaction with spacecraft structures is the focus of this study .

When a plume impinges on a spacecraft's surface there are different types of impingement effects involved, some of them being contamination, heat load transfer, and momentum transfer. Contamination has the possibility of two effects on a spacecraft. The exhaust may leave behind fuel droplets which can adhere to solar array panels, optical sensors, and lens; such an effect would greatly diminish the efficiency or impede the normal operation of those pieces of equipment. The other effect is the possible chemical reaction which might occur when the plume chemically interacts with the alloys of the various components of the spacecraft leading to a decrease of their efficiency. These effects could significantly effect the success of a space mission. Another effect of plume impingement is the heat load transfer from an exhaust plume impinging on spacecraft surfaces of instruments that may be heat sensitive. Surfaces exposed to extreme thermal load may change the rigidity of an alloy making it malleable or brittle and making it vulnerable to space debris. This type of effect can also endanger the entire the spacecraft. There has been extensive research done on contamination effects and heat load transfer effects.

The final effect of plume impingement, momentum transfer, is the subject of this paper. Momentum transfer of the exhaust plume effects the spacecraft's orientation in space. When a plume impinges on a surface, it creates a torque on the spacecraft in a direction away from the intended direction of thrust. It therefore takes additional thrusts to oppose the induced torque. This is important because it takes additional fuel to execute the correcting maneuvers, and the fuel load is a part of the total payload capacity of a spacecraft and is important in deferring cost a launch and the life span of the mission. In this paper our goal is find a method that can accurately measure the momentum imparted to a spacecraft component. [1]

Other work that has been done on plume impingement involved modeling experiments with comparisons to real space flight data retrieved from previous space flight missions. Scientists, such as Parvez[2], Fox[3], and Allegrè , worked on the analytical aspect of the impingement problem. Legge[5], Dettleff[6] others worked on direct measurements of the change in pressure that was applied to a surface. Our approach differs in that it is a direct force measurement that we hope will provide us with a device with a greater range of forces and more accurate measurement of the real force distribution.[2]

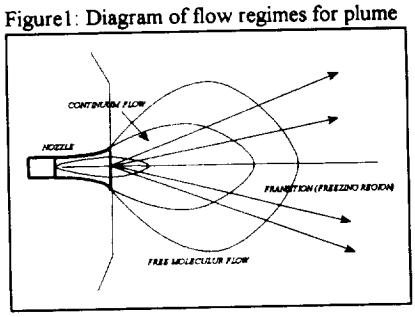
Theory:

A plume is characterized by three regions: the continuum, the transitions region, and the free molecular flow region displayed in figure 1. Each of these regions is characterized by the Knudsen number, Kn, which relates the mean free path z(λ) to a

[1] James Derrick Conner is a graduate student in the physics department and has been matriculating at Fisk University since August, 1993. His research has been on plume impingement studies with emphasis on momentum transfer effects. He will be completing his M.A. degree in May, 1996.

$$K_n = \frac{\lambda}{l_{ref}}$$

Figure1: Diagram of flow regimes for plume



characteristic length(l_{ref}). The mean free path simply means the average distance that a molecule will travel before striking or being struck by another particle. In molecule flow, the rebounding molecules will bounce from a surface unaffected by the impinging particles. For the flow that are in continuum, the Knudsen # is much less the 1(Kn << 1). The transition region has a Knudsen # almost equal to 1 (Kn = 1). The last region is the free molecular flow region, and has a Knudsen # much larger than 1 (Kn >> 1). The subdivision of the flow field is necessary because different methods must be applied to describe the flow field behavior theoretically. The Method of Characteristics, for example, is a proper tool to describe the flow field within the isentropic, continuum region. It, however, can not be applied to the free molecular flow and to the truly expanding boundary layer. In any case forces will be exerted on the wall, and because, on a satellite, the thruster and the impinged walls are mechanically connected, the effective thrust (Feff)is given by

$$F_{eff} = F - \Delta F$$

with F : nominal thrust and ΔF : thrust loss due to impingement.[3]

The flow around a body exposed to the plume is much more complicated than the usual flow around a vehicle moving through a gas at rest. The main differences are that the free stream has gradients in the flow direction, also has static variables and that the body may be larger than the flow field itself. When the gas becomes more and more rarefied the continuum character of the flow disappears. Due to the decreasing number of collisions between the gas molecules they behave like individual particles, and such behavior defines the region of the free molecular flow.[3]

The first deviation from the continuum behavior can be observed when the thickness of the impingement shock wave becomes comparable to the flow field dimension behind the shock wave. As the rarefaction proceeds further, the shock wave will vanish and in the free molecular impingement regime the plume flow will hit the surface without any disturbance by the backscattered molecules.[3]

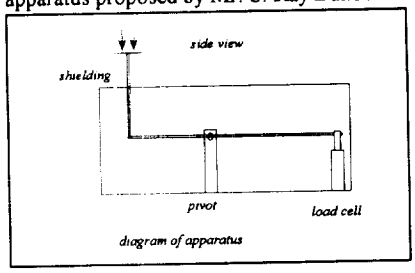
To decide whether we have an impingement situation in rarefied or in continuum gas, a careful analysis of the flow field is necessary. In the case of rarefied impingement the local sheer stress can no longer be neglected. For complete free-molecular impingement Kn>1, and both the normal pressure and the sheer stress, can be calculated in a straight forward way by use of the appropriate formulae. Accommodation coefficients have to be known and one has to remember that the usual free molecular formulae are mathematically not exact for radial flow.[3]

APPARATUS

Design Consideration:

There were many types of designs considered for the momentum transfer device and its measurements, two of which were the fluid piston device and a pendulum force balance. The lever arm force balance with a load cell was the chosen device; as shown in figure 2. It was chosen specifically for the projected ranges of force measurement that it would be capable of delivering. The idea was to have an impingement plate on one side of the lever and load cell on the other side. By changing

Figure 2: The schematic drawing the apparatus proposed by Mr. S. Ray Bullock



the position of the pivot one is able to change the scaling factor of the load cell and the impingement plate.

The lever design was primarily that of Mr. S. Ray Bullock of NASA LeRC and is illustrated in figure 2. The load cell of his design was an Entran elj load cell transducer that works in both compression and tension mode. They come with calibration certificates which give all the pertinent information about the load cells. There was another load cell introduced later into the experiment to help determine the validity of the calibration of the load cells. This was an Unimeasure/80, it also worked in compression and tension modes. This one, unlike the Entran load cell, did not come with calibration data and was used as a tool to determine calibration techniques.

The series of tests ran at NASA were bench top calibration and preliminary vacuum calibration. The test were run at the bell jar 5 facility which is a small vacuum system capable of pressures of approximately 10^{-5} torr. The first test was to establish the scale for the load cell, how many volts equals one gram. This was done by using a set of well characterized weights to monitor the change in voltage. It was then noticed after the first few runs that there was not the linear progression . The next step was to determine why they occurred. Hysteresis test were done to evaluate possible human error or product malfunction. The curve from the load cell showed poor hysteresis, and was determined to be defective. With remaining load cell we made special precautions to ensure that there would be no damage to the load cell before and during placement. This time the results were promising, a good hysteresis curve was achieved and the change of voltage as a function of mass was linear within reasonable values.

It was time for preparations for the calibration test in Bell Jar 5 vacuum facility. Bell Jar5 had previously been used for preliminary proof of concept test and was already fitted with tubing and a gas source. The only modification that was made to the bell jar was the attachments to run the motor. The motor was added to repeat the calibration in a vacuum environment. When the apparatus was placed in the vacuum the reading on the voltage meter changed as a function of time for a specific weight. It was determined that the load cell had to reach thermal equilibrium before calibration could be done. The load cell was allowed an over night warm up period. Afterward, measurements were taken again with promising results.

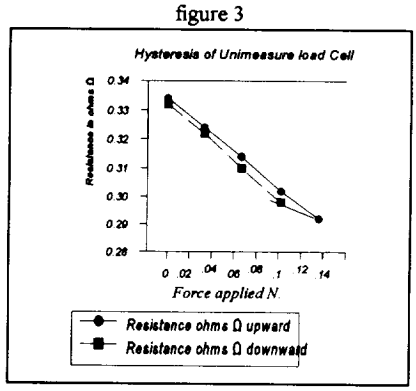
Calibrations:

A series of tests were done to determine the change in resistance of the load cell as a function of applied force. The next set of tests was a response time test where a given mass was lowered and the time it took the reading to equilibrate was measured. The final measurement was to determine the time it took the initial zero reading to return when an applied force (weight) was removed. The results of theses measurements is illustrated in figures 3 through 5.

Results:

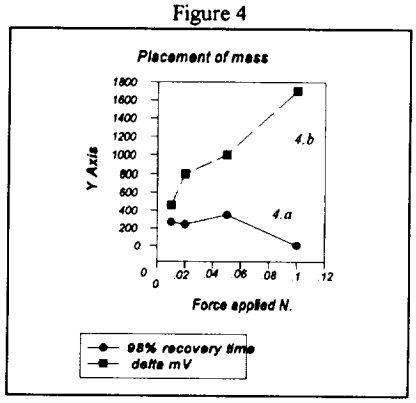
Both the Entran and the Unimeasure load cells displayed hysteresis trials as shown in figure 3, in which a graph of change in voltage is plotted against the applied force as it goes from one gram to ten grams and back again.

figure 3



From the sample data in figure 4, response time can be seen that the amount of time it took for 98% of the maximum in the change in voltage to occur for an applied

Figure 4

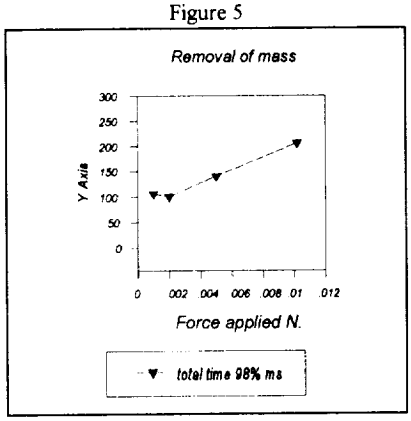


force decreased as the amount of applied force increased. The voltage for the applied force is determined after the voltage for a no-load force. Graph 4.b shows that the measurement for the changes in voltage increased in a nearly linear fashion as a function of applied force . The only nonlinear point corresponds to the 0.05 N. measurement, this point is believed to artificially high. Figure 5 shows the response time as a function of applied force is linear above 0.02 N.

Conclusions:

This device can be used to measure momentum changes, as long as one considers hysteresies, re-zeroing the device for each test and the time between measurement. For every measurement a zero value must be taken. There must be time allow between measurements for the system to equilibrate.

Figure 5



References:

[1] Lengrand, J.C., "Plume Impingement Upon Spacecraft Surfaces", Laboratoire d'Ae'reothemique du C.N.R.S.. Meuden, France.

[2] Parvez, S.A., "GSTAR Satellite Disturbance From Plume Impingement", GTE Spacenet McLean, Virginia, Paper 89-0351 at AIAA 27th Ann. Meeting, Reno NV., Jan 9-12

[3] Fox S.M., "Effects of Plume Impingement on a Momentum Bias Satellite", Journal of Spacecraft and Rockets, Vol 20, No 1 1983, pp 55-60

[4] Allegre, J., Raffin, M;Lengrand, J.C.," Investigation of Freejets Issued From a Small Solid-Propellent Rocket Engine. J Spacecraft, Vol 20, No 2, 1984. pp 131-135

[5] Legge, H., "Plume impingement Forces on Inclined Flat Plates",DLR, Institute for Experimental Fluid Mechanics, 10, D-3400 Goettingen, F.R.G

[6] Boetttcher, R. D., Dankert, C., Dettleff, G., "Studies on Rockets Exhaust Plumes and Impingement Effects on Spacecraft Surfaces; Executive Summary".DFLVR-IB222-88A11, 1982

Acknowledgements:

*Work Supported by NASA Grant numbers NAG3-1430 and NAGW-2925.

Crystal Growth of Undoped and Doped ZnSe

Swanson L. Davis, K.-T. Chen, M. A. George, D. T. Shi, W. E. Collins and Arnold Burger
NASA/Fisk Center for Photonic Materials and Devices
Department of Physics
Fisk University
Nashville, TN 37208

Abstract

The surface morphology of freshly cleaved ZnSe single crystal grown by the physical vapor transport (PVT) method was investigated by Atomic Force Microscopy (AFM) and the results were correlated with Differential Scanning Calorimetry (DSC) data. Selenium precipitates have been revealed in undoped doped ZnSe crystals having a size of about 50 nm. A transition temperature around 221 °C in the DSC measurements is interpreted as the eutectic temperature of Se-saturated ZnSe. The AFM images of doped ZnSe also shows that possible Cr clusters are uniformly distributed and they have an estimated size of about 6 nm.

Keyword: Doped and Undoped ZnSe, Se Precipitation, Atomic Force Microscopy, differential scanning calorimetry, Physical Vapor Transport.

Introduction

ZnSe based epitaxial films have attracted much interest due to its potential for blue-light-emitting diodes [1]. Extensive research efforts have been focused on the growth of thin films of ZnSe on GaAs substrates due to its high availability and relatively low lattice mismatch of 0.28% (5.6532 Å) to ZnSe (5.6676 Å) at room temperatures [2]. The heteroepitaxial approach is still not ideal because of chemical incompatibility [3,4] and structure differences [5]. It may be necessary to grow homoepitaxial layers of ZnSe on bulk ZnSe as substrate and few first attempt have been reported [6-8]. However, low temperature photoluminescence analysis indicated that the layers contain some concentration of impurities which incorporated from ZnSe substrate. In recent years, ZnSe is explored to has potential as tunable solid state laser materials when it is doped with transition metals [5,6]. Among all the dopants, Cr^{+2} shows a room temperature quantum yields of 100% which the highest among Ni and Co. Therefore the growth of high purity, ZnSe bulk crystal of high structure quality is essential to the growth of ZnSe homoepitxial layers and high efficiency of solid state laser when doped with Cr^{+2}

The objective of this study was to grow high quality ZnSe single crystals by physical vapor transport method (PVT), and to evaluate the quality of the as-grown crystals. The freshly cleaved as-grown and doped crystals were characterized by Atomic Force Microscopy (AFM) to reveal the surface morphology. DSC measurements were performed and a correlation with the AFM results was made. The inclusions are identified as having a ZnSe/Se-rich composition and a quantitative analysis of their size and distribution was accomplished.

Experimental

The growth ampoules are made of T08 quartz tubing from Heraeus Amersil with an inside diameter of 15 mm and wall of 1.5 mm thick. The ampoule is about 10 °C/min length with

acid solution and outgassed under vacuum at 1100 °C for 16 hours under higher range of 10^{-7} Torr vacuum. ZnSe starting material was purchased from Eagle-Picher, Inc. Prior to crystal growth, the starting material was heat treated at 800 °C for 4 hours under same level of vacuum to adjust its stoichiometry and to eliminate volatile impurities. The ampoules were then sealed under 10^{-7} Torr vacuum. The temperature profile of the three-zone crystal growth furnace and the ampoule position are schematically shown in figure 1.

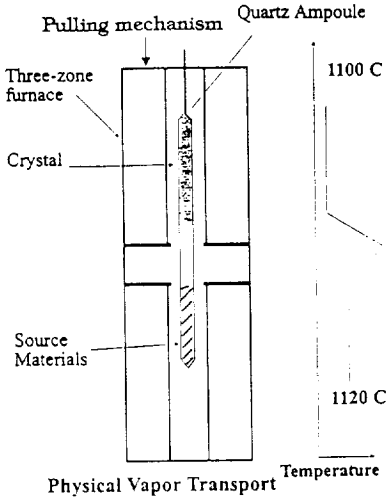


Figure 1. The physical vapor transport crystal growth system.

Sharp temperature gradients at the growth interface as well as uniform temperature plateaus in the growth and source zone were achieved by optimizing the temperature settings for the three individual zones and the outside insolation respectively. The growth ampoule was positioned inside the 3 zone furnace while the furnace was translated with respect to the growth ampoule at a 5 mm/day rate. The cooled down processes consist of a 5 °C/min cooling rate to 700 °C, and then furnace was cooled to room temperature by shutting down furnace. Some of the ZnSe wafers were purchased from Eagle-Pitcher company which were grown from seeded physical vapor transport (SPVT) method. These wafers were then diffusion doped with CrSe powder under isothermal condition. Details of the diffusion doping process are described elsewhere [7]. The AFM analysis was performed on a Digital Instruments Nanoscope II/E which had a piezoelectric tube scanner with an effective scan range from about 0.5 to 14 μm. It was calibrated using mica and a gold grid. The cantilevers were commercial Nanoprobes made of gold-coated silicon nitride. The force constant of these cantilevers employed was 0.06 N/m. The scan rate was in the 0.7-8.3 Hz range. There was no apparent tip-induced deformations of the surfaces when an average optimum force constant of 28 nN/m was used. In spite of 50% relative humidity in the laboratory the effect of humidity on AFM imaging of the ZnSe surfaces was negligable since ZnSe surfaces are very hydrophobic. This prevents sample-water-tip interactions which tend to distort the image due to pronounced capillary forces acting on the cantilever [8].

For DSC measurements, a TA Instruments Differential Scanning Calorimeter model DSC-10 was used in conjunction with a Thermal Analysis 2000 control and analysis system. A typical sample had a mass of about 100 mg. The samples were ground and sealed into stainless-steel capsules. Thermograms were recorded for heating rates of 5 °C/min in the 150-500 °C range.

Results and Discussion

The AFM image of the freshly cleaved PVT and Cr doped SPVT ZnSe surfaces show that inclusions distributed uniformly throughout surfaces. Figures 2 and 3 show the AFM images of these two crystals. The size density of inclusions on these crystals are about the same, 50 nm and 1.6×10^{12} m^{-2}, respectively. Previous studies [9] have reported inclusions in PVT grown ZnSe, and the size and density are 20 nm and 10^{13} m^{-2}, respectively.

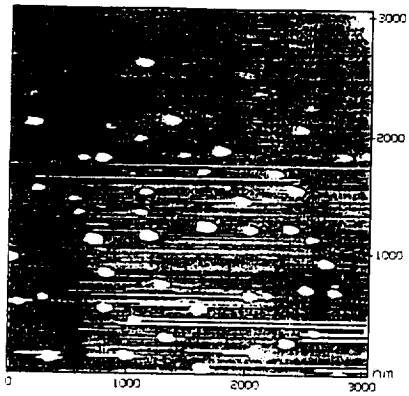


Figure 2. The AFM image of freshly cleaved PVT grown ZnSe surface. The Z rage is 8 nm.

The size of the inclusions revealed in this study is about 2.5 times larger than previously reported [9], but the density is about 10 times lower than Ref. [9]. On the Cr doped ZnSe, AFM image revealed a nano-cluster feature which do not appear on undoped crystal. The cluster size is about 6 nm with a density of about 1.8×10^{13} m^{-2}. In ZnSe crystal, only Se inclusion have been reported. Therefore these small clusters may be the accumulation of Cr instead of Zn or Se, futher verification of this assumption is needed. A typical DSC thermogram is shown in figure 5.

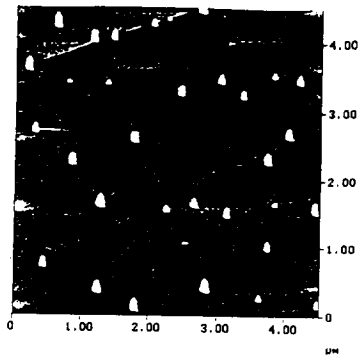


Figure 3. The AFM image of freshly cleaved Cr doped SPVT grown ZnSe surface. The Z range is 10 nm.

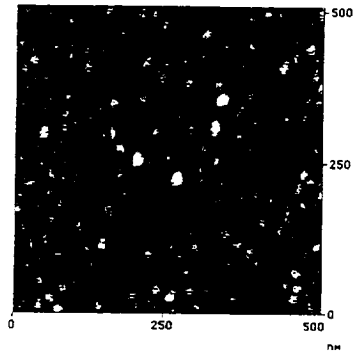


Figure 4. The AFM image of Cr doped ZnSe surface revealing small cluster. The Z range is 5 nm.

There are no other peaks below 200 °C and above 250 °C. Therefore only the 200-250 °C range has been shown. For Cr doped and undoped ZnSe, the DSC thermograms show endotherm peaks around 219 °C which in fair agreement with the eutectic temperature of Se saturated-ZnSe [10].

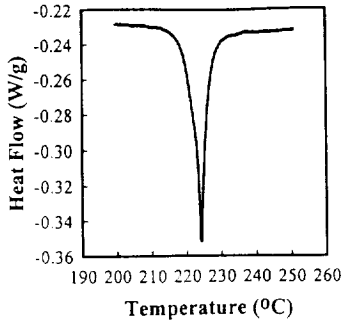


Figure 5. A typical DSC thermogram of undoped ZnSe.

There ther are no phase transition occurred around 419 °C which is the eutectic temperature of Zn saturated-ZnSe [10]. This result comfirmed that the inclusion revealed by AFM is Se-saturated with ZnSe solid solution.

Summary

In conclusion, AFM images revealed that undoped and Cr doped ZnSe consist of Se inclusions under current crystal growth conditions. Further identification was carried out by DSC measurements, and the results confirmed that excess Se has precipitated as inclusions form the composition of solid solution close to its Se-saturated state. The possible main contributing factors for the presence of Se inclusions are the stoichiometry of starting materials and the growth thermal conditions. temperatures [18] leading to deviations from the stoichiometric ratio as large as 10^{-2} (%) at the elevated temperature. Although the postulate we made here is supported by our experimental results, a better knowlege of the phase diagram in the homogeneity range of ZnSe solid solution is necessary.

Acknowledgement

Partial funding was provided by the NASA through the Fisk Center for Photonic Materials and Devices.

References

[1] R. N. Bhargarva, *J. Crystal Growth*, 86, 873 (1988).

[2] "*Table De Constants Des Semiconducteurs*", Aigrain P and Balkanski M, Oxford: Pergamon (1961).

[3] O. Briot, Thesis, Montpellier (1990).

[4] J. O. Williams, A. C. Wright and H. M. Yates, *Proc. 5t Intl. Conf. II-VI Semiconductors*, Okayama, Japan, Sept. 8-13, (1991).

[5] Laura D. DeLoach, Ralph H. Page, Gary D. Wilke, Stephen A. Payne and William F. Krupke, *CALIOPE Annual Conference, Technical Review Proceeding*, II, July (1994).

[6] L. D. DeLoach, R. H. Page, G. D. Wilke, S. A. Payne and W. F. Krupke, *Advanced Solid State Lasers Conference*, Memphis, Tennessee (1995).

[7] T. Journigan, K.-T. Chen, H. Chen, A. Burger, K. Schaffers, R. H. Page and S. A. Payne, Summited to *NASA University Centers First Students' Conference*, Greensboro, NC (1996).

[8] T. Thundat, R. J. Warmark, D. P. Allison, L. a. Bottomley, A. J. Lourenco and T. L. Ferrell, J. vac. Sci. Technol. A 10 (1992) 630.

[9] Kuo-Tong Chen, M. A. George, Y. Zhang, A. Burger, Ching-Hua Su, Yi-Gao Sha, D. C. Gillies and S. L. Lehoczky, *J. Cryst. Growth*, **147**, 292 (1995).

[10] Romesh C. Sharma and Austin Y. Chang, *J. Cryst Growth*, **88**, 193 (1988).

Biography

Swanson L. Davis, is a graduate student in Master degree of Physics/Chemistry joint program. He has been involves in crystal growth and characterization of compound semiconductors in NASA/Fisk Center. He is interest in solid state Physics, electronic materials and application. He plan to continue into Ph.D. program after graduated from Fisk.

Electronic and Vibrational Spectra of InP Quantum Dots Formed by Sequential Ion Implantation

C. Hall[1], R. Mu, Y. S. Tung, A.Ueda and D. O. Henderson,
NASA Center for Photonic Materials and Devices, Fisk University, Nashville TN 37208

C. W. White,
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

KEY WORDS: InP, Quantum Dot, Surface Phonon

ABSTRACT

We have performed sequential ion implantation of indium and phosphorus into silica combined with controlled thermal annealing to fabricate InP quantum dots in a dielectric host. Electronic and vibrational spectra were measured for the as-implanted and annealed samples. The annealed samples show a peak in the infrared spectra near 320 cm^{-1} which is attributed to a surface phonon mode and is in good agreement with the value calculated from Frölich's theory of surface phonon polaritons. The electronic spectra show the development of a band near 390 nm that is attributed to quantum confined InP.

INTRODUCTION

The study of quantum dots has drawn attention both in fundamental and applied science.[1-8] One of the most striking characteristics of these quantum dots is the quantum confinement effect when the photoinduced excitons are spatially confined. This effect will modify the bulk electronic band structure which leads to the alteration of the exciton properties of the quantum dots. Quantum confinement effects become important when particle size is comparable to the excitonic radius. The simple particle in a box concept can be readily applied to understand quantum confinement ofd the exciton.

There have been number of ways developed to fabricate quantum dots in various host materials and in solution phases, such as inorganic synthesis, sol-gel synthesis, impregnation into porous hosts, vapor deposition, and ion implantation. Although ion implantation can be considered as one of the most clean and flexible techniques to obtain the quantum dots in an isulating host, the fabrication of the quantum dots has been limited to single element systems, such as Ge, Si, P, etc. Only recently, we have shown that sequential ion implantation followed by post thermal annealing is a promising and effective technique to fabricate II - VI and III - V quantum dots in the dielectric hosts.
In this paper, we report our first experimental procedure and results on InP quantum dots fabricated in an SiO_2 host. Optical and vibrational spectra have provided unambiguous identification of InP quantum dots formation. Efforts were also made to estimate the average particle size and size distribution via TEM and model calculation.

EXPERIMENTAL

In ions (160 keV energy) and P ions (70 keV energy) were sequentially implanted into Dow Corning 7940 optical quality fused silica substrates. Implantation energies were selected to insure a maximum overlap of the profiles for the two implanted ions. The implantation experiment was conducted at room temperature. Four different concentrations of In and P ions were implanted with 1:1 ratio. The ion doses are $1x10^{16}$, $3x10^{16}$, $6x10^{16}$, and $10x10^{16}$ In (or P)/cm^2.

A step-wise (every 100 °C) thermal annealing process was conducted under 95% Ar and 5% H_2 environment. In addition, two other sets of substrates, In implanted only and P implanted only with same concentrations were also chosen and were annealed under an identical annealing condition. The starting temperature was 200 °C and a typical annealing time was 30 min. Upon each annealing, a set of UV-Vis and far infrared spectra were collected to examine the changes in the electronic and vibrational properties with respect to the previous sets of spectra. A lateral comparison was also made among In implanted, P implanted, and In + P sequentially implanted samples. Optical measurements were made possible with an Hitachi 3501 UV-Vis-NIR spectrophotometer in the

[1]Mr. Hall obtained his B.S. degree in chemistry from Cal. State University, Chico. Currently, he is a second year graduate student working in the Chemical Physics Group in the Department of Physics at Fisk University. He is expected to obtain a M.S. degree in chemistry with a physics minor. His interests are in nanophase materials with a focus on the linear optical properties of quantum dots. His thesis work has concentrated on the electronic and vibration properties of InP quantum dots.

range of 3200 - 185 nm in transmission mode. Far infrared spectra were obtained with Bomem DA3 spectrometer with resolution of 4 cm^{-1} and were normalized against the signal of a virgin silica substrate. Rutherford backscattering measurement were performed and showed maximum at ~0.05 μm and a Gaussian profile. However, In-alone implanted in silica revealed a bimodal distribution

RESULTS AND DISCUSSION

Fig 1 shows the UV-Vis spectra of as-implanted and annealed In + P substrates with concentrations of 1x, 3x, 6x and $10x10^{16}$ ions/cm^2, respectively. As the concentration increases, the absorption in UV region increases and tails into visible region.

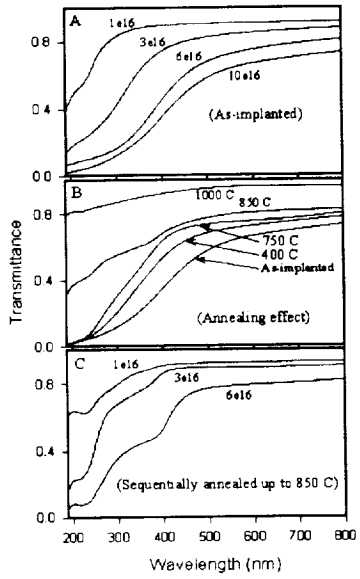


Fig. 1 Electronic spectra of In + P implanted in fused silica.

The figure demonstrates that the step-wise thermal annealing effects on UV-Vis spectra for In + P concentrations of 1, 3, $6x10^{16}$ ion/cm^2 and $10x10^{16}$ ions/cm^2 shows the development of a band between 300-400 nm. When the substrates were annealed at 700 - 850°C, an abrupt change in optical spectrum was observed with both In + P concentrations.This band appears at a wavelength much shorter than the bandgap of bulk InP (969 nm). Therefore we attribue this feature to quantum confined InP

Li *et al.*[9] have employed infrared techniques to investigate InP molecules in rare-gas matrices. Their results show two distinct bands at 257.9 and 249.3 cm^{-1}, respectively. The former is assigned to InP stretch vibration, while the latter is inferred to ν_3 vibration of In_2P species. It is also known that InP semiconductor has two modes in far ir region. The transverse optical mode (TO) is at 304 cm^{-1}, while the longitudinal optical mode (LO) is located at 346.5 cm^{-1}. The far-infrared

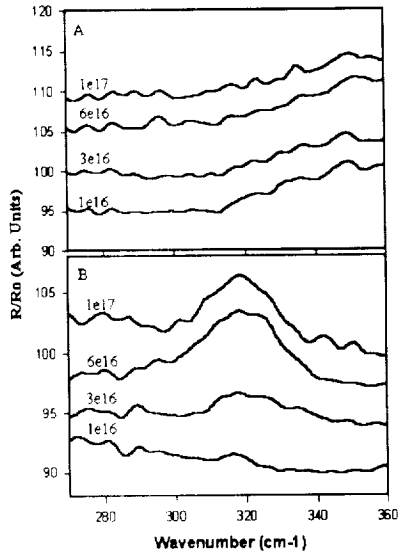


Fig. 2 Vibrational spectra of InP quantum dots formed in fused silica host.

measuremnts show the development of a band near 320 cm^{-1} in figure 2. This frequency falls between the LO and TO modes of the bulk material and therefore it is attributed to a surface phonon mode of InP. This assignment is supported by the fact that the calculated frequency Frölich's theory is within 1% of the observed

value.[10,11]

CONCLUSIONS

Sequential In and P ion implantation into silica and careful thermal annealing have proven to be a novel and effective way to fabricate III - V and II - VI semiconductor quantum dots in a dielectric host. The post thermal annealing can be the critical step to obtain desired quantum dots, size, and size distribution. The infrared reflectance technique is a simple but a very powerful means to identify nanocrystal formation through investigating its surface phonon properties. In the present research, we are able to obtain InP quantum dots in silica host. The particle size is estimated to be in 2 - 3 nm. Due to the bimodal distribution of the indium in silica, it is suggested that the synthesized InP quantum dots also have bimodal distribution based on optical spectra. In addition, because of the mismatch between the indium and phosphorus ion distribution profile, the two absorption peaks 216 nm and 239 nm can be mainly attributed to surface plasma absorption. Higher temperature annealing will result in the loss of both indium and InP materials.

ACKNOWLEDGEMENTS

The work at Fisk is supported by NASA under grant No. NAGW-2925. Research at ORNL is sponsored by the Division of Materials Science, U. S. Department of Energy, under contract DE-AC05-84OR21400 with Lockheed Martin Energy System, Inc.

REFERENCES

1. A. I. Ekimov, A. L. Efros, A. A. Onushchenko, Solid State Commun. **56**, 921 (1985).
2. L. E. Brus, J. Chem. Phys. **80**, 4403 (1984)
3. For example, Y. Wang and N. Nerron, J. Phys. Chem. **91**, 5005 (1987); O. I. Micic, Z. Li, G.Mills, J. C. Sullivan, and D. Meisel, J. Phys. Chem. **91**, 6221 (1987); C. J. Sandroff, S. P.Kelty, and D. M. Hwang, J. Chem. Phys. **85**, 5337 (1986).
4. O. I. Micic, C. J. Curtis, K. M. Jones, J. R. Sprague, and A. J. Nozik, J. Phys.Chem. **98**, 4966 (1994).
5. R. L. Wells, S. R. Aubuchon, S. S. Kher, and M. S. Lube, Chem. Mater. **7**, 793 (1995).
6. J. Warnock, D. D. Awschalom, Phys. Rev. B **32**, 5529 (1985); D. O. Henderson, R. Mu, A.Ueda, A. Burger, K. T. Chen, and D. O. Frazier, Mat. Res. Soc. Symp. Proc. **366**, 283(1995).
7. For example, M. Fujii, S. Hayashi, and K. Yamakoto, Appl. Phys. Lett **57**, 2692 (1990); Y.Sasaki and C. Horie, Phys. Rev. B **48**, 2009 (1993).
8. T. Shimizu-Iwayama, K. Fujita, S. Nakao, K. Saitoh, T. Fujita, and N. Itoh, J. Appl. Phys. **75**, 7779 (1994); H. Atwater, K. V. Shcheglov, S. S. Wong, K. J. Vahala, R. C. Flagan, M.L. Brongersma, and A. Polman, Mat. Res. Soc. Symp. Proc. **316**, 409 (1994).
9. S. Li, R. J. Van Zee, and W. Welner, Jr., J. Phys. Chem. **98**, 2275 (1994).
10. S. Hayashi, Japan. J. Appl. Phys. **23**, 665 (1984) and references therein.
11. R. Ruppin and R. Englman, Rep. Prog. Phys. **33**, 149 (1970); R. Ruppin, *Electromagnetic Surface Modes* edited by A. D. Boardman, Chap. 9, pp345 - 398 (John Wiley & Sons Ltd,1982).

PHOTOREFRACTIVITY IN A TITANIUM DOPED ZnCdTe CRYSTAL

M. Davis[1], L. Collins, K. Dyer, J. Tong, A. Ueda, H. Chen, K-T Chen, A. Burger, Z. Pan, and S.H. Morgan

Center for Photonic Materials and Devices, Fisk University, Nashville, TN 37208

Single crystals of $Zn_{.04}Cd_{.96}Te$ was grown by horizontal physical vapor transport (PVT) method and doped by annealing with $TiTe_2$ powder at 600 C for six days. Photorefractive two-beam coupling, along with photoluminescence and absorption spectroscopy, were used to characterize the ZnCdTe:Ti crystal. At 1.32 μm, the photorefractive gain has been measured as a function of the grating period. A gain of about 0.16 cm^{-1} was obtained at an intensity of about 0.1 W/cm^2. The results of this titanium doped ZnCdTe crystal are compared to that of vanadium-doped CdTe crystals reported previously.

Introduction

A strong motivation for investigation of photorefractive II-VI semiconductors has been the compatibility of their operating wavelength range with currently available semiconductor diode lasers. The 1.3 - 1.5 μm range is also important for fiber communication and eye-safe applications. As one looks for better infrared photorefractive semiconductors, CdTe stands out as having potential. The literature value of the electro-optic coefficient is three times larger than that of GaAs and InP. The comparison of the measured photorefractive sensitivity of CdTe:V to other semiconductors and oxides indicated that doped CdTe provides the most sensitive photorefractive response reported to date. [1-3]

The $Zn_xCd_{1-x}Te$ ternary system offers the capability that its band-gap can be tailored by adjusting the alloy composition, thus the crystals may be optimized for use at desired wavelengths.

Doping CdTe with 3d-transition metals like vanadium or titanium ensures the incorporation of additional deep levels, and allows for the investigation of the role of deep donors in CdTe, assumed to account for the high resistivity. A previous study suggested that Ti may be a more attractive candidate as dopant in CdTe because they found that the optical loss in Ti-doped crystals is smaller [1].

We present here the first results of photorefractive two-beam coupling measurements on a $Zn_{.04}Cd_{.96}Te$ single crystal doped with titanium. The measurements were performed at 1.32 μm using a diode pumped cw Nd:YAG laser. We also present the photoluminescence and absorption spectra, both before and after titanium doping. These results are discussed and compared to that of vanadium-doped CdTe crystals reported previously [1-3].

Theory

In photorefractive two-beam coupling experiment (Fig. 1), two beams of unequal intensity intersect inside the semi-insulating sample to form a spatially sinusoidal interference pattern. In the region of constructive interference the charge carriers are excited to the band states, undergo diffusion, drift, and are recaptured by traps. These effects result in charge redistribution and give rise to the space-charge field. This field, acting through the linear electro-optic effect, modulates the refractive index of the material. If the grating is displaced with respect to the incident optical interference pattern the beam coupling gives rise to energy transfer from one beam to the other. The grating wave-vector depends on the wave-vectors of

[1]Michelle R. Davis received the Bachelor of Science degree in Physics from Fisk University in 1993. Her academic achievements include The National Dean's List, Mortar Board College Senior Honor Society, Beta Kappa Chi National Scientific Honor Society and Who' Who Among Students in American Universities and Colleges. Currently she is pursuing the Master's Degree in the field of physics at Fisk University.

the incident beams and the angle between the two crossed beams, the value of grating wave-vector is

$$k_g = 2k\sin\theta$$

where 2θ is the angle between two crossed beams.

By chopping the pump beam and monitoring the signal beam, the photorefractive gain can be measured. The ratio of the modulation of the signal-beam ΔI_s to the signal beam I_s without pump beam is related to the photorefractive gain Γ. The relation is given by [4]

$$\frac{\Delta I_s}{I_s} = \frac{\exp(\Gamma L) - 1}{1 + \beta\exp(\Gamma L)}$$

where L is the overlap length between two beams and β is the intensity ratio of the signal beam to pump beam.

According to a simplified model given by Kukhtarev and others, the photorefractive gain Γ can be predicted by the formula [1]

$$\Gamma = = \frac{2\pi(n_0^3\gamma_{eff})}{\lambda\cos\theta}\left(\frac{KT}{e}\right)\frac{k_g}{1+(k_g^2/k_D^2)}\xi$$

where ξ is the electron-hole competition factor, k_D is the inverse of the Debye screening length, and γ_{eff} the effective electro-optic coefficient. For a one level model, $k_D^2 = (e^2/\epsilon KT)\,N_{eff}$, where N_{eff} is the effective trap density and ϵ is the dielectric constant. The γ_{eff} depends not only on electro-optic properties of the crystal, but also on the sample orientation and beam polarizations. This relation is

$$\gamma_{eff} = \hat{e}_1^*(Rk_g)\hat{e}_2$$

where R (3×3×3) is the linear electro-optic tensor, k_g (3×1) is the unit grating wave-vector, $\hat{e}_1^*$ (1×3) and $\hat{e}_2$ (3×1) are the unit polarization vectors of signal and pump beams, respectively. $\gamma_{eff} = \gamma_{41}$ for the experimental configuration used in this study.

Experimental

Single crystal $Zn_{.04}Cd_{.96}Te$ was grown by horizontal physical vapor transport (PVT) method. Approximately 4% Zn was added to improve crystallographic properties in reducing the dislocation density and subgrain boundaries. Titanium doping was accomplished by annealing with $TiTe_2$ powder at 600 C for six days. The sample was cut and polished to a 5×4×1.6 mm^3 cube with the orientation shown in Fig. 1. The resistivity was measured by applying silver-paint contacts on the 3 × 3 mm^2 area and measuring the I - V curve. The resistivity is in excess of $10^5\ \Omega$ cm at voltages up to 20 V, further increase of the voltage results in the decrease of the resistivity for the doped sample.

The absorption spectra measured at room temperature before and after titanium doping are shown in Fig. 2. The spectrum of the sample after doping shows a broad absorption shoulder to about 1.4 μm and a long tail extending over 1.8 μm. Because the light scattering increased significantly after doping, the absorption spectrum was measured by using an integration sphere in a Hitachi model U-4001 spectrophotometer. An absorption coefficient of 1.2 cm^{-1} was obtained at 1.32 μm after Ti doping. This value is similar to that of a CdTe:V sample with vanadium concentration 5 × 10^{16} /cm^3 [3].

Photoluminescence (PL) measurements were performed at 10 K. Fig. 3 shows the PL spectra both before and after Ti doping. A prominent, structured broad emission band from 1.4 to 1.5 eV appeared after Ti doping. The zero-phonon recombination occurs at 1.466 eV and is followed by a set of phonon replicas separated by the longitudinal optical-phonon energy of about 0.021 eV. A peak at 1.492 eV was also observed. These emissions are attributed to the donor-acceptor recombinations related to the doping defects. Similar emission peaks were observed in a vanadium-doped CdTe, but the exact nature of the defect centers responsible for these emissions is still unknown or is a matter of speculation [5].

The photorefractive properties were studied by two-beam coupling at 1.32 μm using a diode pumped cw Nd:YAG, Lightwave Model 126-1319-150. Two s-polarized Gaussian beams (beam size is about 1.5 mm in diameter) with a beam intensity

ratio of 10:1 interfered in the (110) plane, forming a grating along the (001) direction. The pump beam was modulated at 84 Hz by a chopper. The beam-coupling gain Γ reached a value of about 0.16 cm^{-1} at intensity levels above 100 mW/cm^2 (I_s) and a grating spacing around 2 μm. This gain coefficient is about 60% of the value reported for a CdTe:V sample at 1.32 μm with an intensity of 10 mW/cm^2 [3]. We attribute the high intensities required to observe photorefractive gain to the relatively low resistivity of our sample, because the photoconductivity must exceed the dark conductivity to observe the photorefractive gain. Fig. 4 shows the two-beam coupling gain coefficient measured at different grating period. The gain coefficient exhibits a large error bar as the data were obtained at different point of the samples. This implies the inhomogeneity of our sample. At present, we do not know whether the dominant photorefractive species in this sample are electrons or holes. We also performed the two-beam coupling experiment at 900 nm by using a cw Ti:sapphire laser, but no gain signal beyond noise can be obtained. The wavelength of 900 nm could be too close to the absorption edge of our crystal. The quality of the titanium doped crystal is of great concern. A recent study reported that titanium doping leads to lower resistivity than when using vanadium as a dopant, and the homogeneity in the resistivity of the CdTe:Ti crystal is not as good as that of the vanadium doped crystal. [6]

Conclusion

We have reported the photorefractivity at wavelength 1.32 μm in a $Zn_{.04}Cd_{.96}Te$:Ti crystal sample for the first time. This material may have potential as a fast and sensitive material for the near-infrared range. But for obtaining sizable crystals with a high crystal quality, controllable defect densities, and a high resistivity, there are material growth challenges to overcome.

This research is supported by NASA Grant NAGW-2925 and NSF Grant HRD-9550605.

References

[1] R. B. Bylsma, P. M. Bridenbaugh, D. H. Olson, and A. M. Glass, "Photorefractive properties of doped cadmium telluride", Appl. Phys. Lett. 51, 889 (1987).

[2] A. Partovi, J. Millerd, E. M. Garmire, M. Ziari, W. H. Steier, S. B. Trivedi, and M. B. Klein, "Photorefractivity at 1.5μm in CdTe:V", Appl. Phys. Lett. 57, 846 (1990).

[3] J.-Y. Moisan, N. Wolffer, O. Moine, P. Gravey, G. Martel, A. Aoudia, E. Repka, Y. Marfaing, and R. Triboulet, "Characterization of photorefractive CdTe:V: high two-wave mixing gain with an optimum low-frequency periodic external electric field", J. Opt. Soc. Am. B 11, 1655 (1994).

[4] P. Tayebati, J. Kumar, and S. Scott, "Photorefractive effect at 633 nm in semi-insulating cadium sulfide", Appl. Phys. Lett. 59, 3366 (1991).

[5] R. N. Schwartz, M Ziari, and S. Trivedi, "Electron paramagnetic resonance and an optical investigation of photorefractive vanadium-doped CdTe", Phys. Rev. B 49, 5274 (1994).

[6] R. Schwarz, W. Joerger, C. Eiche, M. Fiederle, and K. W. Benz, "Closed tube vapour growth of CdTe:V and CdTe:Ti and its characterization", J. Crystal Growth 146, 92 (1995).

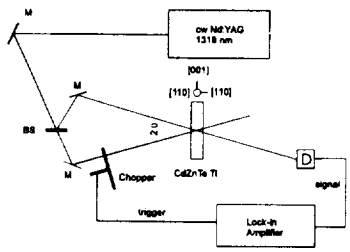


Figure 1: Photorefractive two-beam coupling experimental setup.

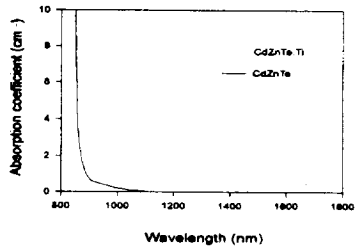


Figure 2: Absorption spectra at room temperature. An absorption coefficient of 1.20 cm^{-1} was observed at 1.32 μm for the Ti-doped sample.

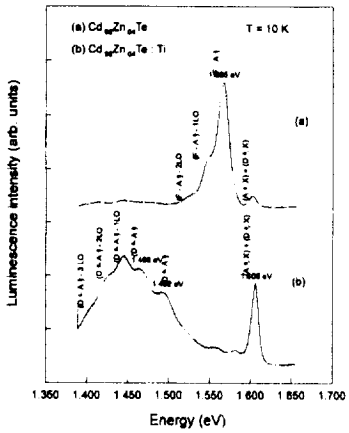


Figure 3: PL spectra at 10 K, showing (D^0 - A^0) band from 1.4 to 1.5 eV for the Ti-doped sample.

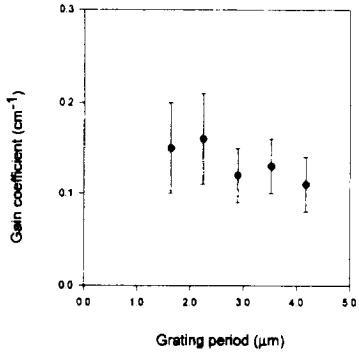


Figure 4: The two-beam coupling gain coefficient as a function of grating period at 1.32 μm.

RAMAN AND PHOTOLUMINESCENCE SPECTROSCOPY OF Er^{3+} DOPED HEAVY METAL OXIDE GLASSES

Keith Dyer[1], Zhengda Pan, and Steve Morgan
Center for Photonic Materials and Devices, Fisk University, Nashville, TN 37208

The potential applications of rare-earth ion doped materials include fiber lasers which can be pumped conveniently by infrared semiconductor laser diodes. The host material systems most widely studied are fluoride crystals and glasses because fluorides have low nonradiative relaxation rates due to their lower phonon energies. However, the mechanical strength, chemical durability and temperature stability of the oxide glasses are generally much better than fluoride glasses. The objective of this research was to investigate the optical and spectroscopic properties of Er^{3+}-doped lead-germanate and lead-tellurium-germanate glasses. The maximum vibrational energy of lead-tellurium-germanate glasses are in the range of 740 -820 cm^{-1}, intermediate between those of silicate (1150 cm^{-1}) and fluoride (530 cm^{-1}) glasses.

Introduction

During recent years, there has been a great deal of interest in optical-frequency upconversion in rare-earth doped solids. One of the rare-earth doped transparent solids of main interest is glass because glass can be fabricated as optical fibers to confine the pumping light from the diode laser in order to maintain a high intensity over a long interaction length. In order to minimize the non-radiative losses, it is best to select glasses in which the maximum phonon (vibrational) energy is small. Heavy metal oxide glasses are good candidates for being the host glass system. A previous study has reported Tm doped lead-germanate glasses to be promising optical fiber materials [1]. In this study lead-germanate and lead-tellurium-germanate glasses doped with erbium and containing modifiers CaO, BaO, KNO_3, and ZnO have been chosen for analysis because of their probable excellent upconversion properties and also because they have good fiber forming properties.

For each of these glasses, measurements were made of the density, index of refraction, Raman spectra, UV-visible absorption spectra, photoluminescence, and upconversion luminescence. The objectives of these measurements are, (1) to determine the mechanism of radiative and non-radiative transitions of Er^{3+} in these glasses, and (2) to investigate the relationship between the efficiency of upconversion luminescence and the structure and composition of the glasses.

Experimental

Sample Preparation

The glasses were prepared by mixing various concentrations of PbO, GeO_2, $CaCO_3$, Er_2O_3, PbF_2 and TeO_2 in 30 gram batches. These concentrations can be seen in Table 1. The GeO_2 was 99.9999% purity (Eagle Pitcher Research Lab), Er_2O_3 was 99.99% purity (Johnson Matthey Rare Earth Products), PbF_2 was 99 + % purity (Aldrich Chemical Company, Inc.). $CaCO_3$ (Mallinckrodt Chemical Works), KNO_3 (Johnson Matthey Electronics), PbO (Johnson Matthey Catalog Company), TeO_2 (Alfa Products Thiokol Ventron Division), and ZnO (Fisher Scientific Company) were all reagent grade.

The mixtures were melted in fused silica crucibles. The temperature and duration of the melting of the glass mixture are as follows: The GPBZKE glass was melted at 1100°C for 25 minutes and annealed at 450°C, the GPFE glass was melted at 1200°C for 10 minutes and annealed at 450°C, and the GPTCE was melted at 900°C for 25 minutes and annealed at 280°C. The glass melts were poured onto a copper plate and pressed from the top with another

[1]Keith Dyer, a native of Nashville Tennessee, attended Hume-Fogg Magnet and Fisk Unversity, both in his home town. During his scholastic career Keith received numerous academic honors including the 1993-1994 & 1994-1995 National Dean's List, Beta Kappa Chi National Scientific Honor Society, Faith Unlimited, S. H. Simpson, Shoney's, and the Eskind Scholarships, etc. He completed both his undergraduate and graduate work at Fisk University in the field of physics. He received the bachelor's degree in 1993 and the master's degree in 1995. He is currently pursuing a career in research and development in industry.

copper plate, forming glass plates of about 1.0 mm in thickness. The glasses were annealed for one hour to relieve strain and allowed to cool to room temperature in the furnace.

The glasses were cut into 10 mm by 15 mm rectangles using a diamond saw, smoothed on an Ecomet polisher-grinder using 600 grit Buehler silicon carbide grinding paper, and polished on an Ecomet polisher-grinder using Buehler micropolish 0.1 micron alpha alumina.

Raman

The Raman spectra of the lead-germanate and lead-tellurium-germanate glasses were excited at 457.9 nm [Fig. 1]. The Raman peaks in the low frequency region between 25 cm^{-1} and 150 cm^{-1} are related to the lead ion vibrational modes. These peaks show up in the low frequency region because larger masses such as lead ions cause low frequency vibrational modes. The Raman peaks in the mid frequency region between 400 cm^{-1} and 500 cm^{-1} are due to deformation vibration modes which could be the combination of the stretching, bending, and rocking vibrations of the network structure of the glasses. The Raman peaks in the high frequency region between 740 cm^{-1} and 840 cm^{-1} are due to the stretching vibrational modes of the glass network structures. These high frequency vibrational modes in the Raman spectra are the most important factors in the coupling to the electronic transitions of the rare-earth ions.

The highest frequency modes in the GPTCE occur at a lower frequency than the highest frequency modes in the GPBZKE and GPFE. This is because the glass network of the GPTCE glass is formed by the combination of Ge-O bonds and Te-O bonds, unlike the other two glasses whose networks are formed by only Ge-O bond. Ge-O-Te has a weaker bond force constant than the symmetric Ge-O-Ge bond. Te is more positive than Ge, which may cause a weaker Ge-O bond in a non-symmetric Te-O-Ge bridge. Weaker bonds have lower frequency modes.

Upconversion Efficiency

The upconversion efficiency is defined as [2]

$$\eta = \frac{I_{emit}}{I_{abs}} \qquad (1)$$

where I_{emit} is the intensity of the emitted light and I_{abs} is the intensity of the absorbed incident light. For a two-step upconversion process, the visible output intensity is proportional to the square of the infrared excitation intensity; i.e.,

$$I_{vis} \propto I_{ir}^2 \qquad (2)$$

The upconversion efficiency defined in (1) therefore depends on the IR pump intensity. Therefore, in order to compare the upconversion efficiency of different materials meaningfully, we have to specify the pump intensity used.

The Experimental Technique

The absorbed infrared light intensity I_{abs} can be calculated using

$$I_{abs} = I_{incident}(1 - \exp(-\alpha t)) \qquad (3)$$

where α and t are the absorption coefficient and sample thickness, respectively. The values of $I_{incident}$, α, and t can be measured experimentally. It is very difficult to measure the I_{emit} directly.The emitted light intensity, however, can be determined by comparing the upconversion luminescence signal with the directly excited luminescence signal intensity upon knowing the direct excitation luminescence quantum yield η_q that is defined as [2]

$$\eta_q = \frac{emitted\ light\ intensity}{absorbed\ radiation\ intensity} = \frac{\tau_{exp}}{\tau_R} \qquad (4)$$

where τ_{exp} is the experimentally determined lifetime from the level in question and τ_R the radiative lifetime. A $\eta_q = 0.82$ has been reported [2] for the $^4S_{3/2} \rightarrow {}^4I_{15/2}$ transition of Er^{3+} in a barium-thorium fluoride glass.

Combining (1) and (4), the formula used to determine the upconversion efficiency η by comparing the upconversion luminescence signal with directly excited signal intensity is:

$$\eta = \eta_q \left(\frac{I_{abs}(vis)}{I_{abs}(IR)}\right)\left(\frac{I_{emit}(upconverted)}{I_{emit}(direct)}\right) \qquad (5)$$

In the above formula, both I_{abs}(vis) and I_{abs}(IR) can be calculated upon measuring the incident intensity, absorption coefficient, and the sample thickness. The luminescence intensities I_{emit}(upconverted) and

I_{emit}(direct) are only the collected portion of the emitted light. The collected luminescence intensities by direct excitation and by IR excitation therefore have to be measured at the same light collecting conditions. The values of upconversion efficiency were calculated using formula (5). See Table 2.

Upconversion Luminescence

In the upconversion photoluminescence study conducted, the glass samples were excited using the wavelength 797 nm (12,547 cm^{-1}). The intense green emission between 514 and 573 nm accounts for 80% of the total emitted light, and is easily visible to the naked eye [Fig. 2]. There is also a weak red emission at 622 nm, a near-infrared band at 885 nm and a weak blue emission at 410 nm.

The 527 nm and 550 nm bright green emission is a result of the $^4S_{3/2} + {}^2H_{11/2} \rightarrow {}^4I_{15/2}$ transitions [Fig. 3]. The absorption of 797 nm photons raise Er^{+3} ions from the $^4I_{15/2}$ ground level to the $^4I_{9/2}$ level, which undergoes multiphonon relaxation to the lower level of $^4I_{11/2}$. The ions in the $^4I_{11/2}$ level sequentially absorb 797 nm photons and are raised to the $^4F_{3/2,5/2}$ levels. The ions in the $^4F_{3/2,5/2}$ levels undergo multiphonon relaxation through $^4F_{7/2}$ to $^2H_{11/2}$ and to the $^4S_{3/2}$ level. Because the energy gap below the $^4S_{3/2}$ level is larger, the nonradiative transition through multiphonon relaxation from $^4S_{3/2}$ becomes smaller. Therefore, the bright green flourescence was emitted through $^4S_{3/2} + {}^2H_{11/2} \rightarrow {}^4I_{15/2}$ transitions.

The 855 nm near infrared emission is a result of the $^4S_{3/2} \rightarrow {}^4I_{13/2}$ transition. The absorption of 797 nm photons raise Er^{+3} ions from the $^4I_{15/2}$ ground level to the $^4I_{9/2}$ level, which undergoes multiphonon relaxation to the lower level of $^4I_{11/2}$. The ions in the $^4I_{11/2}$ level sequentially absorb 797 nm photons and are raised to the $^4F_{3/2,5/2}$ levels. The ions in the $^4F_{3/2,5/2}$ levels undergo multiphonon relaxation through $^4F_{7/2}$ to $^2H_{11/2}$ and to the $^4S_{3/2}$ level. The ions then go through a radiative transition from the $^4S_{3/2}$ level to the $^4I_{13/2}$ level, causing the 855 nm near infrared emission.

The 662 nm red emission is a result of the $^4F_{9/2} \rightarrow {}^4I_{15/2}$ transition. The $^4F_{9/2}$ level can be populated through two paths: (1) from the $^4S_{3/2}$ level through multiphonon relaxation; (2) via the energy transfer between excited ions owing to their Coulomb interaction. The two coupled transitions are identified as $^4S_{3/2} \rightarrow {}^4F_{9/2}$ and $^4I_{9/2} \rightarrow {}^4F_{9/2}$. The energy gaps between the two coupled transitions are matched well with a small red shift, which is favorable to this kind of transfer. The Er^{+3} ions then go through a radiative transition from the $^4F_{9/2}$ level to the $^4I_{15/2}$ level, resulting in the 662 nm red emission.

The 410 nm blue emission is a result of the $^2H_{9/2} \rightarrow {}^4I_{15/2}$ transitions. Because a a weak emission at 382 nm was observed, a third step excitation is responsible for the 410 nm emission. The mechanism for this process has been suggested to be cooperative upconversion [3].

Conclusion

Density, index of refraction, Raman spectra, UV-visible absorption spectra, normal photoluminescence spectra, and upconversion luminescence spectra were made on the GPBZKE, GPFE, and the GPTCE glass samples. These measurements allowed us to successfully determine the mechanisms of radiative and nonradiative transitions of Er^{+3} in these glasses. The measurements also allowed us to investigate the relationship between the efficiency of upconversion luminescence and the structure and composition of the glasses. The effect of tellurium on the glass structure was to lower the vibrational energy of the highest frequency photon, which increased the efficiency of the upconversion luminescence. The upconversion efficiency of Er^{+3} ions in GPTC glass is about twice that in GPBZK glass.

The green upconversion luminescence at 527 nm and 550nm resulted from the $^4S_{3/2} + {}^2H_{11/2} \rightarrow {}^4I_{15/2}$ transition, following the sequential absorption of two 797 nm photons. The near infrared emission at 855 nm, the red emission at 662 nm, and the weak blue emission at 410 nm are identified as the $^4S_{3/2} \rightarrow {}^4I_{13/2}$, $^4F_{9/2} \rightarrow {}^4I_{15/2}$, and the $^2H_{9/2} \rightarrow {}^4I_{15/2}$ transitions respectively.

Acknowledgment:

This research was supported by NASA through Grant NAGW-2925.

References:

1) J. Wang, J.R. Lincolon, W.S. Brocklesby, R.S. Deol, C.J. Mackechnie, A. Pearson, A.C. Tropper, D.C. Hanna, and D.N. Payne, J. Appl. Phys. 73, 8066 (1993).

2) D.C. Yeh, W.A. Sibly, M.J. Suscavage, and M.G. Drexhage, J. Appl. Phys 62, 266 (1987).

3) Z. Pan, S. H. Morgan, A. Loper, V. King, B.H. Long, & W.E. Collins, J. Appl. Phys. 77, 1-5 (1995).

Table 1. Mol% Composition of Glass Samples (all containing 1% Er_2O_3)								
Sample	GeO_2	PbO	PbF_2	$BaCO_3$	ZnO	KNO_3	TeO_2	$CaCO_3$
GPBZKE	55	20	---	10	10	5	----	----
GPFE	73	20	7	----	----	----	----	----
GPTCE	30	30	----	----	----	----	30	10

Table 2. Values of the upconversion efficiency η (focused).

Sample	I_{abs}(vis)/I_{abs}(IR)	I_{emit}(up)/I_{emit}(direct)	η
GPBZKE	0.203	3.54	0.156
GPFE	0.177	3.82	0.182
GPTCE	0.129	5.11	0.330

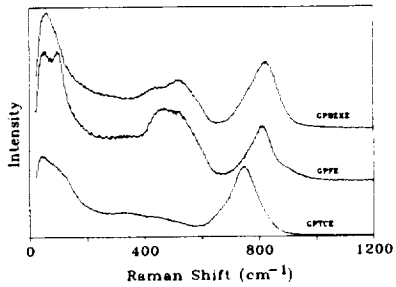


Figure 1: Raman spectra of GPBZKE, GPFE and GPTCE glasses

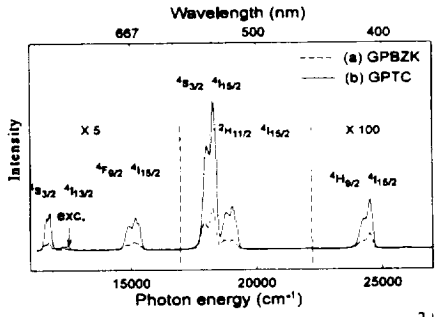


Figure 2: Upconversion luminescence spectra of Er^{3+} in (a) GPBZK glass and (b) GPTC glasss, excited by a 797 nm laser beam.

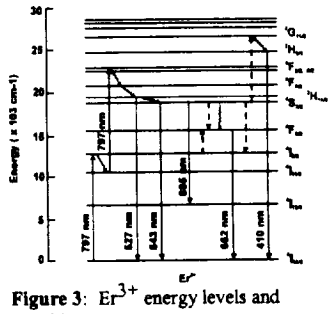


Figure 3: Er^{3+} energy levels and transitions.

Crystal Growth and Characterization of BiI_3

Jalila Hayes, Kuo-Tong Chen and Arnold Burger
NASA/Fisk Center for Photonic Materials and Devices
Department of Physics
Fisk University
Nashville, Tennessee 37208

Abstract

Bismuth tri-iodide (BiI_3) have been grown by physical vapor transport (PVT), and by the Bridgman (melt) method. These crystals along with pure and stoichiometric BiI_3 powder have been investigated by differential scanning calorimetry (DSC). The DSC results shows that pure BiI_3 powder has no phase transition and melts around 408 °C. While we found no evidence for the high temperature dissociation of BiI_3, the DSC measurements show that crystals grown from melt method contain a significant large amount of Bi-rich phases than crystals grown from PVT method, as indicated by phase transition detected at 270, 285, 298 and 336 °C.

Keyword: Bismuth tri-iodide, Differential Scanning Calorimetry, Physical Vapor Transport, Bi-rich Second Phase, Precipitate.

Introduction

Bismuth tri-iodide, BiI_3, is a wide bandgap, layered semiconductor with a large atomic number and a previously reported bandgap energy ranging from 1.73 [1] to 2.18 eV [2]. At room temperature, BiI_3 exhibits a hexagonal [3] structure, black color, large anisotropy, and week van der Waals bonding force between layers normal to the [0001] axis. Compared to the mercury iodide (HgI_2) which is a similar material to BiI_3 and has been developed as room temperature nuclear radiation detector [4], BiI_3 has no phase transition from room temperature up to its melting point [5,6], 408 °C, which attracted interests for its potential usage in the similar applications as HgI_2 with better structural properties. Bulk and single crystal platelet of BiI_3 have been grown by sublimation and recrystallization [7-10]. Crystal platelets of about 10 mm diameter and 1 mm thickness have been cleaved for detector fabrication and tested as room temperature x-ray and gamma ray detector. Although some crystal showed reasonable low leakage current, there were no (or weak) response to radiation. Conclusions from previous results suggest that further purification of the starting materials and optimization of crystal growth parameters are necessary before a evaluation as radiation detector can be made.

Experiments

Physical vapor transport (PVT), which is a vapor phase growth method, and the Bridgman (melt) growth methods have been used for crystal growth of BiI_3. The growth ampoule were first cleaned with acid (solution of 25% Nitric acid + 75% Hydrychroric acid), and rinsed with distilled water. The ampoules were then connected to a vacuum ($< 10^6$ Torr) and at the same time baked in a furnace at 600 °C. BiI_3 starting materials were loaded into ampoules and sealed under vacuum. For PVT growth, the source temperature was about 330 °C with a $\Delta T = 10$ °C, and for melt growth the source temperature was around 410 °C with the same ΔT. A bulk crystal with several small platelets of about 5 mm in diameter and 1 mm thickness were grown by PVT method, and a bulk crystal was grown by melt method. The crystals were investigated by differential scanning calorimetry (DSC), DSC-10 equipped with thermo-analysis program for thermogram analysis. Stainless steel DSC

capsules, 10 °C/min heating rate and temperature raging between 100 and 440 °C were used. The DSC measurements were repeated four times for every samples to ensure that any transition observed from thermograms are accurate.

Results and Discussion

For the pure and stoichiometry BiI_3 starting materials, the DSC thermograms shows no second phase transition until its melting temperature, 408 °C which agree with literature data [5,6]. Figure 1 and 2 shows the thermograms of the vapor and melt grown crystals, respectively. No peaks were observed below 200 °C, and therefore only thermograms between 200 and 440 °C have been shown.

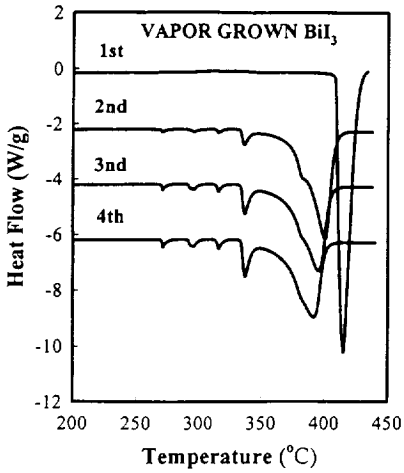


Figure 1. DSC thermograms of PVT (vapor) grown BiI_3 crystals.

The first heating measurement of the vapor grown crystal shows only a distinct melting peak at the melting temperature. On the subsequent heating several second phase transitions occur due to precipitation of Bi-rich phase outside the solid solution and the melting temperature shows a depression to lower temperature.

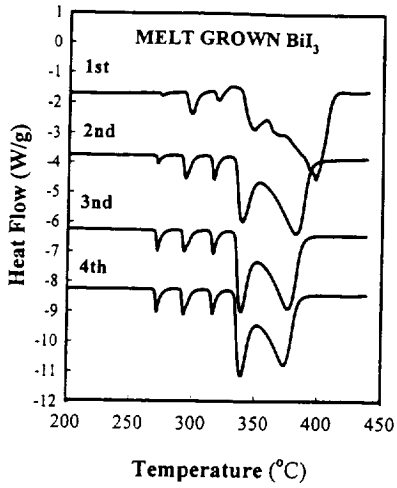


Figure 2. DSC thermograms of Bridgman (melt) grown BiI_3 crystals.

Several second phase transition were observed for the melt grown crystal even from the first heating measurement. These second phases have continued to develop during the subsequent heating measurements.

Table 1. Bi-rich second phase transition and melting temperatures (°C) from phase diagram data [5,6] and our DSC measurements.

Bi-rich Transition & Melting	269	285, 295	300	330, 336	408
Iodine-rich Transition	113				
Melt Growth BiI_3					
1st	272	293	316	339	397
2nd	271	292	315	335	359
3nd	270	291	314	334	358
4th	270	291	314	334	357
Vapor Growth BiI_3					
1st	-	-	-	-	409
2nd	269	293	314	334	400
3nd	270	291	314	334	395
4th	270	291	314	334	391

Table 1 listed the transition temperatures of Bi-rich, Iodine-rich second phases coexisting with BiI_3 solid solution from phase diagram data [5,6] and the transition temperatures obtained from DSC measurements in this study. There is only one Iodine-rich second phase exists above room temperature and the transition occurred at 113 °C. In all DSC measurements, there never been found any peak around this temperature, and this result indicate that there is no Iodine-rich phase in our materials. The observed transition temperatures at around 270, 291, 314 and 334 °C have been identified from phase diagrams [5,6] as Bi-rich second phases coexists with BiI_3 solid solution. For the vapor grown crystal, the stoichiometric composition might be laid within the homogeneity range at the growth temperature, but was close to the Bi-rich than purely stoichiometry, evidenced by the first heating measurement with no observation of any second phase transition. During the subsequent heating the crystal composition reached and followed the solidus line due to precipitation of Bi-rich second phases, and the melting has been depressed to lower temperature than stoichiometry melting. For the melt grown crystals, it is obvious that the crystals had a Bi-rich in composition during crystal growth process as evident by the observation of Bi-rich transitions on all heating measurements.

Summary

Bismuth tri-iodide (BiI_3) crystals have been grown by physical vapor transport and melt methods. DSC technique was employed to investigate the stoichiometry of the as-grown crystals, and the effects due to different growth methods were reported. The DSC results indicate that a crystal grown by the PVT method has a near stoichiometric composition while crystals grown from the melt contained a higher concentration of Bi-rich second phases.

Acknowledgments

Funding for this research have been provided by NASA through NASA/Fisk Center for Photonic Materials and Devices.

References

[1] V. I. Vashchenko and V. B. Timofeev, *Sov. Phys. Solid State*, **9**, 1242 (1967).

[2] V. M. Shmandii, V. K. Miloslavskii and V. V. Mussil, *Sov. Phys. Solid State*, **21(8)**, 1386 (1979).

[3] *"Powder Diffraction File"*, published by JCPDS, 7-269 (1993).

[4] V. Gerrish, *"Characterization and Quantification of Detector Performance"*, Chapter 14 in *"Semiconductors for Room Temperature Nuclear Detector Applications"*, Semiconductors and Semimetals, Academic Press., **43**, Ed. By T. E. Schlessinger and R. B. James (1995).

[5] S. J. Yosim, L. D. Ransom, R. A. Sallach and L. E. Topol, *J. Phys. Chem.*, **66**, 28 (1962).

[6] *"Binary Alloy Phase Diagram"*, 2nd Edition, **3**, Ed.-in-chief, T. B. Massalski, ATM, Metal Park, Ohio (1992).

[7] B. J. Curtis and H. R. Brunner, *Mat. Res. Bull.*, **9**, 715 (1974).

[8] T. Komatsu and Y. Kaifu, *J. Of the Phys. Soc. of Japan*, **40**, 1062 (1976).

[9] S. Takeyama, K. Watanabe and T. Komatsu, *Japan. J. of Appl. Phys.*, **29**, 710 (1990).

[10] M. Schieber, T. J. Davies, M. F. Schnepple, P. T. Randtke and R. C. Carlston, *J. of Apply. Phys.*, **45**, 5371 (1974).

Biography

Jalila Hayes is a senior undergraduate student in Physics/Chemistry program. She has been involves in crytal growth room temperature radiation detector materials and thermodynamic characterization research projects in NASA/Fisk Center for Photonic materials and Devices. She also expresses

interest in advanced degree and medical school.

RADIATION EFFECTS ON DC-DC CONVERTERS USING A NOVEL SPICE MODEL OF POWER MOSFETS

Charles Lawson and John Okyere Attia
Center for Applied Radiation Research
Prairie View A&M University
Prairie View, Texas 77446

ABSTRACT

This paper discusses the development of a spice model that incorporates the effect of total ionizing dose radiation on power MOSFET. The model is constructed using an empirical data taken from an irradiated IRF-150 power MOSFET. The comparison between simulated and empirical data confirms the accuracy of the model. The model is used to analyze the radiation effects on power mosfet dc-dc converter.

KEYWORDS:

Spice, dc-dc converters, radiation, mobility, threshold voltage.

BIOGRAPHY

Mr. Charles Lawson completed his B.S. and M.S. degrees in Electrical Engineering at Prairie View A&M University. He is currently a doctoral candidate at Texas A&M University. His research interests are software and digital hardware design.

1.0 INTRODUCTION

It is normally required that spaceborne systems be hardened against the effects of radiation. To build into systems this extra resilience to errors, it is essential to have accurate models on which to base an initial prototype system design. SPICE is a very powerful tool for circuit development and analysis of integrated circuits [1]. However, the semiconductor device models incorporated into SPICE were developed primarily for low-power devices and, hence, do not accurately simulate all modes of power MOSFET operation. This is an obstacle for any designer building power electronic circuits.

The built-in low power MOSFET models presently available in circuit simulation program, such as SPICE, are not adequate to simulate a power mosfet device for several reasons: (i) The complex variation of capacitance with bias conditions (ii) The presence of a cascade JFET, which affects both static and dynamic operation and (iii) The presence of a parasitic body diode, which has a dramatic impact on the third quadrant operation of the device. In this work, we intend to (i) model power mosfets under the influence of total ionizing dose radiation using SPICE circuit simulation tool [2] and (ii) use the model to analyze radiation effects on dc-dc converters.

2.0 RADIATION EFFECTS SUBCIRCUIT

The Harris model [2] was used as a basis for our development of radiation model of power MOSFET transistors. The approach taken with the Harris model is to use a subcircuit representation of a DMOS power transistor. This is done by adding external components to the built-in SPICE MOSFET models. In this way, a subcircuit with terminal characteristics representing those of a power MOSFET is developed.

See figure 1 for the cross-section of DMOS power transistor.

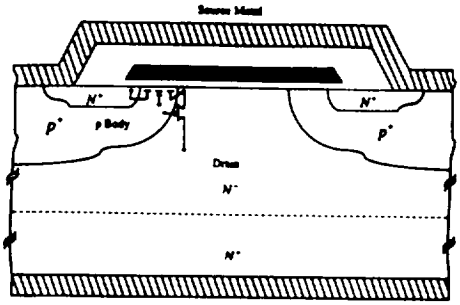


Fig. 1 Cross Section of DMOS Transistor

The electrical characteristics of the DMOS structure is understood in terms of a vertical JFET driven in cascade from a lateral low voltage MOSFET. Figure 2 shows a subcircuit of DMOS transistor developed by Harris [2].

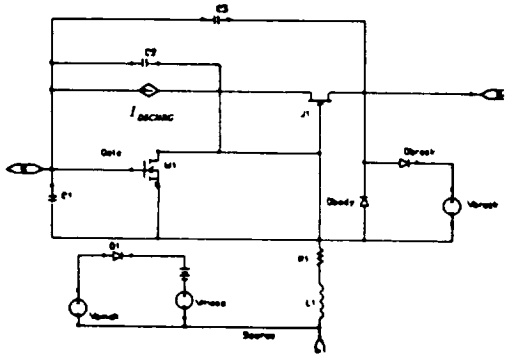


Figure 2 Pre-radiation model of DMOS transistor

The element values in our new model are based on empirical data taken from an irradiated IRF-150 power MOSFET transistor. The IRF-150 was irradiated at two dose rates at varying total doses. The data was recorded at dose rates of .36 mrads (Si)/s and .6rad(Si)/s [3]. Figure 3 shows the effects of the dose rate and total dose radiation on the threshold voltage. As it can be seen, the threshold voltage shift is negative, and increases in magnitude with each increase in both dose rate and total rate.

The threshold voltage shift equation for dose rate of 0.60(Si)/s over the range of total dose values 0-12 krad(Si) was calculated and was found to approximate:

$$V_{shift}(x) = -0.0026528x^2 - 0.0138123x - 0.016117559$$

where x is the total dose at dose rate of 0.6 rad/s.

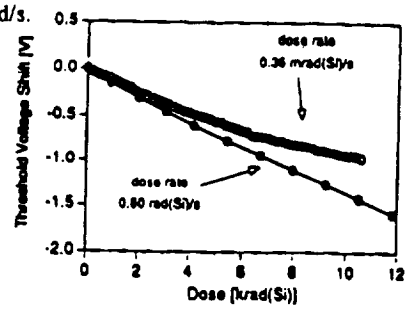


Fig. 3 Threshold Voltage vs. Dose

The corresponding function for the mobility degradation effects, shown in Figure 4, is:

$$\mu_{Norm}(x) = -0.0018615262x^2 - 0.05329208x + 0.98989519$$

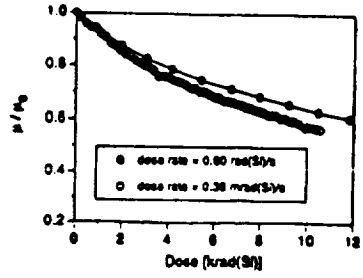


Fig. 4 Mobility vs. Dose

Threshold voltage shifts due to radiation is implemented by placing directly into the subcircuit a dependent voltage source whose value is dependent on the dose rate and total dose radiation.

The reduction of mobility with total ionizing dose results in reduction of device current. To account for the current reduction, it is necessary to scale the amount of current entering the drain by an amount dependent on the various voltages across the source, gate and drain, along with the dose rate and total dose. The mobility degradation is modeled by including in the subcircuit a current probe in series with the lateral MOSFET and a dependant source in parallel with the series combination of the lateral MOSFET and current probe. The dependent current source is scaled up to take into account the reduced mobility of power Mosfets with the total dose radiation.
The subcircuit implementation of the power mosfet under radiation environment is shown Figure 5.

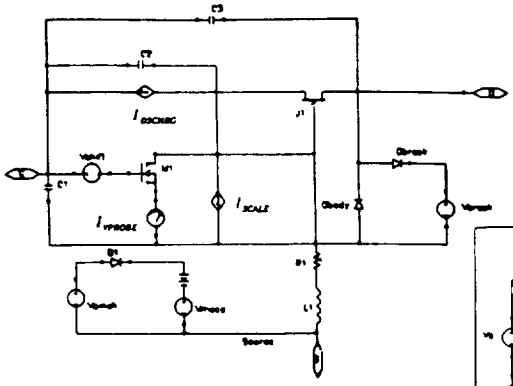


Fig. 5 Radiation Effects Subcircuit for DMOS

A simple test circuit was constructed to test the static I-V characteristics of the model. The subcircuit model I-V characteristics were tested by comparing a direct change in parametric values for the low power MOSFET, which is responsible for the static characteristics of the DMOS device, against the radiation effect model. The test circuit has two branches to provide the same test input to two different transistor models. It was found that the results obtained from our model are exactly the same as those obtained by changing the threshold voltage and mobility parameter, in the schematic capture program [4].

3.0 TID EFFECTS ON DC-DC CONVERTER

DC-DC converters change a fixed voltage source into a variable dc source. The converters are capable of transforming a dc voltage in a step-up or step-down manner. The converters can be used for motor control. They provide smooth acceleration control, high efficiency and fast dynamic response. An example of a dc-dc converter is shown in figure 6.

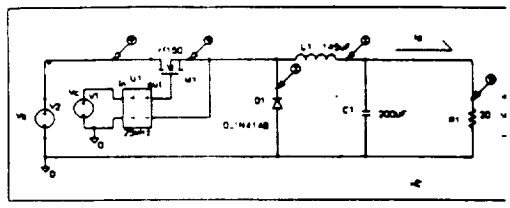


Fig. 5 Buck Regulator Circuit

The spice model of power mosfet, described

in the previous section, was used to obtain the effect of total ionizing dose radiation on buck regulator circuit. The output voltage and the currents through the diode and inductor were obtained before and after radiation. Figures 7 and 8 shows the diode and inductor current before and after radiation.

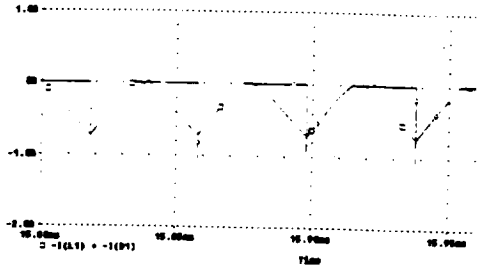


Fig. 7 Diode and Inductor currents before radiation

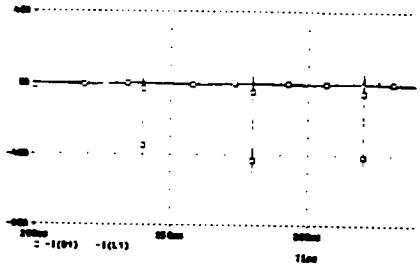


Fig. 8 Diode and Inductor currents after radiation

Figure 7 and 8 indicate that there is a large increase in current spikes due to ionizing radiation.

4.0 CONCLUSIONS

The radiation effect model of DMOS, developed in this work, is a subcircuit for use in the popular circuit simulator SPICE. Threshold voltage shifts due to radiation is implemented by placing directly into the subcircuit a dependent voltage source whose value is dependent on the dose rate and total dose radiation. The mobility degradation is modeled by including in the subcircuit a current probe in series with the lateral MOSFET and a dependant source in parallel with the series combination of the lateral MOSFET and current probe. The dependent current source is scaled up to take into account the reduced mobility of power Mosfets with the total dose radiation. The comparison between simulated and empirical data confirms the accuracy of the model. Simulations of dc-dc converter in a radiation environment indicates an increase in current spikes.

5.0 REFERENCES

1. A. Vlamirescu, "The Spice Book" John Wiley & Sons, Inc, 1994

2. C.F. Wheatley, Jr., H.R. Onan, G.M. Dolny, " Spicing-Up SPICE II Software for Power MOSFET Modeling" Application Note, Harris Semiconductor, No.861.1, May 1992.

3. R.D. Schrimpf, F.J. Wahle, R.C. Andrew, D.B. Cooper and R.F. Galloway, "Dose-Rate Effects on Total-Dose Threshold-Voltage Shift of Power Mosfets" IEEE Trans. Nucl. Sci., Vol 35, No. 6. pg. 1536 -1540.

4. C. Lawson, "Modeling of Total Dose Ionizing Radiation Effects on Power MOSFETs Using SPICE Circuit Simulation", M.S. Thesis, Prairie View A&M University, May 1995.

PROCESS AND RADIATION INDUCED DEFECTS IN ELECTRONIC MATERIALS AND DEVICES

Kenneth Washington and T.N. Fogarty
Center for Applied Radiation Research
Prairie View A&M University
Prairie View, Texas 77446

ABSTRACT

Process and radiation induced defects are characterized by a variety of electrical techniques, including capacitance-voltage measurements and charge pumping. Separation of defect type into stacking faults, displacement damage, oxide traps, interface states, etc. and their related causes are discussed. The defects are then related to effects on device parameters. Silicon MOS technology is emphasized. Several reviews of radiation effects and silicon processing exist [1,2,3,4,5].

PROCESS ENVIRONMENT

As semiconductor IC manufacture [4] now has feature sizes at the submicron level, process induced radiation defects are a yield limiting factor in VLSI & ULSI intergrated circuits. Sources of radiation in the process environment include electron beam evaporation, sputtering deposition, plasma and reactive ion etching, direct writing electron beam lithography and X-ray photolithography.

Because of the shrinking of device dimensions, defects introduced during crystal growth are more critical. In Czochralski (CZ) crystal growth, magnetic fields surrounding the crucible suppress thermal convection in the melt. Therefore, the resulting smaller temperature fluctuations reduce microscopic variation in resistivity. Also, transport of oxygen from the crucible is reduced thus reducing oxygen induced stacking faults and precipitates. Thermal donors from oxygen precipitates can alter carrier concentration thus limiting threshold voltage control in MOS

Kenneth Washington is a second year masters degree candidate. His field of study is the design and fabrication of integrated circuits. His research interest is the use of ferroelectrics in MOS devices and his expected graduation date is in the summer of 1996.

devices.

Gettering during the IC process reduces defects. Gettering processes include mechanical abrasion of the backside, argon ion implant to create damage and phosphorous diffusion. This is effective in removing heavy metals i.e. gold and reducing oxygen induced stacking faults in the active device area. This reduces recombination centers and increases hold time in memory devices.

Immunization to SEU may be increased in CMOS structure by utilizing Twin N and P-Tub and thin epitaxial layers to reduce Charge funneling. Lightly doped drain CMOS structure reduce hot electron effects. [6]

RADIATION DAMAGE IN SEMICONDUCTOR DEVICES

Early efforts were concerned with the formation of displacement defects in bulk crystalline semiconductor materials. Alpha particles, deuterons, electrons, gamma rays and neutrons were all used as radiation sources in these early experiments. The degradation of minority carrier life time is the most sensitive bulk material parameter and it is crucial to devices such as Solar cells and bipolar transistors. [2]

A second form of radiation damage to semiconductor devices is the effect of ionizing radiation in the creation of electron hole pairs. For Ionizing Radiation to occur, the energy of radiation must exceed 1.12 electron volts for silicon and approximately 9 electron volts for silicon dioxide. Surface related phenomena at the Semiconductor Dielectric Interface due to Total Ionizing Dose(TID) Radiation Effects are especially important to MOS Transistors.

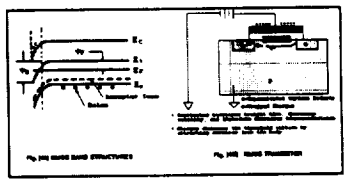
RADIATION EFFECTS ON CMOS TECHNOLOGY

Radiation induced defects in CMOS technology will be emphasized in this paper primarily because of its wide spread use in commercial electronics and space avionics. Also, the expertise developed in the study of the surface and interface related defects provide a firm basis for further study of radiation effects in electronic devices. [2]

Operation of MOS Capacitors and Transistors

An overview of operation of MOS transistors is given by the band diagram Fig. 4B.

$$V_{th} = \phi_{ms} - \frac{Q_{ss}}{C_{ox}} + \frac{Q_s}{C_{ox}} + 2\phi_f$$



Electrically Active Defects

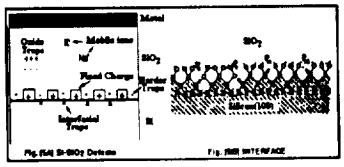
Looking at a MOS structure, it is important to note the type of electrically active defects that may be present at the Si-SiO_2-Gate Structure. A likely source of mobile ions is sodium from human operators during device processing. There is a general surface state charge Q_{ss} which is defined by the following equation.(see Fig.5)

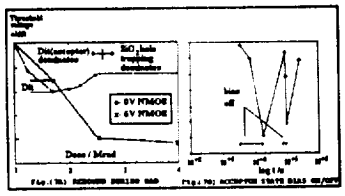
$$Q_{ss} = Q_{flx} + Q_{ot} + Q_{it} + Q_b \qquad (3)$$

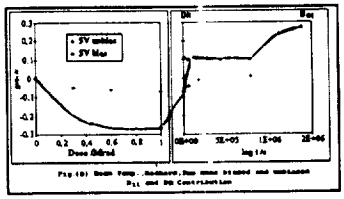
where:

Q_{flx} stands for fixed charge.
Q_{ot} refers to Oxide Trapped Charge.
Q_{it} stands for Interface State Traps.
Q_{bt} are Border Traps.

Threshold voltage shift during total dose radiation is shown during radiation [Fig. 7A] exposure for NMOS devices with 0 bias and under inversion bias. The dose rate is approximately 1 $Mrad_{Si}$ /hr. In both cases, interface states are created. However, under inversion bias, upper half band gap acceptor states are filled producing a negative charge at the interface offsetting the positive charge of deep hole trapping. This D_{it} area effect saturates at approximately 1 $Mrad_{Si}$ after which bias assisted deep hole trapping dominates. Silicon dangling bonds are created during radiation with or without bias. Acceptor states may be filled during post-radiation time under







inversion bias with a sufficiently high Fermi level. In fact, if the bias is cycled on and off, the acceptor states are filled and emptied with only a gradual annealing decay in negative trapped charge at room temperature. This has been called rebound phenomena. (See Fig. 7b)

Similar devices with interface states (D_{it}) as well as border traps (D_{bt}) which are both in communication with the substrate, are shown in Fig. 8. In this case, acceptor interface state filling with associated negative charge is delayed to the post-radiation period. (See Fig. 8) (Table I & II).

$$-\frac{I_{cp}}{f} = qA_G \int_{Ei}^{Eu} D_{it}(E)\, dE$$

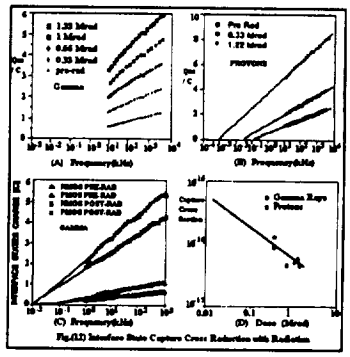
Where E_u and E_i are the conduction and valence band edges and A_g is the area of the gate.

$$v_{th}\sqrt{\sigma_h \sigma_e n_i}\left|\frac{V_{th} - V_{fb}}{\Delta} V_G\right|$$

Where δ_h and δ_e are the hole and electron capture crossections.

OBSERVATIONS AND HYPOTHESIS ON DEFECTS At The SILICON SILICON DIOXIDE INTERFACE

We have been examining radiation and process induced defects at the silicon/silicon-dioxide interface in the MOS system for some time [1;1b]. In the previous data, MOS devices were exposed to gamma and proton irradiations at equivalent dose rates. Various electrical parameters were measured before, during and after the irradiations. The following conclusions were drawn:



CONCLUSIONS

1. The average capture cross section of interface states was reduced with 1 Mrad radiation of 1 MeV protons. It was hypothesized that new interface states were created during radiation due to the H^+ nature of the protons. This was negated because similar results were obtained using gamma radiation. (Fig. 12A, B, & D)

2. For radiation of NMOS

devices under inversion bias, acceptor states in the upper half band gap are filled. If the bias is removed after annealing, electron transport will remove the charge but not the states so that the effect is reversible with bias charge. It was hypothesized that the primary reason for the average capture cross section reduction is the creation of more electron trapping interface states (assuming that electrons have a smaller capture cross section). This was negated because under inversion bias, both PMOS and NMOS devices exhibit similar capture cross section reductions. (Fig. 12C)

3. The density of interface states increased at approximately the same rate with total dose for both proton and gamma radiations. This implies that the change in capture cross section may depend on two separate species (possibly the P_{b0} and P_{b1} states). It is not clear however, the exact nature of these states. Electron of Spin Resonance experiments with P. Lenahan (Penn State) may resolve this issue.

4. Radiation at liquid nitrogen temperature does not create interface states. In fact, D_{it} interface state density remains approximately constant through liquid nitrogen gamma radiation and post-radiation anneal. However, again we find that the capture cross section of the interface states is reduced thus a structural charge in the interface state dangling bond is again suggested. (See Fig. 13) The results of the AC Conductance measurement shown in Fig. 14 shows the appearance of a second maxima at the lower frequency in the post-gamma irradiation that may again be due to structured charge in the P_b center even the possible introduction of border traps.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the work of former and current graduate students: Kurt Kloesel, Tania Thomas, Zhong You, Danzhi Lin and Gordon Wilhite; staff and associates of the PVAMU Laboratory for Radiation Studies.

This work is partially supported by NASA grants NAG5-929, NAG9-331, NAG9-333 and NAG9-659.

REFERENCES

1. V. Zajic, K. Kloesel and T.N. Fogarty. "Single Event Upset and Total Dose Radiation Effects on Rad-Hard SRAMs," In the Second Workshop on Radiation-Induced and/or Process Related Electrically Active Defects in Semiconductor-Insulator Systems, Microelectronics

Center of North Carolina, Sept. 1989, p. 170.

TABLE I

	PreRad		Post Radiation		
		1MRad	1hr	3hrs	9days
D_{it}(#/cm^3)	5.98×10^{10}	1.17×10^{11}	1.15×10^{11}	1.24×10^{11}	1.25×10^{11}
f_0(Hz)	24.88	10.00	5.47	4.52	5.016
σ(cm^2)	1.11×10^{-16}	4.45×10^{-17}	2.43×10^{-17}	2.6×10^{-17}	2.23×10^{-17}
V_{th}(V)	1.115	.848	1.044	1.224	1.217
ΔV_{th}(V)	0	-.262	.071	.109	.102

Charge pumping and threshold voltage results for 5μm radhard nmos device biased during and after radiation at room temperature.

TABLE II

	PreRad	DuringRadiation			Post Radiation		
		.3MRad	.6MRad	1MRad	1.5hr	4hrs	16hrs
D_{it}(#/cm^3)	9.21×10^{10}	9.4×10^{10}	1.1×10^{11}	1.33×10^{11}	1.36×10^{11}	1.39×10^{11}	1.48×10^{11}
f_0(Hz)	18.11	.60	.26	.20	.31	.37	8.9
σ(cm^2)	8.08×10^{-17}	2.68×10^{-18}	1.16×10^{-18}	8.92×10^{-19}	1.38×10^{-18}	1.65×10^{-18}	3.97×10^{-17}
V_{th}(V)	1.234	1.185	1.172	1.163	1.181	1.186	1.194
ΔV_{th}(V)	0	-.049	-.062	-.071	-.053	-.048	-.04

Charge pumping and threshold voltage results for 5μm nmos radhard nmos device radiated at room temperature with no bias.

SCIENCE AND COMMUNICATIONS ON WHEELS

William Nazario-Irizarry, Dr. Juan López-Garriga

Contribution from the Chemistry Department, University of Puerto Rico, Mayagüez Campus, Mayagüez, P.R. 00680-5000.

* To whom correspondence should be addressed. E-mail address: J_lopez@rumac.upr.clu.edu
# J.L.G. is a Camille and Henry Dreyfus Scholar.

BIOGRAPHY

The author is an undergraduate student of the University of Puerto Rico at Mayagüez in his fourth year of the Electrical Engineering Department. National Hispanic Scholar in 1992, and member of the IEEE since 1992. He has been working under the Science and Wheels program for four years.

ABSTRACT

We have developed a Science and Space Communications on Wheels Program focusing on diverse population of precollege students and teachers. Initially this program evolved from a chemistry demonstration shows program consisting also of teachers training and the incorporation of demonstrations by teachers into their precollege chemistry courses. In a period of three year, the Program has exposed over 5,000 students, 335 teachers and 289 relatives and administrators to science phenomena. The Program also offers workshops to precollege teachers to train them in the presentation of chemical demonstrations in their own chemistry courses. The evaluations made indicated that close to 90% of these students would like to see more activities of this kind included as part of their science courses. Workshops and direct communication of between precollege and college educators is recognized as one solution to increase teachers confidence. An evaluation of the demonstrations presented by precollege teachers in their science classes showed an overwhelming acceptance by their students, with around 80% of the students giving the maximum rating to the demonstrations presented by their teachers . Regarding this we are expanding the program to explore remote sensing data and its applications to earth studies. Initially we are focusing our efforts to seminars and demonstrations consisting of the discussion of electromagnetic waves, satellite orbitals mechanics and the applications. Our results indicated an excellent response and motivation among the students to participate in this initiative.

PROGRAM DEVELOPMENT

In these increasingly complex times, studying a career in science has become an interacting phenomena of disciplines. Interdisciplinary education, experimental knowledge and hands-on experience appear to be components of a successful science education. To reinforce these abilities in junior high school and early high school students and their teachers, higher education institutions should strive to supply effective ways to enhance their knowledge and experimental talents. The students need to learn, from their science teachers, how to search for opportunities to develop adequate skills in writing, speaking, reading critically, reasoning, judging, supervising and hands-on laboratory experiences, so that they may choose and aspire to a career in the sciences. The students can no longer be taught, exclusively, through a process of comprehending and memorizing facts. They should also develop skills to think critically and logically through scientific arguments. The achievement of these goals is fundamental in the development of human resources for the next decades.

Science education, an essential part of our competitive environment, is hardly accessible to a large number of diversified secondary students. This is, partially, because their schools and teachers have limited access to adequate information, science materials, scientific instrumentation and hands-on experiences. The lack of these materials translates into apathy from the students, and

creates an aura of inaccessibility towards the sciences for all but the brightest of them. To overcome such educational barriers it is necessary to develop an efficient link between higher and secondary education. This coherent link between these two educational perspectives can improve the general understanding and the background knowledge which a student should have if he or she wishes to pursue and succeed in a science career (1-9).

The University of Puerto Rico at Mayaguez (UPRM) is an institution with approximately 11,000 Hispanic students and is located on the western region of Puerto Rico. The institution offers sixteen different B.S. degrees in four major colleges; Engineering, Arts and Sciences, Agriculture, and Business Administration. Approximately 60% of UPRM undergraduate students are science, mathematics or engineering majors, and from 1988 to 1995 the number of undergraduate students in science increased from 1,134 to 2,239. During this period, UPRM conferred 4,356 baccalaureate degrees in science and engineering. Efforts have been made by the Mayaguez Campus in general, and by the departments of Chemistry and Physics in particular, to encourage students to pursue careers in science and engineering. In order to do this effectively, we must not only be able to identify those high school students who have shown potential for careers in the sciences, but also those who have the potential for pursuing these disciplines and have not been properly oriented or motivated to continue studies in these areas.

At the Departmental level, we have developed a cost effective plan which consists of chemistry science shows and a training program for secondary school science teachers, coupled to a mobile unit. The Science on Wheels mobile unit commutes resources from the UPRM Campus to schools in the Island as a way to achieve the exposure of K-12 grade students to the "science phenomena" through science demonstrations.[1] The process is helping to consolidate and mature the knowledge of precollege science teachers, and to enhance the opportunities of their students in the sciences. It is very important to realize that this dual integration will benefit both secondary school and college education. Such networking of information has the potential to prepare and motivate students in science and engineering toward the, not so far, twenty first century. The Program's success is preceded by the effectiveness of analogous programs, such as mobile spectroscopy laboratories (10-11), "Chemistry on Wheel"(12), "The MIT Chemistry Outreach Program" (13), "The Twelve Days of Chemistry (14), "Science Bag" (15), "The Scientist-in-Residence Program" (16), "Chemistry is Fun"(17), "Not So Late Night Chemistry with USD"(18) and others initiatives (19-20).

The Program's targets are rural areas and towns in which schools have little access to hands-on experiences and technology in science education. The main link with the K-12 teachers is a series workshops where they have the opportunity to develop their science skills through hands-on experiences. However, it is important to realize that to present demonstrations and workshops without the proper scientific preparation can be misleading in terms of the message and purpose of the activity (21-24). Simple and clear exploration and discussion of the scientific background of the demonstrations and workshops, followed by clearly designed metaphors and analogies used to explain the scientific phenomena, can reinforce the learning experience of the attendees (25-29). As part of the workshop training, the teachers learn how to produce interesting and safe experiments for their students with easily obtainable materials, so that the experiments can be reproduced by the students at home. This is a way to achieve the

exposure of teachers, students and parents to the science phenomena.

The response from high school's students to our efforts have been so overwhelming that we are expanding our program to the communications environment. By developing conference with certain demonstrations of what are electromagnetic waves and how they behave, and the mechanics of satellite orbits we explain them how this physics principles are used daily to link the whole world through radio and space communications. Showing them with simple homemade antennas we demonstrate the simplicity to receive the signals sent by satellites in the stratosphere. In this way they start to realize that space communications are as familiar to them as turning on their home radio sets.

Using simple models of orbits we explain them how is possible to launch a satellite and keep it in orbit, change its orbit and how its use determines the type of orbit that best suites it. Analogies, like our Moon are used in order to explain this phenomena. establishing this concepts we move to the uses that we give to this advantages in all the fields like the scientific, communications and entertainment industry. We pay a greater attention to the concepts of Remote Sensing an Global Positioning Systems because of their widely spread use in the scientific matters. Science the Weather uses are one of the most widely known we provide them a more detailed explanation of certain orbits used to put those satellites and how we can reach the signals of the satellites from NOA.

The Science and Communications on Wheels Program allows the UPRM : a) to implement a cost effective program in communications for the K-12 students and teachers, b) to motivate precollege students to enroll in science courses, c) to motivate student to participate more actively in science oriented programs, d) to provide the means to integrate student development in the area of science with faculty in-campus development, e) to make available scientific equipment to K-12 science teachers after providing hands-on workshops and laboratory experience in the UPRM facilities, f) to involve high school students and teachers in research to develop relevant demonstrations to be performed in their schools during the regular academic year. Both science demonstrations and laboratory experiences are means of triggering student interest, focusing attention and, in many cases, initiating student learning (7,22).

FUTURE PLANS

We are developing some demonstrations using small satellite dish antennas in order to receive geostationary satellites data at the rural schools that we might visit. In this way the impact of the program will be more widely spread over our island. Also we are testing other types of ways to get our data like through Amateur Radio Bands and linking our mobile unit with a soon to be establish information center at our campus via cellular phone to get access to the Internet or the data from our future antennas.

Certain workshops are planned to give hands on experience to the teachers and students in our laboratories for them to be able to construct antennas and get and interpret the data. In the meantime we are establishing links to get the equipment and information needed to continue our applied science demonstrations.

REFERENCES

1. Crosby, G.A. *J. Chem. Educ.* **1989**, *66*, 4-7.

2. Bok, D. *Higher Learning*; Harvard University Pres: Cambridge, Mass. 1986.

3. Haderlie, H. *J. Chem. Educ.* **1994**, *71*, 1058-1062.

4. Smith, S.G. and Jones L.L. *J. Chem. Educ.* **1989**, *66*, 8-13.

5. Moore, J.W. *J. Chem. Educ.* **1989**, *66*, 15-19.

6. Gardner, M. *Chem. Soc. Rev.* **1988**, *17*, 135-146.

7. Swinehart J.H., *J. Chem. Ed.* **1987**, *64*, 429-430.

8. Schwartz, A.T.; Bunce, D.M.; Silberman, G.R.; Stanitski, C.L.; Stratton, W.L.; Zipp, A.P. *J. Chem. Ed.* **1994**, *71*, 1041-1046.

9. Winther, A.A.; Volk, T.L. *J. Chem. Ed.* **1994**, *71*, 501-505.

10. Roberts, T.D. *J. Chem. Educ.* **1978**, *55*, 279-301.

11. Smith, D.H. *J. Chem. Educ.* **1986**, *63*, 68-69.

12. Newman, A.R. *Anal. Chem.* **1990**, *62*, 449A-450A.

13. Noviwick, J.S. and Brisbois, R.G. *J. Chem. Educ.* **1989**, *66*, 668.

14. Gammon, S.D.; Brauer, R.D.; Brauer, T.M.; Daniel, N.; Farah, F.M.; Fisher, J.; Pharr, C.M.; Phelps, C.L.; Richardson, R. *J. Chem. Ed.* **1994**, *71*, 1077-1079.

15. Shaw III, C.F.; Greenler, R.G.; Lasca, N.P.Jr.; Brooks, A.S. *J. Chem. Ed.* **1992**, *69*, 1020 1023.

16. Kelter, P; Hughes, K; Murphy, A.; Condon, K.; Heil, P; Lehman, M.; Netz, D.; Wager, T. *J. Chem. Ed.* **1994**, *71*, 864-866..

17. Sherman, M.C *J. Chem. Ed.* **1992**, *69*, 413-415.

18. Koppang, M.D.; Webb, K.M.; Srinivasan, R.R. *J. Chem. Ed.* **1994**, *71*, 929-931.

19. Deavor, J.P.; Deavor, J.W. *J. Chem. Ed.* **1995**, *72*, 798.

20. Turner, K. *J. Chem. Ed.* **1995**, *72*, 203.

21. Kelter, P. *J. Chem. Ed.* **1994**, *71*, 119-120.

22. O'Brien, T. *J. Chem. Educ.* **1991**, *68*, 933-936.

23. Roadruck, M.D. *J. Chem. Educ.* **1993**, *70*, 1025-1028.

24. Lin, H.S.; Lawrenz, F. *J. Chem. Educ.* **1991**, *68*, 904-907.

25. Miller, L.T. *J. Chem. Ed.* **1993**, *70*, 187-189.

26. Bowen, C.W. *J. Chem. Ed.* **1992**, *69*, 479-482..

27. Katz, D.A. *J. Chem. Ed.* **1991**, *68*, 235-244.

28. Bhushan, N.; Rosenfeld, S. *J. Chem. Ed.* **1995**, *72*, 578-582.

29. Ditzler, M.A.; Ricci, R.W. *J. Chem. Ed.* **1994**, *71*, 685-688.

Footnote

[1] Presented in part at the ChemEd 93, Indianapolis, Indiana, August 1993; at the 13th International Chemical Education Conference, San Juan P.R., August 1994; and at ChemEd 95, Norfolk, Virginia, August, 1995.

COMPUTER CONTROLLED MAGNETOTRANSPORT SETUP FOR THE CHARACTERIZATION OF SEMICONDUCTOR THIN FILMS

G. O. Ducoudray[1], R. Collazo, and A. Martínez, Tropical Research Center of Earth and Space Studies, University of Puerto Rico, Mayagüez, 00682, Puerto Rico and Department of Physics, University of Puerto Rico, Río Piedras, San Juan, Puerto Rico 00931-3343

ABSTRACT

We have constructed a computer controlled magnetotransport setup using LabWindows environment. It allows for measurements of resistivity, Hall resistance, carrier concentration and charge mobility in semiconductor thin films using a van der Pauw configuration. The setup features an electromagnet (B = 0.7 Tesla), a 80486-DX 33 computer with a National Instrument AT-MIO 16 AD/DA and a GPIB interface board. A Keithley 224 current source and a Keithley 196 digital voltmeter are also used in the setup. Plans for the addition of capabilities to allow for magnetic field sweeping and the performance of measurements as a function of temperature will be presented.

[1] Gladys Omayra Ducoudray is a master student in the Physics Departmnet of the University of Puerto Rico in Río Piedras Campus. She holds a BS in Physics from the same university. Her major interest in research is growth and characterization of thin film wide band gap semiconductor materials.

I. INTRODUCTION

The values of the electrical properties of semiconductors vary in orders of magnitude with changes in temperature, optical excitation and impurity content. This variability of electrical properties makes the semiconductor material natural choices for electronic device development.

There are numerous semiconducting materials and the wide variety of electronic and optical properties of these semiconductors provides potentially great flexibility in the design of electronic and optoelectronic functions. One of the most important physical properties in the determination of the applicability of a particular semiconductor for a specific device application is its energy band gap. For example, the requirements for higher component packing densities in microcircuits and the desirability of semiconductor lasers that operate in the violet and ultra-violet region of the electromagnetic spectrum have fueled recent growth and characterization efforts in the group-III nitrides family of wide band gap semiconductors.[1-2]

The development of these novel device technologies based on new semiconductor materials requires intense efforts to establish the optimal growth conditions. In particular, an understanding of the effects due to the introduction of impurities in the electrical properties of the materials, e.g. the type and concentration of the charge carriers introduced through doping, are of critical for any future device fabrication.

We describe here the construction and operation of a computerized setup to measure the magnetotransport properties of semiconductor thin films. The setup implements the van der Pauw technique to measure the electrical properties, i.e. charge carrier density, mobility and resistivity, of these materials.

II. VAN DER PAUW METHOD

The van der Pauw configuration consists of four point contacts used in pairs to alternately inject a constant current through the sample and to probe the electric potential difference across the sample. The contacts used should be small and should be placed on the perimeter of the sample. A typical van der Pauw probe arrangement in a square configuration, is shown in Figure 1. The resistivity and the carrier concentration of arbitrarily shaped and flat samples can be determined using this method. The samples must be of constant thickness and should contain no isolated holes, to insure that there are no alternate non equivalent current paths through the sample.

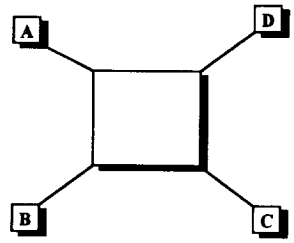


Figure 1 *Typical van der Pauw configuration*

In order to obtain the resistivity of the sample, a total of four measurements are needed. Providing a constant current source through probes A and B and measuring the voltage between C to D to the sample, V_{CD}. is obtained. Changing the probes so that the constant current flows from probe B to C and measuring the voltage between D and A, V_{DA} is obtained. Repeating this process and inverting the current sense, values for V_{AD} and V_{DC} are obtained. We define the voltage

$$|V_{DA}| = \frac{|V_{DA} - V_{AD}|}{2},$$

$|V_{CD}|$ is defined analogously. Defining a "resistance", $R_{AB,CD} = |V_{CD}| / I_{AB}$ and $R_{BC,DA} = |V_{DA}| / I_{BC}$.,van der Pauw proved that the resistivity of the sample and these "resistances are related through the expression:

$$\exp(-\pi\frac{d}{\rho}R_{AB,CD}) + \exp(-\pi\frac{d}{\rho}R_{BC,DA}) = 1.$$

In the previous equation we can not solve for ρ analytically. We define a factor f such that:

$$\rho = \frac{\pi}{\ln 2} d \frac{R_{AB,CD} + R_{BC,DA}}{2} f$$

where d is the thickness of the film. Note that f is a geometrical factor that satisfies the following equation

$$\cosh\left(\frac{\ln 2}{f}\,\frac{\frac{R_{AB,CD}}{R_{BC,DA}-1}}{\frac{R_{AB,CD}}{R_{BC,DA}+1}}\right)=\frac{1}{2}\exp\left(\frac{\ln 2}{f}\right).$$

This expression can be solved numerically for f in terms of the resistance ratio $R_{AB,CD}/R_{BC,DA}$, and hence, the resistivity can be calculated also. A value of the resistance ratio larger than 2 is indicative of either bad contacts or of inhomogeneous sample doping.

For Hall measurements, the choice of probes is made such that electric potential is measured transverse to the injected current. The resistance $R_{AC,BD}$ is measured with the sample placed in a perpendicular homogeneous magnetic field, The resistance change, $\Delta R_{AC,BD}$, is related to the Hall coefficient through the expression

$$R_H=\frac{d}{|B|}\Delta R_{ACBD}$$

where B is the intensity of the magnetic field and d is the thickness of the sample. The Hall coefficient is inversely proportional to the charge carrier concentration, thus by knowing the Hall coefficient we are able to determine carrier concentration by

$$R_H=\frac{1}{ne}.$$

III. EXPERIMENTAL SETUP

Our magnetotransport setup characterization is composed of an electromagnet and associated power supply, a Keithley 196 digital multimeter, a Keithley 705 scanner, a Keithley 224 current source, and a Hall Probe. These instruments are controlled by a 80486 DX-33 computer outfitted with a National Instruments AT-MIO 16 AD/DA board and a General Purpose Interface Bus (GPIB) interface board. The scanner has eight input/output data arranged in a matrix array. The scanner is programmable via the GPIB interface and allows for the independent connection of the digital multimeter and the current source to any of the sample probes. A constant current source Keithley 224 provides the current for the sample. The AT-MIO 16 AD has 16 channels for data information and is used to measure the voltage from the Hall probe to monitor the intensity of the magnetic field. A home-made circuit, shown in figure 3, provides a constant current for the Hall probe and an amplifier circuit converts the output of the Hall probe into a voltage which in millivolts equals the magnetic field intensity in Tesla. This amplified output is fed into the computer and into a panel meter that allows for instantaneous visual inspection of the magnetic field intensity. The automation was done using LabWindows environment. This environment is compatible with common computer languages such as BASIC or C.

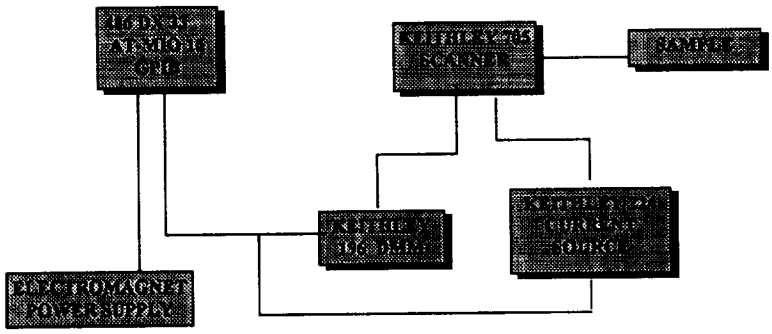


Figure 2 *Schematic Diagram of Automation Set up*

The LabWindows environment also contains various libraries containing instrument drivers for most of the instruments in the setup. This environment is compatible with common computer languages such as BASIC or C and features complete control with window-like commands.

The van der Pauw characterization automation program first initializes the instruments in the GPIB. After this process is completed, the user inputs his/her name, the sample name, the sample thickness and the temperature and magnetic field at which the measurement will be performed i. e. 300 or 77K. If a value of zero is entered for the magnetic field, the program will proceed to make as a resistivity measurement. A non zero value to the magnetic field will cause the program to perform resistivity and Hall measurements.

The program calculates the standard deviation of the voltage being measured and provides an on screen warning when noisy data is detected. After the data acquisition is over, the computer stores the data in ASCII format making the data retrievable for further studies.

IV. ACKNOWLEDGMENTS

This work was sponsored by the NASA Minority University Research Center Program.

V. REFERENCES

1. S. Strite, H. Morloc, J, Vac. Sci. Technol. B 10, 1237, 1992.

2. H. Morcok, S. Strite, G.B. Gao, M. E. Lin, B. Sverdlov, and M. Burns, J. Appl. Phys. 76,1363, 1994

3. K. Seeger, Semiconductor Physics, Springer-Verlag, New York, 1973, pp 483-488

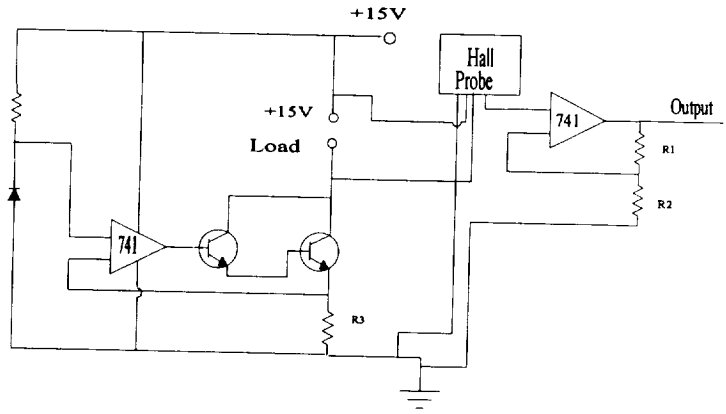


Figure 3 *Hall probe constant current source and readout circuit*

NUMERICAL SOLUTION OF RANDOM ALGEBRAIC EQUATIONS

Jenell Sargent and S. Sathananthan
CASS Center, Tennessee State University, Nashville, TN 37203

ABSTRACT

In mathematical and engineering applications, it is well known that the knowledge of the solutions of algebraic equations is important. We have investigated several methods involving random parameters in an algebraic equation and some numerical methods to produce approximate solutions of random nonlinear equations. An example is provided to demonstrate the obtained results.

Keywords: Random polynomials, algebraic equations, convergence of Numerical schemes.

INTRODUCTION

Algebraic equations play an important role in many scientific fields. Whenever the known parameters in an algebraic equation are subject to uncertainty, a more accurate mathematical model can be obtained by expressing the parameters in the equation as random variables.

Jenell Sargent is a graduate student in Mathematics at Tennessee State University. She plans to finish her M. S. in May 1996. She has worn many hats throughout her matriculation . Ranging from math/computer lab assistant to Adjunct Faculty. Her interests include mathematical computing and computational simulation.

This leads to a random equation, which, in general, is more difficult to solve than its deterministic counter parts [1, 3, 4]. Most of the recent research [1, 4] on random equations has concentrated on measurability of roots, number of roots, and distribution of roots of random algebraic equations. Usually numerical methods are often used to approximate solutions of deterministic equations, and they also can be applied successfully to approximate random algebraic equations. In this paper, we compared several numerical methods with different orders of convergence for a random algebraic equation in one unknown. An example is provided to demonstrate the effectiveness of these numerical methods.

MAIN RESULTS

The Newton's method for approximating the solution $R(\omega)$ of the random nonlinear equation $G(x, \omega) = 0$ is

$$X_{n+1}(\omega) = X_n(\omega) - \frac{G(X_n(\omega),\omega)}{G'(X_n(\omega),\omega)}, n \geq 0 \quad (1)$$

where X_0 is given, converges to $R(\omega)$ with probability 1 and order of convergence 2.

The following theorem establishes the convergence of the Newton's method.

Theorem 1: Let $G(x,\omega)$ be defined on the probability space (Ω,β,P). Let $\Omega_0 \subset \Omega$ be such that $P(\Omega_0)=1$. For each $\omega \in \Omega_0$, let $G(R(\omega),\omega)=0$, $G'(R(\omega),\omega)\neq 0$, and $G(x,\omega), G'(x,\omega), G''(x,\omega)$ be continuous for all values of x in the finite interal I. Let

$$M \equiv M(\omega) = \sup_{\omega \in \Omega_0} \max_{I} \left| \frac{G''(x,\omega)}{2\ G'(x,\omega)} \right| < \infty \qquad (2)$$

Let ϵ exist such that $0<\epsilon<\min(1,1/M)$, and $[R(\omega)-\epsilon, R(\omega)+\epsilon] \subset I$ for all $\omega \in \Omega_0$. Let $x_0(w)$ be a random variable such that for all $\omega \in \Omega_0$,

$$|X_0(\omega)-R(\omega)|<\epsilon \qquad (3)$$

Then the sequence $\{x_n(\omega)\}$ defined by equation (1) converges with probability 1 to $R(\omega)$ and the order of convergence is 2.

Proof: For each $\omega \in \Omega_0$, it can be easily shown [2] that $\{x_n(\omega)\}$ converges to $R(\omega)$ and

$$\lim_{n\to\infty} \frac{|X_{n+1}(\omega)-R(\omega)|}{|X_n(\omega)-R(\omega)|^2} = \frac{1}{2}\left|\frac{G''(R(\omega))}{G'(R(\omega))}\right| \le M_1 \qquad (4)$$

Since M_1 is constant, the order of convergence is 2.

Next we discuss the Secant Method. The Secant method for approximating the solution $R(\omega)$ of the random nonlinear equation $G(x,\omega)=0$ is

$$X_{n+1}(\omega) = X_n(\omega) - G(X_n(\omega),\omega) \cdot \left[\frac{X_n(\omega)-X_{n-1}(\omega)}{G(X_n(\omega),\omega)-G(X_{n-1}(\omega),\omega)}\right] \qquad (5)$$

$n \ge 1$, where X_0 and X_1 are given.

The following theorem establishes the convergence of the secant method.

Theorem 2: Let $G(x,\omega)$ be defined on the probability space (Ω,β,P). Let $\Omega_0 \subset \Omega$ be such that $P(\Omega_0)=1$. For each $\omega \in \Omega_0$, let $G(R(\omega),\omega)=0$, $G'(R(\omega),\omega)\neq 0$, and $G(x,\omega), G'(x,\omega), G''(x,\omega)$ be continuous for all values of x in the finite interval I. Let

$$M = \sup_{\omega \in \Omega_0} \max_{I} \left| \frac{G''(x,\omega)}{2\ G'(x,\omega)} \right| < \infty \qquad (6)$$

Let ϵ exist such that $0<\epsilon<\min(1,\frac{1}{M})$ and

$[R(\omega)-\epsilon, R(\omega)+\epsilon] \subset I$ for all $\omega \in \Omega_0$,

$$X_0(\omega) \neq X_1(\omega),$$
$$|X_0(\omega)-R(\omega)|<\epsilon,$$
$$|X_1(\omega)-R(\omega)|<\epsilon,$$

Then the sequence $\{X_n(\omega)\}$ defined by equation (3) converges with probability 1 to $R(\omega)$ and the order of convergence is at least $\frac{1+\sqrt{5}}{2}$.

Proof: For each $\omega \in \Omega_0$ it can be easily shown [2] that $\{X_n(\omega)\}$ to $R(\omega)$ and

$$\lim_{n\to\infty} \frac{|X_{n+1}(\omega)-R(\omega)|}{|X_n(\omega)-R(\omega)|^p} = \frac{1}{2}\left|\frac{G''(R(\omega))}{G'(R(\omega))}\right|^{p-1} \le M_2 \qquad (7)$$

where $p = \frac{1+\sqrt{5}}{2}$. Since M_2 is a finite constant, the order of convergence is p. Finally , we will discuss the Muller's method.

The Muller's method for approximating the solution $R(\omega)$ of the random nonlinear equation $G(x,\omega) = 0$ is,

$$X_{k+1}(\omega) = X_k(\omega) - \frac{2G'(X_k,\omega)}{W_2 \pm \sqrt{W_2^2 - 4G(X_k,\omega)G[X_k,X_{k-1},X_{k-2},\omega]}} \quad (8)$$

where,

$W_2 = G[X_k,X_{k-1},\omega] + G[X_k,X_{k-2},\omega] - G[X_{k-1},X_{k-2},\omega]$ where $X_0(\omega)$, $X_1(\omega)$ $X_2(\omega)$ are given random variables.

The following theorem establishes the convergence of the Muller's Method.

Theorem 3: Let $G(x,\omega)$ be defined on the complete probability space (Ω,β,P). Let $\Omega_0 \subset \Omega$ be such that $P(\Omega_0) = 1$. For each $\omega \in \Omega_0$, let $G(R(\omega),\omega) = 0$, $G'(R(\omega),\omega) \neq 0$ and $G(x,\omega)$, $G'(x,\omega)$, $G''(x,\omega)$, $G'''(x,\omega)$ be sample continuous for all values of x in the finite interval I. Let

$$M \equiv M(\omega) = \sup_{w \in \Omega_0} \max_{I} \left| \frac{G^{(3)}(\eta_k,\omega)}{6[G'(\xi_k(\omega)+K(\eta_k,\omega))]} \right| < \infty \quad (9)$$

where $K(\eta_k(\omega)) = [(R(\omega) - X_k) + (R(w) - X_{k-1}) - (R(\omega) - X_{k+1})] \cdot \frac{G''(\eta_k,\omega)}{2}$.

Let $X_0(\omega)$, $X_1(\omega)$, $X_2(\omega)$ be random variables such that

$$M|R(\omega) - X_i(\omega)| < 1,\ i = 0,1,2 \quad (10)$$

Then, the sequence $\{X_k(\omega)\}$, defined by (8) converges to $R(\omega)$ $\omega \cdot p \cdot 1$ and the order of convergence is $q \approx 1.84$. Moreover,

$|R(w) - X_k(w)| \leq M^{-q_k^2 \delta_{k-1}^q}$, where $0 < \delta < 1$, such that $M|R(w) - X_i(w)| \leq \delta$, i= 0, 1, 2, q_{k+1} is the solution process of $q_{k+1} = q_k + q_{k-1}, + q_{k-2}, q_0 = q_1 = q_2 = 1$ and q^0 is the solution process of $q_k^2 = q_{k-1}^2 + q_{k-1}^1$, $q_k^1 = q_{k-1}^2 - q_{k-1}^0$, $q_k^0 = q_{k-1}^2$, $[q^0,q^1,q^2] = [0,1,0]$.

Proof: The proof can be easily obtained by following [3].

Numerical Example

To illustrate the effectiveness of these methods, we will provide an example [3]. In this, we consider a random algebraic equation. Since Newton's method is analogous to the Secant method we will compare the Secant and Muller's method.

Let us find the real roots of $x^2 + Bx - C = 0$ where B and C are independent random variables with respective probability density functions:

$$f_B(b) = \begin{cases} 1, & 1 \leq b \leq 2 \\ 0, & 0 / w \end{cases}$$

$$f_C(c) = \begin{cases} 0.5, & 2 \leq c \leq 4 \\ 0, & 0 / w \end{cases}$$

The following table represent estimates for the positive root of $x^2 + Bx - C = 0$ for a sample of size 100.

Method	$E(X_k)$	$V(X_k)$
Secant Method	1.33889	0.003175
Muller's Method	1.13388	0.003175

We notice that the Secant method started at $x_0 = 1, x_1 = 2$ and the number of iterations never exceeded 6 until the successive iterates differed by less than 10^{-4}. In the Muller's method the average number of iterations was 2. In the Newton's and the Secant method the cost of evaluating G' as compared to G is a factor. We also note that in the Secant method real choices $X_0(\omega)$ and $X_1(\omega)$ will lead to a real value of $X_2(\omega)$ but in Muller's method real choices of

$X_0(\omega)$, $X_1(\omega)$, $X_2(\omega)$ may lead to complex root of $G(x, \omega)$.

References

1. Bharuda-Reid, A. T. , and Sambandham, M., "Random Polynomials, Academic Press, New York, 1986.

2. Burlish, R., and Stoer, J., "Introduction to Numerical Analysis", Spiner-Verlag,1993.

3. Ladde, G. S., and Sambandham, M., "Numerical Treatment of Random Polynomials", Applied Mathematics and Computation 55:13-30 (1993).

4. Lax, M.D., "Numerical Solution of random nonlinear equations", Stoch-Anal-Appli. 3-163-169 (1985).

290517 P5

An Analytical Study of the Effects of Environmental Contaminants on Global Warming Using Finite Element Computer Modeling

Patricia Moon and Darrell Hinton
Center for Computational Methods
Tennessee State University
Nashville, Tennessee

Abstract

As a means of addressing one of the many important questions currently circulating in the scientific community, namely whether the presence of CO_2 gas in the atmosphere contributes to the phenomena of global warming, we calculated the CO_2 distribution in a spherical atmosphere do to a point source at the center of the sphere. We calculated both the steady-state and transient diffusion, and diffusion flux assuming constant diffusion coefficient and a diffusion coefficient which varied with temperature. We found no appreciable difference when using the variable coefficient as opposed to the constant coefficient. We found the steady-state diffusion of CO_2 in air to vary as $[a+br;\ 0\le r\le 1;\ a+b\ln(r);\ r>1]$ and the steady-state diffusion flux to vary as $a(1-b\tan^{-1}(r))$, where a and b are arbitrary constants (not necessarily the same) and r is the radial coordinate. For a fixed radial value, we found the transient difusion flux of CO_2 in air to vary as $a(1-b\tan^{-1}(t))$ and the transient diffusion to vary as $[a(1-t^{-1});\ r>0]$, a and b again being arbitrary constants (not necessarily the same).

Introduction

It has long been under debate in the scientific community as to whether or not the presence of CO_2 gas in the atmosphere contributes to global warming. One way to prove (or disprove) this supposition is to calculate the heat transfer in both an atmosphere containing air only (with standard concentrations of CO_2) and an atmosphere containing air plus an abundance of CO_2. If it can be shown that the atmosphere having the abundance of CO_2 impedes the flow of heat (i.e., acts as an insulator), then it can be argued that the presence of CO_2 serves to retain heat in the atmosphere and subsequently contribute to global warming. A logical first step towards this objective is to determine the distribution of CO_2 gas in the atmosphere do to a steady CO_2 source. This is precisely the objective we undertook in the present study.

Theory

For the case of cylindrical symmetry, the diffusion of any constituent is governed by

$$\partial\zeta/\partial t = D[\partial^2\zeta/\partial r^2 + (1/r)\partial\zeta/\partial r] \qquad (1)$$

where ζ represents the ratio of the moisture concentration C to the initial moisture concentration C_o, t is the time, and r represents the radial coordinate. The steady state diffusion is given by

$$D[\partial^2\zeta/\partial r^2 + (1/r)\partial\zeta/\partial r] = 0 \qquad (2)$$

and the diffusion flux J is given by

$$J = -D\nabla\zeta. \qquad (3)$$

Finite Element Model

As a means of computing the distribution of CO_2 in the air we chose to model the atmosphere as a sphere containing a point source of CO_2 at its center. The moisture concentration ratio ζ was set to one at the center of the sphere and zero along the boundary of the sphere. Due to the symmetry in the boundary conditions, the surfaces of equal ζ should be concentric spheres centered about the origin of the atmosphere. If one were to cut the spherical atmosphere with a plane which passed through its origin, the distribution of equal ζ on this plane should form concentric circles centered about the origin. This distribution, however, should be approximately the same as the distribution on a plane which intersects a circular cylinder normal to its axis, provided the value of ζ along the axis of the cylinder is equal to one and its value on the surface of the cylinder is zero. Therefore, we chose to model the spherical atmosphere using 51 2-D axisymmetric ring elements. We arranged the elements such that more elements were concentrated near the axis of the cylinder where the value of ζ should be changing the fastest. We chose the following shape functions to approximate the distribution of ζ across each element

$$\zeta = \frac{1}{4}\left[\zeta_I(1-s)(1-t) + \zeta_J(1+s)(1-t) + \zeta_K(1+s)(1+t) + \zeta_L(1-s)(1+t)\right] \qquad (4)$$

where ζ_I, ζ_J, ζ_K, and ζ_L are the nodal values of ζ at each corner of the element and s and t are the vertical and horizontal

element coordinates. We employed the Newmark method [3] to integrate equation (1) in time.

Results

We computed the solutions to equations (1), (2) and (3) on a cylindrical domain with $\zeta=1$ along the axis and $\zeta=0$ along the boundary. The variation of the diffusion coefficient D with temperature is presented in Figure 1. We fit a regression curve to the five data points [2] which were available for the diffusion of CO_2 in air. The steady-state moisture concentration ratio $\zeta = C/C_0$ of CO_2 in air is shown in Figure 2. We performed the analysis assuming both a uniform value for D and a D which varied with temperature. We found the results to be virtually identical, as can be seen in Figure 2. Furthermore, we found the steady-state diffusion of CO_2 in air to vary as

$$[a+br;\quad 0\le r\le 1;\quad a+b\ln(r);\quad r>1], \tag{5}$$

where a and b are arbitrary constants. The steady-state diffusion flux J is presented in Figure 3. We found the steady-state diffusion flux to vary as

$$a(1-b\tan^{-1}(r)), \tag{6}$$

a and b again being arbitrary constants, not necessarily the same as in equation (5). As can be seen in Figure 3., the diffusion flux is greatest in the immediate vicinity of the CO_2 source as it should be.

We present the transient moisture concentration ratio of CO_2 in air at five distinct radial values in Figure 4. We found the moisture concentration ratio to vary as

$$[a(1-t^{-1});\quad r>0], \tag{7}$$

a being an arbitrary constant different from any previous value. We present the transient diffusion flux of CO_2 in air at five distinct radial values in Figure 5. We found the transient flux to vary as

$$a(1-b\tan^{-1}(t)). \tag{8}$$

Future Work

Having obtained the distribution of CO_2 in a spherical atmosphere for a point source, we next plan to calculate the heat flow in the same spherical region having the aforementioned distribution of CO_2. In order to do so, we will have to obtain a suitable temperature distribution throughout the atmosphere.

References

1. Jost, W., 1952, *Diffusion in Solids, Liquids and Gases*, Third Edition, Academic Press Inc., New York, pp. 1-4.

2. Lide, D. R. and Kehiaian, H. V., 1994, *CRC Handbook of Thermophysical and Thermochemical Data*, First Edition, CRC Press, Boca Raton, pp. 421-423.

3. Bathe, K. J., 1982, *Finite Element Procedures in Engineering Analysis*, First Edition, Printice Hall, Englewood Cliffs.

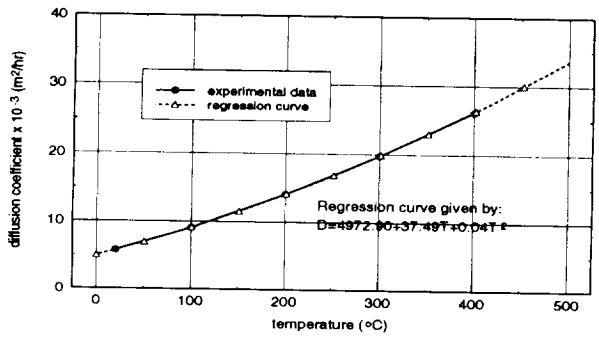


Figure 1. Diffusion Coefficient for CO_2 in Air.

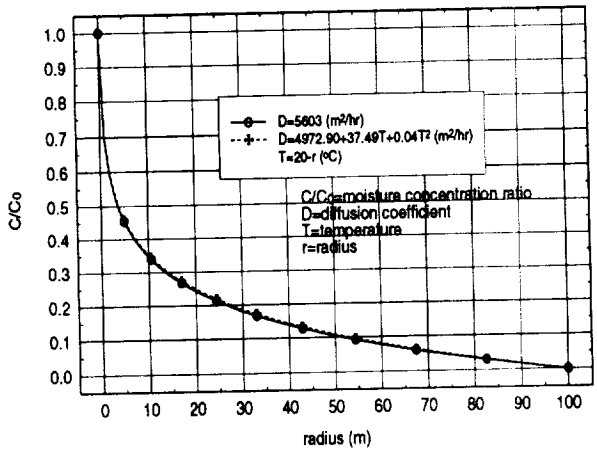


Figure 2. Steady State Diffusion of CO_2 in Air for Uniform and Variable Diffusion Coefficients.

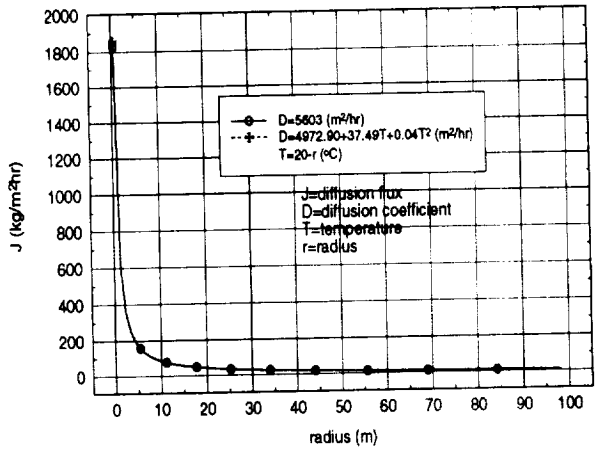


Figure 3. Steady State Diffusion Flux for CO_2 in Air for Uniform and Variable Diffusion Coefficient.

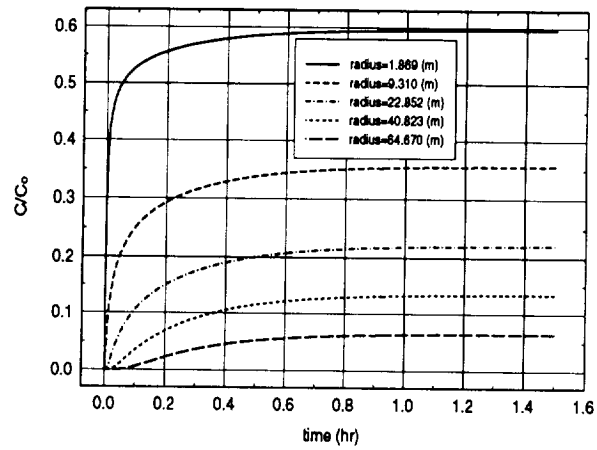


Figure 4. Transient Diffusion of CO_2 in Air for Uniform Difusion Coefficient.

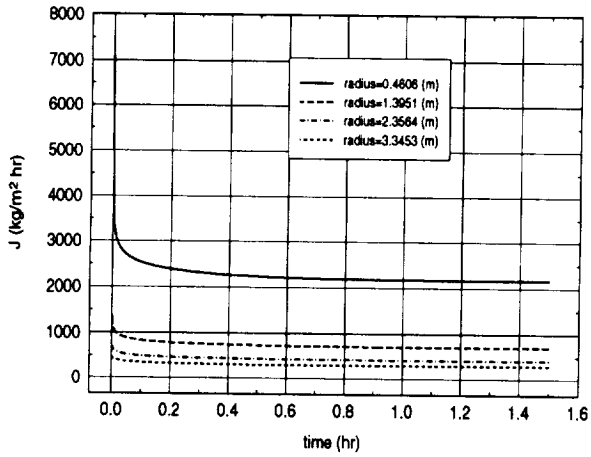


Figure 5. Transient Diffusion Flux for CO_2 in Air for Uniform Diffusion Coefficient.

SOME QUEUING MODELS IN DISCRETE EVENT DYNAMICAL SYSTEMS

Mario Davidson and S. Sathanathan
CASS Center, Tennessee State University, Nashville, TN 37221

ABSTRACT

Queuing models play an important role in "Discrete Event Dynamical Systems" (DEDS). We have investigated several queuing models such as M |M|1, M|M|1|K and M|M|C|K models. We have explored the mathematical theory underlying these models. We will also provide examples that occur naturally in simulation of these models.

INTRODUCTION

A simple queuing system is an example of Discrete Event Systems (DES), in which the system is driven by events instead of time. The study of queues ranges from the simple single-server systems modeled as birth and death processes to the analysis of arbitrarily complex networks of queues. Queuing systems form a very broad and practically useful class of DES, especially when dealing with problems of resource sharing frequently encountered in the design and control of computer, manufacturing and communication systems, etc. Queuing theory deals with the random nature of the demand behavior of customers implies that probabilistic measures such as average delay, average throughput, and delay percentiles are required to assess the performance of such systems. Queuing theory is very useful as "Event Scheduling Schemes" in which sharing is a common resource. Such as a computer, a communication systems, manufacturing or even a telescope. In this paper, we discuss simple queuing systems, develop basic theory and some applications.

Mario Davidson is a graduating senior in Mathematics at Tennessee State University. He plans to obtain a M. S. in Mathematics. He has excelled in his studies, making the Dean's List numerous times. His interest include Statistics and Simulation.

PRELIMINARIES

In this section we provide definitions, notations, and simple basic results to specify a queuing model. There are, in general, three basic components that constitute a queuing model. They are stochastic models of the arrival and service times, structural parameters of the system such as storage capacity of a queue, the number of servers, etc. and the operating policies such as conditions under which arriving customers are accepted, preferential treatment of customers (First in First Unit... FIFO) by the server etc.

Definition 1: In a queuing system, a stochastic sequence $\{Y_1, Y_2, \ldots\}$ where Y_k is the kth inter arrival time, that is the time between the (k-1)st and kth events. The sequence $\{Y_k, k \geq 1\}$ is called the sequence of interarrival times. We will assume that Y_k's are independent identically distributed random variables. Let $Y_1, Y_2, \ldots \sim Y$ and Y follows an exponential distribution, then $E[Y] = \frac{1}{\lambda}$; where λ is the average arrival rate of customers.

Definition 2: In a queuing system, a stochastic sequence $\{Z_1, Z_2, \ldots\}$, where Z_k is the kth service time, that Z_k is a random variable that represents time required for the kth customer to be served. The sequence $\{Z_k, k \geq 1\}$is called the sequence of service times. If we assume that $\{Z_k\}$ are independent identically distributed random variables, such that $Z_1, Z_2, \ldots, \sim Z$ and Z follows an exponential distribution, then $E[Z] = \frac{1}{\mu}$, where μ is the average service rate of the server in the system.

Definition 3: The storage capacity of the queue, usually denoted by K, k=1,2,... is the # of space provided for customers in service. The number of servers usually denoted by c, above C=1,2,... For example, $k = \infty$ means infinite storage capacity, c=1 means single server etc.

The operating policies are categorized as number of classes. Scheduling policies, queuing disciplines and admission policies. In a single-class system, all customers have the same service requirements and the server treats them all equally. Queue discipline in a single element as First Come First Out (FIFO), Last in First Served (LIFO), and random order. The queuing model for a single-class system with a FIFO queuing discipline can be written as:

$$A \mid B \mid C \mid K$$

where A: the inter-arrival distribution
B: the Service time distribution
C: the number of Servers present c=1,2,....
K: the storage capacity of the queue K=1,2,...

Some simple models we plan to consider are $M \mid M \mid 1 \mid \infty$, $M \mid M \mid C \mid \infty$. and $M \mid M \mid C \mid K$.

where M | M | 1: A single server system with infinite capacity, the inter-arrival and service time are exponentially distributed. That is

$A(t) = 1 - e^{-\lambda t}$,

$B(t) = 1 - e^{-\mu t}$,

$\lambda > 0, \quad \mu > 0$.

M | M | C | K : C Server System with Capacity K.

M | M | C: C- Server system with infinite capacity.

We also adopt the following notations:

Let S_k is the arrival time of the kth customer.
D_k is the departure time of the kth customer.
W_k is the waiting time of the kth customer (from arrival until beginning of service).
T_k is the time spent in the system (including service).

N(t) is the total number of custo-mers in the system, A(t) be the number of arrivals up to time-t and D(t) is the number of customers left at time t.

We also utilize Little's formula.

$$E[N] = \lambda E[T]$$

The average arrival number of customers in the system is equal to the product of the average rate and average time spent in the system.

Definition 4: A state automation is a five-tuple

$$(E, X, \Gamma, f, X_0)$$

where E is a countable event set X is a countable state space, $\Gamma(x)$ is the set of feasible events, $f: X \times E \rightarrow X$ is a state transition function, X_0 is an initial state.

MAIN MODELS

For the M | M | 1 model, the steady state probability density of this system is given by,

$$P[N(t) = j] = (1-\rho)\rho^j, j = 0,1,2,\ldots \quad \text{where } \rho = \frac{\lambda}{\mu} < 1. \tag{1}$$

The average number of customers in the system is given by,

$$E[N] = \frac{\rho}{1-\rho} \tag{2}$$

The mean total delay in the system,

$$E[I] = \frac{1}{\mu - \lambda} \tag{3}$$

The average waiting time in queue,

$$E[W] = \left(\frac{\rho}{1-\rho}\right) \cdot \frac{1}{\mu} \tag{4}$$

The mean number in queue;

$$E[N_q] = \frac{\rho^2}{1-\rho} \tag{5}$$

The server utilization is given by

$$1 - \rho_0 = \rho = \frac{\lambda}{\mu} \tag{6}$$

For the M| M| 1| K model, the steady state probability density of the system is given by

$$P[N(t) = j] = \frac{(1-\rho)\rho^j}{1-\rho^{k+1}}, j=0,1,2,\ldots k \tag{7}$$

The average number of customers in the system is given by

$$E[N] = \begin{cases} \frac{\rho}{\rho+1} - \frac{(k+1)\rho^{k+1}}{1-\rho^{k+1}}, \text{for } \rho \neq 1 \\ \frac{k}{2}, \text{for } \rho = 1 \end{cases} \tag{8}$$

For the M| M| C model, the steady state probability density of the system is given by

$$P[N(t) = j] = \begin{cases} \prod_0 \frac{(c\rho)^n}{n!}, n = 1,2,\ldots,c-1 \\ \prod_0 \frac{\rho^n}{c!}\rho^n, n = c, c+1\ldots \end{cases} \tag{9}$$

where $\rho = \frac{\lambda}{c\mu}$ and $\prod_0$ can be obtained using the normalization condition. Let $B \varepsilon \{0,1,2,\ldots m\}$ denotes the number of busy servers, then

$$E[B] = \frac{\lambda}{\mu} \tag{10}$$

The average system length,

$$E[N] = c\rho + \frac{(c\rho)^c}{c!}\frac{\rho}{(1-\rho)^2}\prod_0 \tag{11}$$

Let P_Q denotes the probability that an arriving customer does not find an idle customer and therefore forced to wait in the queue. Then,

$$P_Q = \frac{(c\rho)^c}{c!} \cdot \frac{\pi_0}{1-\rho} \tag{12}$$

Example: Queuing systems form one of the most useful classes of DES. In this example, we will show how this works for a simple queuing system.

For a simple queuing system with events a customer arrival, d: customer departure. A state automation model is given by

$\lambda E = \{a, d\}. X = \{0, 1, 2, \ldots\}$

$\Gamma(x) = \{a, d\}$, for all $x > 0, \Gamma(0) = \{a\}$

$f(x, a) = x + 1$, for all $x \geq 1$

$f(x, d) = x - 1$, if $x > 0$.

Here the state variable x represents the number of customers in the system. If we look at the state of the server rather than the system, I-Server idle, B-Server busy, D- Server breaks down, α: Server starts, β; service completes, λ :server breaks down and μ : server repaired. A state automation model is given by $E = \{\alpha, \beta, \lambda, \mu\}$, $X = \{I, B, D\}$,

$\Gamma(I) = \{\alpha\}$, $f(I, \alpha) = B$, $\Gamma(B) = \{\beta, \lambda\}$, ,

$f(B, \beta) = I$, $f(B, \lambda) = D$, $\Gamma(D) = \{\mu\}$,

$f(D, \mu) = I$.

References

1. Christos, G. Cassandras, " Discrete Event Systems", Modeling and Performance Analysis, 1993, Aksen Associates, U. S. A.

2. Sheldon M. Ross, "Introduction to Probability Models," Academic Press, 1993.

3. Howard M. Taylor and Samuel Karlin,"An Introduction to Stochastic Modeling," Academic Press, 1994.

STRUCTURAL PARAMETER CHANGES DUE TO ELEMENT DAMAGE

Kevin L. Joyce and Jiann-Shiun Lew
NASA Center for Automated Space Science
Tennessee State University
Nashville, Tennessee 37203

ABSTRACT

This paper presents a novel study of the characteristics of the parameter changes due to the damage of structural components. First, the system models for various damaged conditions are obtained by applying finite element techniques. Then a coherence approach is used to study the characteristics of the parameter changes of the transfer function due to the damage of structural components. A nine bay truss example is used to demonstrate the results of this research.

Keywords: Damage Detection, Finite Element Model, Parameter Change.

INTRODUCTION

The operation of a large space structure is sensitive to the damage of the structural components. To maintain the performance and safe operation of the system, it is necessary to remotely monitor the mechanical health of the structure.

In the last decade different approaches[1-3] have been proposed for damage detection. Much research has focused on the finite element model (FEM) update techniques. In general FEM update techniques require a large number of sensors to measure mode shapes accurately. But the number of sensors is limited in practical applications. Recently Lew[1] has proposed a coherence approach based on the fact that the system transfer function parameter change due to damage is uniquely determined by the damage type and location. Only a few sensors are required for this approach. This paper presents a novel study of the characteristics of parameter changes due to structural damage by applying the developed coherence technique. A nine-bay truss example is used to demonstrate the results.

Kevin L. Joyce is a junior student in Electrical and Computer Engineering at Tennessee State University. He plans to finish his Bachelor degree in 1997. He was with AT & T for his internship during the summer of 1995. He is currently a research student in the Center for Automated Space Science. His interests include communication and system identification for damage detection.

MODELING

Finite element analysis has been recently widely applied in the space and automobile industries for the computer solution of the dynamic problem of structures. The finite element of a nine-bay truss structure as shown in Figure 1 is chosen as the system for the study. The truss is cantilevered to the ground at the left end. This structure has 9 bays with 45 bar elements, and each nondiagonal element is a 3-m-long aluminum tube with cross-sectional area of 1.38×10^{-4} m^2. The vibration of a bar element, as shown in Figure 2, along the axial direction is governed by the partial differential equation[4]

$$\frac{\partial^2 u}{\partial t^2} = \frac{E}{\rho}\frac{\partial^2 u}{\partial x^2} \tag{1}$$

where u is the displacement at x, E is Young's modulus, and ρ is mass density. By applying the finite element analysis[4], we can obtain the system dynamic equation

$$M\ddot{q} + Kq = Bf \tag{2}$$

where M is the mass matrix, K is the stiffness matrix, B is the input matrix, q is the generalised coordinate vector, and f is the input force. The displacement output measurement can be written as

$$y = Cq \tag{3}$$

Taking the Laplace Transformation of Eq. (2) yields

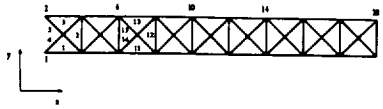


Figure 1: Nine bay truss structure.

Figure 2: Displacement of bar element.

$$(Ms^2 + K)q(s) = Bf(s) \tag{4}$$

From Eqs. (3) and (4), we can obtain the transfer function

$$G(s) = \frac{y(s)}{f(s)} = C(s^2M - K)^{-1}B \tag{5}$$

In this paper, we consider the system with one input at node 20 and four displacement outputs at nodes 20, 14, 10, and 6. All of the outputs are in the y direction and the input is in the z direction. The transfer function of this system can be written as

$$G(s) = [g_1(s)\ g_2(s)\ g_3(s)\ g_4(s)]^T \tag{6}$$

The transfer functions with the first k modes can be written as

$$g_i(s) = \sum_{l=1}^{k} \frac{b_{il}}{s^2 + \alpha_l}, \quad i = 1, 2, 3, 4 \tag{7}$$

where α_l is the square of the lth natural frequency, and b_{il} is the amplitude related to the lth mode for the ith output.

COHERENCE APPROACH

We will give a brief description of the coherence approach[1] used to study the characteristics of the parameter changes due to structural damages. To present this approach, we consider the preceding nine-bay truss structure. We only consider the first 4 modes and their natural frequencies are listed in Table 1.

The parameter vectors of the healthy structure are defined as

$$\mathbf{p}_0 = [a_1\ a_2\ a_3\ a_4]^T$$

Table 1: Natural Frequency of Nine-Bay Structure

	Mode	HZ
1	Bending	4.986
2	Bending	27.717
3	Torsion	39.787
4	Bending	67.434

$$\mathbf{p}_i = [b_{i1}\ b_{i2}\ b_{i3}\ b_{i4}]^T, \quad i = 1, 2, 3, 4 \tag{8}$$

where $\mathbf{p}_0$ is related to natural frequencies, and $\mathbf{p}_i$ is related to the amplitudes of the ith output. Here we consider the damage case when one beam is broken. The changes of the parameter vectors due to the damage of the ith element are defined as

$$\Delta\mathbf{p}_{ij} = \mathbf{p}_{ij} - \mathbf{p}_j, \quad j = 0, 1, 2, 3, 4 \tag{9}$$

where $\mathbf{p}_{ij}$ are the corresponding parameter vectors of the system with the ith element broken. The coherences between the tested system and the system with the ith element damage are defined as

$$C_{ij} = \frac{(\Delta\mathbf{p}_j^W)^T \Delta\mathbf{p}_{ij}^W}{|\Delta\mathbf{p}_j^W||\Delta\mathbf{p}_{ij}^W|}, \quad j = 0, 1, 2, 3, 4 \tag{10}$$

where $\Delta\mathbf{p}_{ij}^W$ are weighted change vectors[1] of the ith damage case, and $\Delta\mathbf{p}_j^W$ are weighted change vectors of the tested system. The coherence C_{ij} represents the cosine between two vectors. The value of the coherence C_{ij} is between -1 and 1. When C_{ij} is less than 0, the change vector $\Delta\mathbf{p}_j^W$ of the tested system is in a different direction ($> 90^0$) from the change vector $\Delta\mathbf{p}_{ij}^W$ due to the ith element damage. It strongly infers that the ith element is not damaged. The magnitude ratios between the tested system and the system with the ith element damage are defined as

$$R_{ij} = \frac{|\Delta\mathbf{p}_j^W|}{|\Delta\mathbf{p}_{ij}^W|}, \quad j = 0, 1, 2, 3, 4 \tag{11}$$

The properties associated with C_{ij}, C_i, and R_{ij} can be found in [1]. Examples will be used to demonstrate how to apply these parameters for damage detection.

RESULTS

We will apply the approach in the last section to the nine-bay truss structure to study the characteristics of the parameter changes due to structural damage. In this paper, we consider the damage condition where one beam loses stiffness. When a beam loses 100% stiffness, this beam is broken. Here we consider different levels of stiffness loss from 20% to 100% with an increment of 20%.

Table 2: Changes of α_i due to Beam Broken

	Healthy	Beam 11	Beam 12	Beam 15
α_1	9.816e02	-4.347e02	0	-1.010e01
α_2	3.033e04	-2.659e03	0	-1.791e03
α_3	6.249e04	-1.736e04	-3.898e03	-1.719e03
α_4	1.795e05	-2.768e04	0	-6.572e02

Table 2 shows the changes of α_i when beam 11, 12, or 15 is broken.

In Table 2, the results corresponding to the healthy structure are α_i and the results corresponding to each broken beam are $\Delta\alpha_i$, the parameter changes. Because of the symmetry of this structure, the parameter changes $\Delta\alpha_i$ due to the failure of beam 11 and beam 13 are the same; $\Delta\alpha_i$ due to the failure of beam 14 and beam 15 are the same. All variables $\Delta\alpha_i$ are negative, and this coincides with the fact that the system natural frequencies reduce due to beam failure. The parameter changes $\Delta\alpha_i$ due to the failure of beam 11 are significant, so the system natural frequencies reduce significantly due to this beam failure. Especially, the variable α_1 related to the first mode reduces 44.3% due to beam 11 failure. The failure of beam 12 only affects the third mode, which is a torsional mode. The failure of beam 15 has much less effect on the structure than the damage of beam 11.

The next example is used to demonstrate the coherence approach. Figure 3 shows the coherence C_{i0}, which is related to the change of natural frequencies, for the tested system for which beam 11 is broken. As expected, both C_{110} and C_{130} are 1. Table 3 shows the results of the coherence C_{ij} and the magnitude ratio R_{ij} for the cases with high C_{i0} as shown in Figure 3. From the results, the following observations are noted.

1. From Table 3, the coherence C_{130}, which is related to the change of natural frequencies, is 1. However, the coherences $C_{13j}, j = 1, 2, 3, 4$, which are related to the transfer function amplitude, are all less than 0. This implies that the change vector $\Delta\mathbf{p}^W_{13j}$ related to the amplitude of the transfer function due to the failure of beam 13 is totally in a different direction from $\Delta\mathbf{p}^W_j$ of the tested system. This implies that the tested system is not the system with beam 13 broken.

2. Figure 3 shows that both C_{60} and C_{80} are close to 1. From Table 3, the minimum coherences between the tested system and the system with beam 6 and beam 8 failure are 0.198 and -0.354, respectively. These low coherences show that neither beam 6 nor beam 8 is broken.

3. All the parameters C_{11j} and R_{11j} are 1. The tested system is identified as the system with beam 11 broken.

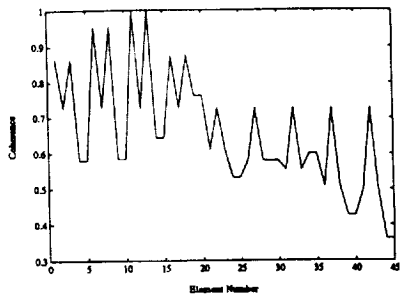


Figure 3: Coherence C_{i0} for the beam 11 broken case.

Table 3: Coherence and Magnitude Ratio

C_{ij}	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$
$i=11$	1.000	1.000	1.000	1.000	1.000
$i=13$	1.000	-0.943	-0.954	-0.891	-0.464
$i=6$	0.954	0.198	0.302	0.319	0.710
$i=8$	0.954	0.054	0.011	-0.137	-0.354
R_{ij}	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$
$i=11$	1.000	1.000	1.000	1.000	1.000
$i=13$	1.000	1.355	1.394	1.706	1.643
$i=6$	0.987	1.497	1.435	1.329	1.288
$i=8$	0.987	2.358	1.810	1.414	1.626

Table 4: Coherence Related to Natural Frequency

C_{i0}	$i=11$	$i=15$
100%	1.0000	1.0000
80%	0.9924	0.9998
60%	0.9856	0.9996
40%	0.9812	0.9994
20%	0.9781	0.9992

Next we will discuss the results for the structure with different levels of stiffness loss. Table 4 shows the coherence C_{i0} for the tested systems with different levels of stiffness loss. The results in Table 4 are the coherences between the tested system with stiffness loss and the system with the beam broken, 100% stiffness loss, for the same beam element i. When beam 11 loses stiffness from 20% to 80%, the change vector $\Delta\mathbf{p}^W_0$ is almost in the same direction as $\Delta\mathbf{p}^W_{110}$ due to the failure of beam 11. The results for $i = 15$ show that the direction of $\Delta\mathbf{p}^W_0$ has negligible change when the stiffness loss changes from 20% to 100%. From Table 2, the changes of α_i due to beam 11 broken is much larger than that due to beam 15 broken. The system transfer function changes signifi-

cantly due to beam 11 broken. This causes little change of the direction of Δp_0^W when the stiffness loss changes from 20% to 100%. Tables 5 and 6 show the results of the coherence and magnitude ratio.

Table 5: Coherence Related to Amplitude

C_{ij}	$i = 11$ $j = 1$	$i = 11$ $j = 4$	$i = 15$ $j = 1$	$i = 15$ $j = 4$
100%	1.000	1.000	1.0000	1.0000
80%	0.993	0.971	1.0000	1.0000
60%	0.984	0.951	1.0000	0.9998
40%	0.978	0.941	1.0000	0.9997
20%	0.973	0.934	1.0000	0.9997

Table 6: Magnitude Ratio

R_{15j}	$j = 0$	$j = 1$	$j = 4$
100%	1.000	1.000	1.000
80%	0.561	0.561	0.567
60%	0.323	0.323	0.329
40%	0.175	0.175	0.179
20%	0.074	0.074	0.075

From the results, the following observations are noted.

1. The high coherences C_{11j} in Table 5 imply that the direction of Δp_i^W related to the transfer function amplitude for each output changes little when the stiffness loss of beam 11 changes from 20% to 100%.

2. The coherences C_{15j} in Table 5 for beam 15 with different levels of stiffness loss is almost exactly 1. The direction of Δp_i^W has negligible change. The results in Table 5 coincide with the results in Table 4.

3. From Table 6, the magnitude ratios R_{150}, related to the natural frequency, and R_{15j} $(j = 1, 4)$, related to the amplitude, for each level of stiffness loss are almost the same. The magnitude ratio is close to a quadratic function of the stiffness loss. It can be used to indicate the intensity of the stiffness loss.

4. From the results in this paper, the following procedures are proposed to locate the damage position.

(i) The coherence related to the natural frequency is less sensitive to the level of beam damage. It can be used as the first step to locate the candidates of the damage position.

(ii) The coherence related to the amplitude is then used to locate the damage position.

(iii) The magnitude ratio is used to identify and verify the damage position. It also shows the intensity of the element damage.

CONCLUSIONS

A novel study of the characteristics of the parameter changes due to structural damage has been presented. A coherence technique is used to study the characteristics of the parameter changes. The results show that the direction of the parameter change vector related to the natural frequency changes little due to different levels of damage and the coherence C_{i0} provides reliable information for damage detection. The coherence related to the amplitude can be used to locate the damaged element when several elements have high C_{i0}. Then the magnitude ratios can be used to verify the damage position and identify the intensity of damage.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Michael Busby, Director of Center of Excellence in Information Systems at Tennessee State University, for his help of this work.

REFERENCES

1. Lew, J.-S. "Using Transfer Function Parameter Changes for damage Detection of Structures," *AIAA Journal*, Vol. 31, No. 11, 1995, pp. 2189-2193.

2. Peterson, L. D., Alvin, K. F., Doebling, S. W., and Park, K. C., "Damage Detection Using Experimentally Measured Mass and Stiffness Matrices," *Proceedings of the 34th AIAA/ASME/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA, Washington, DC, 1993, pp. 1518-1528.

3. Zimmerman, D. C., and Kaouk, M., "Structural Damage Detection Using a Subspace Rotation Algorithm," *Proceedings of the 33th AIAA/ASME/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA, Washington, DC, 1992, pp. 2341-2350.

4. Craig, R. R. *Structural Dynamics: An Introduction to Computer Methods*, Wiley, New York, 1981.

Introduction to Fuzzy Logic Control With Application to Mobile Robotics

Edward Tunstel, Tanya Lippincott[1] and Mo Jamshidi
NASA Center for Autonomous Control Engineering
Department of Electrical and Computer Engineering
University of New Mexico
Albuquerque, NM 87131

ABSTRACT:
A brief introduction to fuzzy set theory and its application to control systems is provided. Fuzzy sets do not have sharp boundaries and are therefore able to represent linguistic terms which may be considered "gray" or vague. Aspects of fuzzy set theory and fuzzy logic are highlighted in order to illustrate distinct advantages, as contrasted to classical sets and logic, for use in control systems. Using a mobile robot navigation problem as an example, the synthesis of a fuzzy control system is examined.

Keywords: mobile robots, fuzzy logic control, fuzzy sets, rover, autonomy

1. INTRODUCTION

"The world is not black and white but only shades of gray." In 1965, Zadeh [1] wrote a seminal paper in which he introduced fuzzy sets, sets with unsharp boundaries. These sets are considered gray areas rather than black and white in contrast to classical sets which form the basis of binary or Boolean logic. Fuzzy set theory and fuzzy logic are convenient tools for handling uncertain, imprecise, or unmodeled data in intelligent decision-making systems. It has also found many applications in the areas of information sciences and control systems.

In this paper, fundamental concepts of fuzzy sets and logic are briefly presented. Its utility for synthesis of control systems is discussed in the context of an application to mobile robot motion control. In mobile robotics, a fuzzy logic based control system has the advantage that it allows the intuitive nature of collision-free navigation to be easily modeled using linguistic terminology. Due to the relative computational simplicity of fuzzy rule-based systems, intelligent decisions can be made in real-time, thus allowing for uninterrupted robot motion. Moreover, accurate (expensive) sensors and detailed models of the environment are not absolutely necessary for autonomous navigation [2].

[1] B.S. degree in electrical engineering, University of New Mexico. Currently pursuing M.S. degree, electrical engineering, with an expected graduation date of May, 1997; Networks and Control Systems concentration with emphasis on fuzzy control of mobile robots.

2. FUZZY SET THEORY

In classical set theory a set, C, is comprised of elements, $x \in U$, whose membership in C is described by the characteristic, or membership function

$$\mu_C(x) : U \rightarrow \{0,1\} \tag{1}$$

where U is the universe of discourse, a collection of elements that can be continuous or discrete. The membership function $\mu_C(x)$ implies that the element x either belongs to the set ($\mu_C(x) = 1$) or it does not ($\mu_C(x) = 0$). In fuzzy set theory a fuzzy set, $\tilde{F}$, is described by the membership function

$$\mu_{\tilde{F}}(x) : U \rightarrow [0,1] \tag{2}$$

where elements, $x \in U$, have degrees of membership in $\tilde{F}$ with any value between 0 and 1 inclusive. Note that a fuzzy membership function is a so-called possibility function and not a probability function. A membership value of zero corresponds to the case where the element is definitely not a member of the fuzzy set. A membership value of one corresponds to elements with full membership in the fuzzy set. Membership values in the open interval (0, 1) correspond to partial membership and indicate a measure of uncertainty or imprecision associated with the element.

A comparative example of a crisp set and a fuzzy set can be illustrated by using the linguistic term 'far' in reference to relative distance between objects. The term 'far' can take on different meanings to different individuals, and in different contexts. For illustrative purposes, let 'far' be 2 meters (approximately 2 meters in the fuzzy set case). A graphical representation of a crisp set and a fuzzy set for 'far' is shown in Figure 1.

Membership functions can be defined as functions which take on a variety of possible shapes determined at the discretion of the fuzzy system designer. Commonly used function shapes (fuzzy logic terminology given in parentheses) include triangular (Λ), trapezoidal (Π), delta (singleton), positively sloped ramp (Γ), and negatively sloped ramp (L). These are shown in Figure 2. The ramp functions are sometimes referred to as right shoulders (Γ) and left shoulders (L).

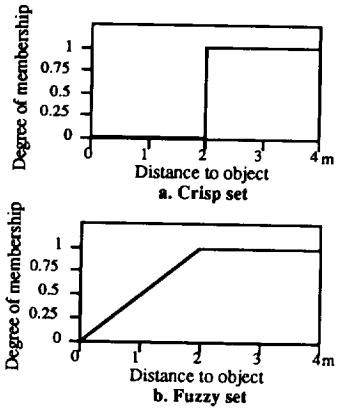


Figure 1 Graphical representations of 'far'.

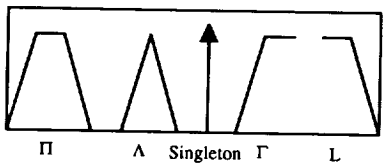


Figure 2 Common fuzzy membership functions.

Fuzzy sets, like classical crisp sets, are subject to set operations such as union, intersection, and complement [1] which are used to express logic statements or propositions. The union of two fuzzy sets $\tilde{A}$ and $\tilde{B}$ with membership functions $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$ is a fuzzy set $\tilde{C}=\tilde{A}\cup\tilde{B}$, whose membership function is related to those of $\tilde{A}$ and $\tilde{B}$ as follows:

$$\mu_{\tilde{C}}(x)=\mu_{\tilde{A}\cup\tilde{B}}(x)=\max[\mu_{\tilde{A}}(x),\mu_{\tilde{B}}(x)] \qquad (3)$$

The operator in this equation is referred to as the max-operator and is represented by the logical term OR. The intersection of $\tilde{A}$ and $\tilde{B}$ is a fuzzy set $\tilde{D}=\tilde{A}\cap\tilde{B}$ whose membership function is given by:

$$\mu_{\tilde{D}}(x)=\mu_{\tilde{A}\cap\tilde{B}}(x)=\min[\mu_{\tilde{A}}(x),\mu_{\tilde{B}}(x)] \qquad (4)$$

The operator in this equation is referred to as the min-operator represented by the logical term AND. For details on complements and other fuzzy logical operations see [1] or [3].

Consider the Cartesian product of two universes U and V defined by

$$U\times V=\{(u,v)\mid u\in U;v\in V\}$$

which combines elements of U and V in a set of ordered pairs. A fuzzy relation **R** is a mapping:

$$\mathbf{R}:U\times V\rightarrow[0,1]$$

where

$$\mu_{\mathbf{R}}(u,v)=\mu_{\tilde{A}\times\tilde{B}}(u,v)=\min[\mu_{\tilde{A}}(u),\mu_{\tilde{B}}(v)] \qquad (5)$$

The composition of two relations, R(u,v) and S(v,w), is denoted by $\mathbf{T}=\mathbf{R}\circ\mathbf{S}$. Its membership value can be determined by the following expression

$$\mu_{\mathbf{T}}(u,w)=\max[\mu_{\mathbf{R}}(u,v)\bullet\mu_{\mathbf{S}}(v,w)] \qquad (6)$$

which is called the max-product composition. Another common compositional rule of inference is the max-min composition [3].

Fuzzy relations can be represented linguistically by natural language statements in the form of fuzzy *if-then* rules. A collection of such rules is referred to as a rule-base. Accompanied by suitable membership functions, the rule-base is a core ingredient of any fuzzy rule-based expert system.

3. FUZZY LOGIC CONTROL

Fuzzy logic based controllers are expert control systems that smoothly interpolate between rules. Rules fire to continuous degrees and the multiple resultant actions are combined into an interpolated result. Processing of uncertain information and savings of energy using common-sense rules and natural language statements are the bases for fuzzy logic control. As pointed out by Lee [4], fuzzy logic controllers provide a means of transforming the linguistic control strategy based on expert knowledge into an automatic control strategy.

Fuzzy controller rule-bases typically take the form of a set of if-then rules whose antecedents ('if' parts) and consequents ('then' parts) are propositions involving fuzzy membership functions. If X and Y are input and output universes of discourse of a fuzzy controller with a rule-base of size n, the usual if-then rule takes the following form

IF x is $\tilde{A}_i$ THEN y is $\tilde{B}_i$

where x and y represent input and output fuzzy linguistic variables, respectively, and $\tilde{A}_i \in X$ and $\tilde{B}_i \in Y$ $(1 \leq i \leq n)$ are fuzzy sets representing linguistic values of x and y. Typically in robotics applications, the input x refers to sensory data and y to actuator control signals. In general, the rule antecedent consisting of the proposition "x is $\tilde{A}_i$" could be replaced by a conjunction of similar propositions; the same holds for the rule consequent "y is $\tilde{B}_i$". We can formally define a fuzzy system behavior (rule-base) as a function (B), from sensor space (S) to actuator space (A), i.e. $B: S \rightarrow A$, where the universes of discourse for S and A are such that $S \subset \Re^n$ and $A \subset \Re^n$. Embodied in this function is a fuzzy relation between fuzzy sets defined over S and fuzzy sets over A. This fuzzy relation is the actual rule-base of the fuzzy control system.

3.1 Mobile robot application

Mobile robots are typically equipped with several sensor modalities which may include range sensors, tactile/contact sensors, encoders, and vision systems. Given such sensor modalities, the usual procedure for fuzzy control synthesis consists of first defining linguistic terminology for the inputs and outputs, partitioning the sensor space and actuator space using appropriate fuzzy sets (membership functions), and formulating fuzzy rules that satisfactorily govern the desired response of the robot in all practical situations.

The subject of discussion in this paper is a mobile robot modeled after LOBOT, a custom-built robot driven by a 2-wheel differential configuration with two supporting casters. It is octagonal in shape, stands about 75 cm tall and measures about 60 cm in width. Range sensing is achieved using a layout of 16 ultrasonic transducers (arranged primarily on the front, sides, and forward-facing obliques); optical encoders on each driven wheel provide position information. Assuming a constant linear speed of 5cm/sec, we synthesize a fuzzy controller that uses four inputs and one output. The inputs are relative obstacle ranges to the front, left and right, and the angle in the direction of a designated goal location. Their respective linguistic terms are: FS (front sensor), LS (left sensor), RS (right sensor), and DIR. The individual sensor inputs are derived from pre-processed data from multiple sensors on the corresponding sides of the robot. The output is a direction in which to turn in order to satisfy avoiding obstacles and navigating to the goal. Its linguistic term is TURN-ANGLE. The range input space was partitioned based on a relevant maximum sensor range measurement of 4m, i.e. the universe of discourse for range spans the interval [0m, 4m]. The goal direction input covers a universe spanning $\pm\pi$ radians. The actuator control, or steering direction, covers $\pm\pi/2$ radians. The corresponding fuzzy sets are shown in Figure 3 where the labels of Figure 3b are listed in Table I.

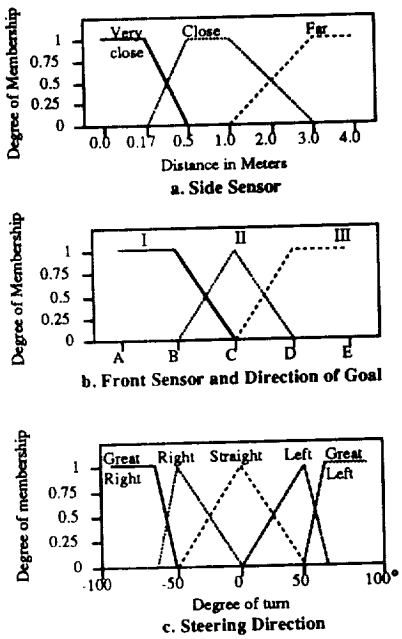


Figure 3. Membership functions

Table I.

Input	Front sensor	Goal Direction
I	Very close	Right
II	Close	Zero
III	Far	Left
A	0.0	$-\pi$
B	0.2	$-\pi/2$
C	0.3	0
D	0.75	$\pi/2$
E	1.0	π
Units	meters	radians

Based on the membership functions selected, a rule-base was designed to effect motion behavior

suitable for collision-free navigation to designated goal locations. A total of 36 rules were formulated. An example of one of these is:

IF: LS is *Close* and RS is *Close* and FS is *Far* and DIR is *Zero*

THEN: TURN-ANGLE is *Straight*.

4. SIMULATION

A simulation is described here to demonstrate the behavior of the fuzzy controller described above. It is a two-dimensional mock simulation of a Mars rover navigation task. The scenario is as follows. A planetary rover is deployed at a scientifically interesting landing site near Ares Tiu on Mars. Human operators on Earth command the rover to navigate to a designated location where experiments are to be performed. The rover's immediate task is to autonomously navigate to the goal under sensor-based control, i.e. no internal map of the environment is used.

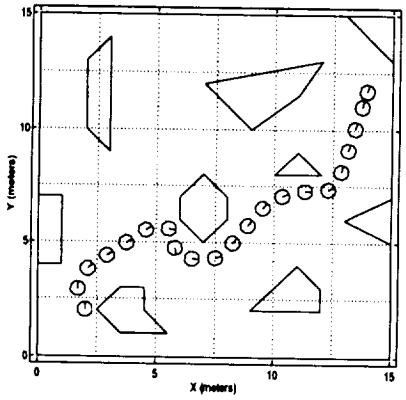


Figure 4. Fuzzy controlled sensor-based navigation.

The simulation result is illustrated in Figure 4 which shows a $225m^2$ region cluttered with an arbitrary distribution of obstacles. The simulated mobile robot (rover) is displayed as an octagon with a radial line segment indicating the robot's heading. It's initial location is $(x\ y\ \theta)$ = (2.0m 2.0m π rad.) and the goal is located at (X, Y) = (14.0m, 12.0m).

Using a fuzzy controller as developed above, the rover was able to successfully negotiate a smooth path to the goal location. In the figure, the robot icon is displayed every 10 seconds as it traverses the path.

5. SUMMARY

A brief introduction to fuzzy sets and logic was given with emphasis on its application to intelligent control of mobile robots. Insights into the synthesis procedure of such fuzzy control systems is provided via an exercise in developing a fuzzy controller for autonomous navigation. The performance of the resulting control system was demonstrated using a simulated navigation task described in the context of a simplified two-dimensional Mars rover mission scenario.

At the ACE Center, research is ongoing in the area of intelligent control of autonomous mobile robots. Focal areas include hierarchical fuzzy control, genetic programming applications to intelligent controller design, and embedded fuzzy control at the microprocessor and integrated circuit level.

REFERENCES

[1] Zadeh, L.A., "Fuzzy Sets", *Information and Control*, Vol. 12, pp. 338-353, 1965.

[2] Martinez, A., E. Tunstel and M. Jamshidi "Fuzzy Logic Based Collision Avoidance for a Mobile Robot", *Robotica*, Vol. 12 No. 6, pp. 521-527, 1994.

[3] Jamshidi, M., T. Ross and N. Vadiee (Eds.), *Fuzzy Logic and Control: Software and hardware applications*, Prentice Hall, Englewood Cliffs, NJ, 1993.

[4] Lee, C.C., "Fuzzy Logic in Control Systems: Fuzzy logic controller - Part I", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 20 No. 2, pp. 408-418, 1990.

Embedded Mobile Robot Control Using Fuzzy Logic Integrated Circuits

Edward Tunstel†, Dean E. Gonzales and Mo Jamshidi
NASA Center for Autonomous Control Engineering
Department of Electrical and Computer Engineering
University of New Mexico
Albuquerque, NM 87131

ABSTRACT:
Recent advanced VLSI technology has made robust embedded control possible for autonomous robots. For many applications, untethered intelligent mobile robots are required. State-of-the-art integrated circuit (IC) chips make it possible to apply fuzzy control in embedded systems such as micro-robots. Autonomy is further enhanced by augmenting embedded systems with pattern-matching capabilities as facilitated by neural networks. A chip-level architecture is proposed that utilizes Fuzzifier, Defuzzifier, and Fuzzy Neuron ICs as components of an embedded controller for autonomous mobile robots. It represents a general fuzzy-neuro architecture for realizing embedded intelligent control and decision systems.

Keywords: fuzzy logic, embedded control, mobile robots, microrover, integrated circuits

1. INTRODUCTION

For planetary exploration missions, NASA research is focused on the development of smaller, faster, and cheaper solutions. Smaller spacecraft and reductions in mission scope are among the goals for future missions. An example is the 1996 Mars Pathfinder mission. For planetary surface exploration, the mission will employ a microrover controlled by an embedded microprocessor-based control system. A related program is the Mars Rover Technology Program [1] which aims to develop and test enabling technologies for later, more robust missions to Mars or other planets like Venus.

At the ACE Center we are investigating control technology based on fuzzy logic for potential application to mobile robot and planetary rover navigation. In particular, we describe state-of-the-art integrated circuit chips and their potential application to embedded mobile robot control.

2. EMBEDDED FUZZY CONTROL

A controller employing fuzzy inference is referred to as a *fuzzy logic controller* [2]. Dedicated fuzzy logic controllers are implemented via software or hardware according to the design constraints of the target system. Hardware implementations are generally orders of magnitude faster their software counterparts with inference speeds on the order of microseconds as opposed to milliseconds. The reason for this is that fuzzy software controllers process information in a sequential manner, whereas fuzzy hardware processing is done in parallel. VLSI circuitry using exclusive

†B.S. and M.E. degrees in mechanical engineering, Howard University. Member of Robotic Intelligence Group at the Jet Propulsion Laboratory, Pasadena, CA since 1989. Presently a JPL Minority Fellow anticipating completion of the Ph.D. in electrical engineering, University of New Mexico, 1996. Research interests include fuzzy logic control of mobile robots, soft computing for autonomous systems, and rover technology.

fuzzy logic gates and devices have been developed at various research labs and electronics corporations. A brief survey is provided in [3]. Embedded hardware control systems are implemented in digital mode or analog mode. The fuzzy logic chips described herein were developed by Yamakawa [4] and operate in analog mode. Analog mode simplifies sensor integration, and processing analog signals directly allows for high speed data acquisition since A/D conversions are not necessary. The following section describes ICs that are potentially useful for embedded fuzzy control of small mobile vehicles.

3. FUZZY CHIPS

The components of typical fuzzy logic controllers have been fabricated in VLSI devices. Brief descriptions of several chips are provided below. For more detailed information such as circuit building blocks, photomicrographs, overviews, and response characteristics consult reference [5].

3.1 *Fuzzy Inference Chip*

Decision making based on human intuition and expert knowledge can be modeled using a fuzzy inference engine. The human knowledge is represented by defining membership functions representing the uncertainty of system variables, and a collection of fuzzy if-then rules.

The FP-9000 is a fuzzy rule chip capable of inference speeds of more than one Mega fuzzy logic inferences per second (MFLIP). Its integrated circuitry consists of membership function circuits and fuzzy *min* and *max* circuits. The rule chip includes four fuzzy inference engines, each of which achieves one fuzzy inference. Each inference engine is characterized by a fuzzy if-then rule, so that a given rule chip processes four rules simultaneously. The footprint area for the FP-9000 covers less than 2.25 cm^2.

3.2 *Defuzzifier Chip*

A fuzzy logic controller delivers a deterministic signal to the system under control. The procedure for calculating a deterministic value from a fuzzy set resulting from rule-base inference is commonly known as *defuzzification*.

The FP-9001 is a defuzzifier chip which processes a fuzzy set output by the rule chip(s) by converting it to a real number (analog voltage) necessary for input to a controlled system. The chip is designed to perform defuzzification by the center-of-gravity method. The defuzzified value must be a deterministic value because it drives electronic actuators such as motors. The footprint for the FP-9001 covers less than 1.5 cm^2.

Using the rule and defuzzifier chips, chip-level fuzzy control systems can be realized with response times of less than 2μs, independent of the number of rules employed.

3.3. *Fuzzy Neuron Chip*

Embedded fuzzy controllers that involve pattern recognition in the control loop can be realized with the aid of the fuzzy neuron chip [6]. This IC implements a fuzzy neuron model in which the synaptic weights and summation of conventional neuron models is modified respectively by fuzzy membership functions and the *min* operation of fuzzy set theory. The output of the fuzzy neuron is a value between 0 and 1

representing the degree or grade of pattern matching. Although the chips are intended for high speed pattern recognition, control applications can be achieved by using the fuzzy neuron chips in cascade with rule and defuzzifier chips as described below.

4. ARCHITECTURE

The fuzzy IC's can be used to realize an autonomous micro-robot controller capable of exhibiting goal-oriented navigation. Potential applications include tasks such as collision-free navigation in cluttered spaces in search of designated goals. In the case of rover mission like Mars Pathfinder, a goal may be a desired soil sampling site or waypoints leading to locations where experiments will be performed [7].

The fuzzy neuron chip can be used a sensor fusion device that allows the robot to recognize meaningful sensor patterns that might describe a number of obstacle configurations, hazards, or other situations to be assessed by the motion control system. Sensor patterns would consist of readings from subsets of the robot's available sensor suite. In the case of the Pathfinder microrover, these may include rate gyros, optical range sensors, encoders, and various other sensors. For common mobile robots, sensors include ultrasonic rangers, infrared proximity sensors, and encoders. The chip-level architecture being considered is shown in Figure 1.

Each fuzzy neuron can be programmed to recognize a particular situation in the robot's immediate environment. Respective degrees of pattern classification are used as input to the rule chips. Encoded rules formulated for the navigation controller map obstacle or hazard situations into a motion control command appropriate for collision-free navigation to the goal(s) at each instance.

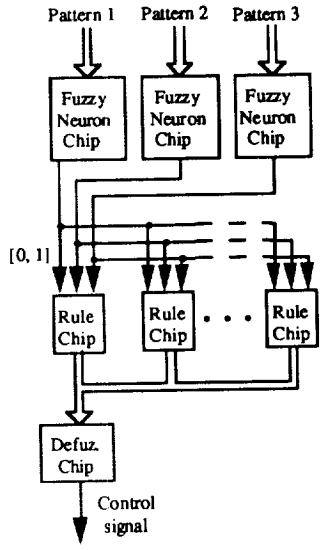


Figure 1. Fuzzy control architecture

Given the physical size of the fuzzy chips, a small circuit board can be constructed to comprise the bulk of the automatic control system for the robot, thus leaving additional space for on-board science instruments.

5. SUMMARY

Fuzzy integrated circuit devices have been described for potential use in embedded control of mobile robots. An associated chip-level architecture is proposed that may endow small robots, including planetary

microrovers, to reactively navigate in unstructured environments while performing goal-directed activities. Real-time inference speeds on the order of microseconds are possible which would facilitate sensor-based reactive/reflex control.

Future work will include real-time experimentation using actual mobile robot hardware. A control board based on the architecture described is currently in the design stage.

REFERENCES

[1] Zimmerman, W. "JPL Rover Technology Program", *5th International Symposium on Robotics and Manufacturing*, Maui, HI, pp. 345-354, 1994.

[2] Jamshidi M., N. Vadiee and T. Ross (Eds.), *Fuzzy Logic and Control -- Software and hardware applications*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[3] Tunstel, E. and M. Jamshidi, "On Embedded Fuzzy Controllers," *1st World Automation Congress*, Maui, HI, pp. 619-624, 1994.

[4] Yamakawa, T. "A Fuzzy Inference Engine in Nonlinear Analog Mode and its Application to a Fuzzy Logic Control," *IEEE Transactions on Neural Networks*, Vol. 4 No. 3, pp. 496-522, May 1993.

[5] Miki, T., H. Matsumoto, K. Ohto, and T. Yamakawa, "Silicon Implementation For a Novel High-Speed Fuzzy Inference Engine: Mega-Flips analog fuzzy processor." *Journal of Intelligent and Fuzzy Systems*, Vol. 1 No. 1, pp. 27-42, 1993.

[6] Yamakawa, T. "A Fusion of Fuzzy Logic and Neuroscience – A fuzzy neuron chip and its application to a pattern recognition", *36th Annual Conference of the Institute of Systems, Control and Information Engineers*, pp. 45-48, 1992.

[7] Matthies, L. et al "Mars Microrover Navigation: Performance evaluation and enhancement", *Autonomous Robots Journal*, Special Issue: *Autonomous Vehicles for Planetary Exploration*, to appear.

DSP IMPLEMENTATION OF EVOLUTIONARY FUZZY CONTROL

by

Mohammad-R. Akbarzadeh-T.
M. Elmer Medina[1]
&
Mo Jamshidi

NASA Center for Autonomous Control Engineering (ACE)
Dept. of Electrical and Computer Engineering
University of New Mexico, Albuquerque, NM 87131, USA
Email: emed@eece.unm.edu

ABSTRACT

Evolutionary algorithms such as Genetic Algorithms (GA) and Genetic Programs (GP) have recently been used for enhancing expert knowledge for fuzzy controllers. One issue in successful hardware implementation of such adaptive fuzzy controllers is the need for immense computational capability and dual parallel processing of the optimizing algorithm as well as the fuzzy controller. The other issue is the development of a dynamic fuzzy controller which can change its knowledge-base in real time. The Dynamic Fuzzy package developed at the University of New Mexico allows for the scaling/shifting of membership functions, creating/eliminating rules, and changing the number of rules in real-time. A Digital Signal Processing (DSP) board is used in combination with a Pentium-133 Processor to accommodate the intense computational requirements and the dual parallel processing of the GA as well as a fuzzy logic controller. In addition, a custom data acquisition system facilitates rapid parallel data transfer.

Keywords: Fuzzy Logic, Genetic Algorithms, Hardware Implementation, Digital Signal Processing

1. Introduction

Numerous simulations of evolutionary fuzzy controllers have established utility of GAs and GPs in optimizing knowledge-bases for adaptive fuzzy controllers [1,2]. Linkens and Nyongesa [3] applied GA's to adaptive fuzzy control of a non-linear multi-variable system and addressed various issues such as coding of fuzzy parameters, fitness scaling, and population size. However, real-time implementation of such adaptive controllers remains a challenging problem due to several implementation requirements such as the computational intensity, the parallel dual processing of GA and fuzzy algorithms, and the dynamic nature of adaptive fuzzy controllers. This is mainly due to the fact that a GA is strongest in a large population size. The larger the population size, the faster is the rate of convergence to an optimum solution and the higher rate of success. In real-time systems, though, the population consists of only one individual. Efforts in developing an on-line model of the system by which various individuals may be evaluated is also hampered by the computational requirements of such task. This has given rise to various approaches such as an incremental GA where only one individual is evaluated in each generation [4].

GA-based evolutionary algorithm can be suitably applied to control of flexible robots in unstructured hazardous environments. At present, a majority of robots used in industry operate in structured environments and perform slow and simple repetitive

[1] M. Elmer Medina received the B.S. degree in electrical engineering from the University of New Mexico, Albuquerque, in 1995. He is currently a masters student specializing in signal processing and communications at the department of electrical and computer engineering, UNM; with an expected graduation date of May 1997. He has worked as a research assistant at NASA's ACE Center since August 1995. His current research interest include signal processing, artificial intelligence, and image processing.

tasks that require only end-effector positioning or motion along fixed paths. However, fast and efficient operations in unstructured environments, such as in space operation and waste handling, require robots to autonomously perform more complex tasks that involve flexing of the links and joints among many other robot issues such as autonomous navigation, compliant motion, operational redundancy and fault tolerance. Moreover, these operations have to be completed in dynamic, remote, hazardous, or in general, uncertain surroundings.

Flexible robot control has already been approached using fuzzy-neuro techniques [5]. In [6], input gain tuning and direct look-up tables are used to adapt a fuzzy controller for various applications such as temperature and elevation control. The control architecture proposed here is shown in Figure 1. The fuzzy hierarchical structure, without the GA's adaptation, has shown a stable and converging behavior in [7]. When developing the hierarchical controller, some initial knowledge is expected to be supplied by the expert for feature extraction and lower level control [8]. However, because of increasing complexity of the distributed parameter system and the changing environmental conditions, it is reasonable to assume that the initial knowledge is not optimal. By incorporating genetic algorithms in the architecture, the control structure now has the capability of learning and optimizing its knowledge base.

In the following sections, different aspects of implementing evolution-based adaptive fuzzy controllers are discussed. Section 2 briefly discusses the hierarchical GA-adaptive architecture for fuzzy control of flexible robots. Various aspects of the Dynamic Fuzzy package is discussed in Section 3. Hardware aspects of digital interface and the DSP board is discussed in section 4.

2. Flexible Robot Control

The control objective for a flexible robot is to eliminate undesired oscillations while moving the robot tip through a pre-specified trajectory. In [7], a dynamic model of a single link flexible robot is derived and the above two-level hierarchical fuzzy control structure is simulated. Due to the potentially large number of spatial sensory information, such as from multiple strain gages, a pre-processing module reduces the strain gage readouts to only two variables, the average and variance of the readouts. The higher level of hierarchy then determines certain spatial features of the arm such as *Straight*, *Oscillatory*, and *Gently Curved*. The lower level of the hierarchy receives this feature determination and other low level sensory information such as the angle and velocity at the robot's tip to determine the desired torque of the motor.

A typical rule from the higher level fuzzy controller is as follows,

If *Strain Average* is Zero and *Strain Variance* is Zero,

Then,

Straight is *High*, *Oscillatory* is *Low* and *Gently Curved* is *Zero*.

A typical rule from the lower level fuzzy controller is as follows,

If *Straight* is *High*, *Oscillatory* is *Low*, *Error* in *Position* is *Zero* and *Error* in *Velocity* is *Positive*,

Then, Torque is Negative

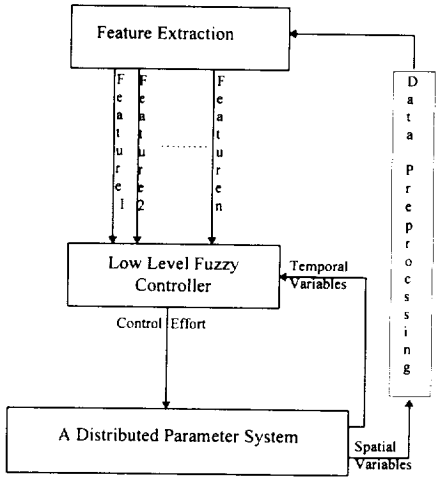


Figure 1. A block diagram of a hierarchical control system for a distributed parameter system.

3. A Dynamic Fuzzy Package

During the past decade, many commercial packages have been developed for analysis of fuzzy control systems such as the *Togai Infra-Logic* and *FIDE*. Each of these packages cater to a particular type of

need such as linear and non-linear systems and control, real-time control, simulation, among others. Some of the newly developed packages such as *O'INCA* have responded to the growing need of hybrid control. *O'INCA* is a package which allows for separate modules of neural networks and fuzzy logic to communicate sequentially or in parallel. However, none of the commercial packages has yet accounted for hybrid adaptive fuzzy control.

The program described in this paper, *Dynamic Fuzzy*, allows for adaptation parameters to be passed back and forth between an optimization algorithm and the rule-based controller through a block of memory. In our particular hardware arrangement, this memory block is located on the DSP board. The optimizing algorithm can be arbitrary as long as communication is performed through the memory block. Therefore, another fuzzy module or a module based on GA or Neural Networks can be used to modify a controller's rule set. The *Dynamic Fuzzy* is a dynamic program which allows for incorporating new rules in the rule-set while disposing of undesirable rules, changing of the number of rules, and tuning of the input and output membership functions.

The combination of DSP board and the Pentium-133's main CPU allows for simultaneous dual processing of tasks. The GA-optimizing algorithm runs on the Pentium-133 and the fuzzy controller runs on the DSP board. Since the tuning of the fuzzy controller is performed by adjusting only a memory block, the update time during each adaptation becomes minimal. In other words, the whole fuzzy routine no longer needs to be updated and compiled after each adaptation. The main routine can permanently rest in the DSP board's memory and only a small memory block will be updated each time with the new parameters.

4. Hardware Design

Some of the issues in successful hardware implementation of fuzzy logic controllers is the need for immense computational capability and dual parallel processing of the optimizing algorithm as well as of the fuzzy controller. To remedy this situation, a Digital Signal Processing (DSP) board may be used in combination with a Pentium Processor. Specifically, DSP Research's TIGER 30 board utilizes Texas Instruments TMS320C30 DSP chip and is fully compatible with a Pentium-133 Processor board. The TMS320C30 digital signal processor has a powerful instruction set, operates at 40MFLOPS, and can process data in parallel. The TIGER board has the ability of being interfaced with a PC therefore allowing dual processing of the genetic algorithm as well as the fuzzy logic controller.

A medium for data transfer between the control system and the DSP board was deemed necessary. A generic data acquisition system was designed to implement this need. The data acquisition system is generic because it is not specifically for this application; it has the capability of being used with other control systems. The capability of asynchronous and 16-bit data transfer were two features the data acquisition system needed. The data acquisition system designed consists of eight ports: four input and four output ports.

A program written in C is used to transfer data to and from the control system. The data acquisition board is assigned a memory location; the first port of the board is assigned address location 80400 with each other port offset by hex 2. The ports of the data acquisition system are accessed by either reading for writing to the mentioned address locations. Every time the memory location 8040X is accessed the output of one bit of a 3-to-8 decoder chip (SN54S138) is asserted low with the other seven bits asserted high. The location of the asserted low bit corresponds to the port selected. The asserted low bit of the 3-to-8 decoder chip is ORed with the IOSTRB signal of the TMS320C30's expansion bus using a 2-input positive -or gate (SN5432). The IOSTRB signal is asserted low when the DSP board is ready to read data. The result of the OR operation is used to select the specific port. Figure 2 is a block diagram of the circuit used to select a port.

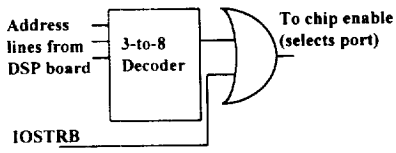


Figure 2. Block diagram of port selection circuit

The input ports of the data acquisition system consist of a pair of octal tri-state transceivers (SN54LS245) in series with a set of octal D-type flip-flop (SN54273), figure 3. The mentioned pair of chips are two chips with the enable pins wired together to accommodate 16-bit data transfer. It was necessary

to use a tri-state transceiver in order to isolate the ports from one another, as the output of the ports are all paralleled on the expansion bus. The flip-flops act as storage registers keeping the input signals from being compromised.

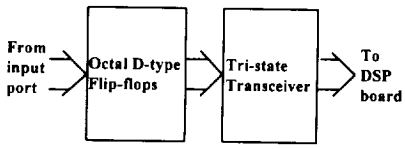


Figure 3. Block Diagram of Input Ports

The output ports of the data acquisition system are also pairs of octal tri-state transceivers (SN54LS245) with the enable pins wired together, figure 4. The input to the transceivers are connected to the output of the DSP boards expansion bus. When an output port is selected, control signals generated by the PC are transmitted from the expansion bus to the input pins of the transceiver. The transceiver enable pin is asserted high and the signal is transferred to the control system..

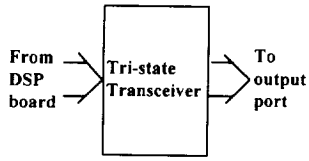


Figure 4. Block Diagram of Output Ports

5. Conclusion

Flexible robot operation in unstructured environments requires autonomous adaptability to the environmental conditions. When using fuzzy logic controllers, GA's can offer a suitable approach to this adaptation. This paper has outlined various aspects of implementing such an adaptive fuzzy controller. First, an example of a hierarchical GA-adaptive fuzzy controller is provided. The processing of both the Genetic algorithm and fuzzy controller is often too demanding on a single processor. Therefore, the hardware proposed here offers a versatile quick data acquisition system combined with a fast Digital Signal Processing board and a Pentium-133 processor to accommodate the dual processing requirement. In addition, a Dynamic Software is developed here which is compatible with the GA and allows for changing the knowledge-base in real-time. Virtually, every parameter in the knowledge base can be modified in real-time using Dynamic Fuzzy.

At present, the hardware portion and the data-acquisition of the control system has been completed. A Genetic Algorithm package is currently being debugged and will interface with Dynamic Fuzzy. Future directions of this research involves performing experiments to test the performance of the proposed routine and the DSP board.

References

[1] C. L. Karr and E. J. Gentry, "Fuzzy Control of pH Using Genetic Algorithms," *IEEE Transactions on Fuzzy Systems*, Vol. 1 No. 1, February 1993.

[2] M. A. Lee and H. Takagi, "Dynamic Control of Genetic Algorithms using Fuzzy Logic Techniques," *Proceedings of the Fifth International Conference on Genetic Algorithms*, pp. 76-83, 1993.

[3] D. A. Linkens and H. O. Nyongesa, "Genetic Algorithms for Fuzzy Control, Part 1: Offline System Development and Application," *IEEE Proc.-Control Theory Appl.*, Vol. 142, No. 3, pp. 161-176, May 1995.

[4] D. A. Linkens and H. O. Nyongesa, "Genetic Algorithms for Fuzzy Control, Part 2: Online System Development and Application," *IEE Proc.-Control Theory Appl.*, Vol. 142, No. 3, pp. 177-185, May 1995.

[5] A. Tzes, P. Peng, and F. Khorrami, "Fuzzy Neural Network Control for a Single Flexible-Link Manipulator," *Jnrl. of Intelligent & Fuzzy Systems*, 1993.

[6] S. Shenoi, K. Ashenayi, and M. Timmerman, "Implementation of a Learning Fuzzy Controller," *IEEE Control Systems*, 15 (3), pp. 73-80, 1995.

[7] Akbarzadeh-T., M.-R (1994), "A Fuzzy Hierarchical Controller For A Single Link Flexible Arm," *Proceedings of the 1994 World Automation Congress*, Maui, Hawaii, pp. 411-418.

[8] Akbarzadeh-T., M.-R., M. Jamshidi, and N. Vadiee, "A Hierarchical Fuzzy Controller Using Line-Curvature Feature Extraction for a Single Flexible Arm," *Proceedings of the 1994 IEEE Conference on Fuzzy Systems*, Orlando, Fl., 1994.

290523 P4

Adaptive Neuro-Fuzzy Controller

Kishan Kumar Kumbla, Mohammad Jamshidi and Steve Rodrieguiz
Autonomous Control Engineering
Department of Electrical and Computer Engineering
University of New Mexico
Albuquerque, NM 87131, USA

Abstract
Fuzzy logic and neural networks have been successfully utilized in design of intelligent controllers for complex systems. Such systems in real world are often operated under unstructured, unpredictable, noisy, and/or unknown environments. Such environmental conditions warrants an adaptive as well as a robust controller design to compensate for the uncertainties in the operating conditions. The enormous success in implementation of the learning ability of the neural networks has made them particularly attractive in design of adaptive fuzzy controllers. The proposed neuro-fuzzy controller monitors the behavioral pattern of the controlled variable by classifying it based on temporal data. Based on the NN output a suitable adjustment can be made in the parameter of the fuzzy controller. A fuzzy controller can be adapted by adjusting its rules, membership functions, and/or input-output gains based on the observed performance. The neuro-fuzzy controller is applied to a temperature control loop of a desalination plant process. Result show the tuning of the fuzzy controller.

Key Words: Fuzzy Logic, Neural Networks, Self-Organizing Controller.

Introduction

Fuzzy concepts have already been incorporated into neural nets in control problems, in modeling output possibility distributions, in learning and extrapolating complex relationships between antecedents and consequents of rules and in fuzzy reasoning. Integration of neural networks and fuzzy logic and its application in controls are discussed in [1-7]

In [4] neural network driven fuzzy reasoning is discussed. Fuzzy reasoning can express the qualitative aspect of human logic. There are two problems in conventional fuzzy reasoning. One is the lack of definite method to determine the membership function and the second is the lack of learning function or adaptability. These two problem inherent to fuzzy reasoning can be handled by neural networks (NN). For fuzzy reasoning, stable reasoning can always be attained despite deviations because the backbone logic is manifested as a rule in IF-THEN form. On the other hand, however, the rule cannot be expressed unless the logic is identified. Since the NN self-organizes the mapping relationship by learning, it is applicable even for unknown logical relationships. Moreover, it is capable of expressing any nonlinear relationship because it is itself nonlinear. From these differences, fuzzy reasoning is employed mainly for well-identified logic cases such as control, and the NN is used mainly for unidentified recognition rules such as pattern recognition.

Consider now the problem of designing a membership function. Even if it is not clear, the inference rule can be automatically derived from the fuzzy rule partition of learning data using the learning function of

the NN. Furthermore, an adaptive modification of the membership function is possible, because the NN has a learning function. Fuzzy controllers for most applications are made of a collection of linguistic control rules. These rules are based on the knowledge or experience of experts. In certain cases this knowledge or experience may not be available. We know that artificial neural networks are capable of learning. Therefore, this learning capability of the neural network can be made use of in designing the fuzzy controller.

The proposed neuro-fuzzy controller monitors the behavioral pattern of the controlled variable by classifying it based on temporal data. This can be for example, exponentially converging or diverging pattern, converging but oscillatory pattern, oscillatory pattern, etc. A neural networks can do this classification very efficiently and will be able to do the job in a robust manner. Based on the NN output a suitable adjustment can be made in the parameter of the fuzzy controller. A fuzzy controller can be adapted by adjusting its rules, membership functions, and/or input-output gains based on observed performance.

In this paper the adaptive fuzzy controller is simulated on a process control problem. The temperature of a complex nonlinear plant is monitored and the output of the neural network is used to adapt the fuzzy controller. A simulation of the temperature control is presented

Self-Organizing Neuro-Fuzzy Controller

The learning capability of the neural network can be made use of in designing the fuzzy controller. The *self-organizing fuzzy controller* is one such combination of a neural network and a fuzzy controller.

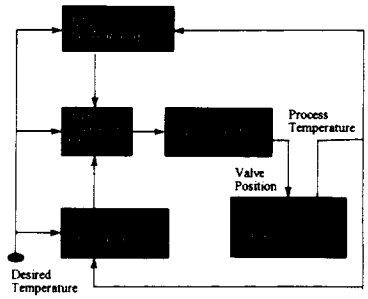


Figure 1. Schematic of Neuro-Fuzzy Controller

Figure 1 shows a schematic diagram of the system forming the self-organizing fuzzy controller. The aim of this system is to automatically form the fuzzy controller. It uses two neural networks of the *back propagation learning type*, NN1 and NN2. NN1 acts as a classifier of the dynamic responses of the object system being controlled (process control). NN2, set in judgment mechanism 2, has knowledge of the dynamic characteristics of the object system. Judgment mechanism 2 has a self tuning mechanism to automatically determine the normalizing values of the membership functions to control the object system
adequately.

In this particular case NN1 classifies the error in the temperature profile of the process to several typical patterns such as desired response or an oscillating and diverging pattern or an oscillating and slowly converging pattern or any other pattern. The result of the classification is sent to judgment mechanism 1. NN2 is made to learn the dynamic characteristics of the object system through pairs of input and system response. NN2 can then be used to

simulate the object system in cases where it is too risky to control the object system with an incomplete fuzzy controller.

The judgment mechanism 1 decides whether a next fuzzy rule has to be formed or not and whether tuning operation is needed or not for the scaling of membership functions at the currently established rules. Next the labels of the fuzzy variables in the antecedent part of the fuzzy control rule are determined from the information of the dynamic response of the object system and the rule is integrated using the fuzzy variable in the consequent part whose label is calculated in the judgment mechanism 2. If there are no rules at the initial state in the fuzzy controller, then the judgment mechanism sets up some initial control value to control the object system. The object system is controlled by the established fuzzy rules and the information of the dynamic response is sent to NN1. The self tuning operation of the normalizing values for currently established fuzzy control rules is repeated until the dynamic response is classified to desired response by NN1. If the dynamic response of the object system is classified to the desired response by NN1, the judgment mechanism 1 decides to form a next new fuzzy control rule.

The judgment mechanism 2 mainly plays a role in determining the quantities for the label of the fuzzy variable in the consequent part and in tuning values for the fuzzy variables in both the antecedent and consequent parts.

Simulation Results and Conclusions

As described above there are two neural nets NN1 and NN2 which is used to adapt the fuzzy controller. NN2 is basically used to map the input-output relation of the controlled system. Since this is a simulation, this particular NN is not used. The NN1 is the one which maps the temporal response of the system output. A multilayer perceptron is used to perform this function. O'INCA [8] software is used to develop and train this neural net. The perceptron consists of two hidden layer consisting of 10 neurons. The input layer consists of 20 inputs corresponding to the samples of the temporal data of the output. This NN has a binary output consisting of 3 layers. The binary number can map up to 8 different patterns. Initially this network is made to learn signals such as oscillating and diverging pattern, an oscillating and slowly converging pattern, an asymptotically and slowly converging pattern etc. This is done by feeding the network with standard signals.

The adaptive fuzzy control algorithm is developed in C language. The neural network developed using the O'INCA can be converted to a C code and compiled together with the adaptive fuzzy controller code.

The temperature of a process plant is used to simulate the fuzzy controller. A series of step changes in the desired temperature is performed. Figure 2a shows the response before tuning of the fuzzy controller. The tuning scaling function of the membership function *only* is performed here. Figure 2b shows the response after the fuzzy controller is adapted.

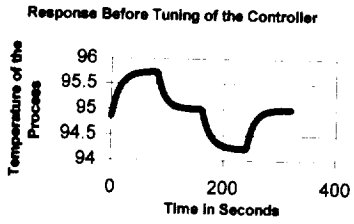


Figure 2a. Response before tuning

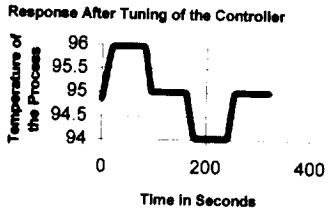


Figure 2b. Response after tuning

References

[1]S. Nakanshi, T. Takagi, K. Unehara and Y. Gotoh, "Self-Organizing Fuzzy Controllers and Neural Networks", *Proceedings of the International Conference on Fuzzy Logic & Neural Networks* ,Vol.1 pp 183-186, Japan, July 1990.

[2] Kiyohiko Uehara, Masayuki Fujise, ``Learning of Fuzzy-Inference Criteria with Artificial Neural Network", *Proceedings of the International Conference on Fuzzy Logic & Neural Networks*, Vol.1 pp 187-191, Japan, July 1990.

[3] Shin-ichi Horikawa, Takeshi Furuhashi, Shigeru Okuma and Yoshiki Uchikawa, "A Fuzzy Controller using a Neural Network and its capability to learn Expert's Control rules", *Proceedings of the International Conference on Fuzzy Logic & Neural Networks*, Vol.1 pp 103-106, Japan, July 1990.

[4] Hideyuki Takagi and Isao Hayashi, "NN-Driven Fuzzy Reasoning", *International Journal of Approximate Reasoning* 1991, 5, pp 191-212.

[5] Hisao Ishibuchi, Ryosuke Fujika and Hideo Tanaka, "Neural Networks that Learn from Fuzzy If-Then Rules", *IEEE Transactions on Fuzzy Systems*, Vol.1 No. 2, May 1993.

[6] Jyh-Shing R. Jang, "Self-Learning Fuzzy Controllers Based on Temporal Back Propagation", *IEEE Transactions on Fuzzy Systems*, Vol.1 No. 2, May 1993.

[7] Shin-ichi Horikawa, Takeshi Furuhashi, and Yoshiki Uchikawa, "On Fuzzy Modeling Using Fuzzy Neural Networks with the Back-Propagation Algorithm", *IEEE Transactions on Neural Networks*, Vol.3, No.5, September 1992.

[8] O'INCA Design Framework, User's Manual, Intelligent Machines Inc.

290524 p9

IMPROVING THE SPATIAL RESOLUTION OF SATELLITE IMAGES BY MERGING WITH NAPP IMAGES AT DIFFERENT RESOLUTIONS

[1]Venu G. Oddiraju, Tommy L. Coleman, and Wubishet Tadesse

Alabama A&M University, Normal AL 35762

ABSTRACT

The spatial resolution of Landsat MSS and TM images was improved in the past by merging with either the SPOT panchromatic band (10 m) or digitized aerial photographs having higher resolutions. However, merging at different resolutions to determine how resampling effects the spatial resolution and the statistical properties have not been studied.

In this study, two CIR NAPP photographs scanned to 1 m digital images were resampled to 15, and 8 m. When these CIR NAPP images at three different resolutions, 1, 8, and 15 m, were merged using the IHS transformation technique with corresponding Landsat-TM images, which were also resampled to the same resolutions, the spatial resolution of the satellite images at all the above resolutions was improved and the topographic shadows in the satellite images were suppressed. However, due to lack of correlation between the two data sets, the statistical properties of the satellite images were significantly changed at all resolutions after being merged with the NAPP images ($p < 0.01$).
(Key words: spatial resolution, IHS, NAPP)

INTRODUCTION

In the last few years there has been a substantial increase in the use of multispectral data from aircraft and satellite-based platforms for land use planning. This interest is mainly due to the improved spatial and spectral quality of the data (Burrough et al., 1988). The ease by which these data can be integrated with geographic databases has also provided and sparked interest in the user and scientific communities (Catlow et al., 1984; Janssen et al., 1990; and Kenk et al., 1988). However, due to the small scale of Landsat-TM satellite imagery, its usefulness to classification of forest stands is limited (Schriever and Congalton, 1993). But this problem can be alleviated, to some extent, by successful merging of low resolution satellite data with high resolution images such as SPOT panchromatic (10 m) (Chavez and Bowell, 1988) and scanned airphotos (Chavez, 1986) without changing the radiometric integrity of the original satellite data.

A digital color enhancement technique which has gained much attention as an image merging tool is the IHS (intensity, hue, saturation) transformation. The IHS transformation is a mathematical procedure that operates on digital, raster-based images. Notable virtues of the IHS transformation are its ability to effectively separate spatial (intensity) and spectral (hue and saturation) information from a standard RGB image, and its ability to convert IHS elements back to RGB display components. Siegal and Gillespie (1980) described the geometric transformation of a pixel from RGB coordinates to IHS coordinates.

In this study, scanned CIR photographs obtained as a part of National Aerial Photography Program (NAPP) were merged with the satellite images of the same geographic area at three different resolutions to improve the spatial resolution of the satellite images. The merging of these images was accomplished using an Intensity Hue Saturation (IHS) merging technique. The radiometric properties of the satellite images before and after merging were tested.

The basic objectives of this study were:
(i). To determine if the spatial resolution of a Landsat-TM image can be improved by merging with a scanned NAPP image which was poorly correlated with the satellite image.
(ii). To asses the effect of IHS merging on the large topographic shadows in the satellite image

Study Site

The Satolah quadrangle in Georgia and the Tamassee quadrangle in South Carolina were used for this study. The Tamassee quadrangle

[1] Venu Oddiraju is a Doctoral student in the Department of Plant and Soil Science. His research interests are in the applications of remote sensing and geographic information systems for studying forest vegetation, agricultural crops, soils, and wetlands. He is expected to graduate in May, 1996.

has higher relief than the Satolah image and these two quadrangles are part of the Chattooga river basin.

Data Acquisition

Color infrared photographs of the Satolah and Tamassee quadrangles, acquired as part of the National Aerial Photography Program (NAPP) (1:40,000), were purchased from USDA-ASCS, Salt Lake City, UT. The latest aerial photographs covering this area were of March 4th 1990. A Landsat-TM subscene of April 19, 1993 for the Chattooga River Basin was purchased from Earth Observation Satellite Company, Lanham, Maryland. Although the Landsat-TM image had the spectral information for differentiating various vegetation types, the spatial resolution was still coarse for a detailed classifications or for creating photo resected images which require good ground control points (GCPs). An experiment was conducted to determine if the spatial resolution of the satellite images which were poorly correlated with the NAPP images, can be improved by merging with the NAPP images using the IHS transformation procedure. Merging of these two data sets was tested at three different resolutions to evaluate the effect of resampling factors. ERDAS 7.5 and IMAGINE 8.2 (ERDAS, 1995) on Sun Sparc 10 workstation were used for all the image processing required for this study.

The IHS transformation procedure for merging scanned aerial photographs and satellite data described by Grasso (1993) was followed for this study. The NAPP CIR photographs were scanned at a resolution of 22.5 microns resulting in a resolution of approximately 1 meter.

Geometric correction

The scanned NAPP images were geometrically corrected and projected to Universal Transverse Mercator (UTM) coordinates to coincide with the satellite data. The ground control points for rectification were obtained by digitizing 7.5 minute quadrangle maps of this area using Arc/Info software. During rectification the NAPP images were resampled to different pixel sizes, 15.0 m, 8.0 m, and 1.0 m, using the nearest neighbor resampling technique. Although this technique is known to cause some blockiness, it does not alter the radiometric integrity of the images. The same resampling technique was used to adjust the pixel size of the satellite data to match with the NAPP images of the same resolution. Output images that were of slightly different sizes compared to their corresponding images were scaled in X and Y direction for proper registration.

RGB-IHS-RGB Transformation

RGB to IHS Transformation (forward transformation)

Both the NAPP (bands: 3, 2, 1) and satellite images (bands: 4, 3, 2) were orthogonally transformed into their corresponding IHS components. This step separates the spatial component (intensity) from the spectral component (hue and saturation). The intensity component contained the spatial information or the radiant brightness of all the three bands of imagery and hue and saturation contained the average color of surface reflected light.

Histogram matching

The intensity band of the NAPP image and the intensity band of the satellite image were histogram matched to compensate for changes in sun angle and atmospheric effects as these two images were obtained at different times. The resultant image, obtained as a result of histogram matching, was substituted for the intensity band of the satellite image. This step basically replaced the spatial component of the satellite data and substituted the spatial component of the NAPP image in its place.

IHS to RGB (reverse transformation)

The IHS image of the satellite, with its intensity component replaced by the histogram matched image, was transformed into its original RGB space. The RGB color composites developed in this way had the spatial resolution of the scanned NAPP images and spectral properties were comparable to the original satellite images.

RESULTS AND DISCUSSION

Improvement in Spatial Resolution of Low Resolution Satellite Image by Merging with High Resolution CIR NAPP Image

The results of this study infer that the IHS transformation is an effective technique for improving the spatial resolution of the satellite images. The spatial resolution of 1 m merged images were the closest to the spatial resolution of the original NAPP images. However, they looked blocky at a magnification of one. The nearest neighbor resampling method that was used is believed to have caused such blockiness especially since the satellite image was resampled to a pixel size that was 25-fold smaller. Similar findings were reported by

Jensen (1986) who also used this nearest neighbor method. The 8 m merged images had the spatial resolution comparable to the original NAPP images and did not have the blockiness that was observed in the 1 m image at magnification of one (Plates 1, 2, and 3). The 15 m merged image, although it did not have any blockiness at a magnification of one, lacked the spatial resolution of the 8 m image. These results are in agreement with the study conducted by Grasso (1993) in which MSS data when merged with digitized aerial photographic data resulted in a spatially improved hybrid image. Statistical properties were not tested before and after merging in their study.

Suppression of Topographic Shadows

Satellite images characteristically have large shadows that are due primarily to the topography of the area. By merging these data with the digital NAPP images, which have less shadows compared to the satellite images, the shadows in the satellite images may be eliminated. Guo and Moore (1993) using RGB-IHS-RGB transformation found that by replacing the intensity band of a color composite image by the thermal band, the cloud shadows were effectively removed and both spectral and spatial information in the shadowed areas was successfully restored. By converting a color composite image to an IHS image, the spectral variations are converted to the hue component and illumination variation caused by shadows and topography to the intensity component. Therefore, by replacing the intensity component of the satellite image with the intensity component of the digital NAPP image, topographic shadows may be suppressed.

Sabins (1978) successfully used the multispectral image band ratios for eliminating topographic shadows. However, when the input bands are highly correlated this method results in an output image that exhibits residual noise (Grasso, 1993). In this study it is observed that by merging satellite images, with large topographic shadows, with digital NAPP images using IHS transformation technique, the shadows in the original satellite image are eliminated. The intensity component of the satellite image which consists of the illumination variations due to shadows and topography, is replaced by the intensity component of the digital NAPP image which has less shadows compared to the satellite image.

Statistical Properties Before and After Merging

Although the main reason for using IHS transformation for merging two data sets was to improve the spatial resolution of the satellite images while maintaining the radiometric integrity of the original satellite image, the statistical properties significantly changed after being merged with digital NAPP images. The statistical properties such as minimum, maximum, mean, standard deviation and the t-test results of each band in both the Satolah and Tamassee images, at three different resolutions before and after merging are presented in Tables 1 and 2. The t-test analysis shows a general increase in the changes between the spectral properties of images before and after merging as the resolution increased. The t-values of the 15 m images were lower than the t-values of the 8 m images and these were inturn lower than the t-values of 1 m images. The large changes in 1 m images were attributed to a 25-fold increase in the pixel resolution compared to an approximately 3-fold increase in the case of the 8 m image and a 2-fold increase in the case of 15 m image. Only band 4 in the Tamassee image at 15 m and band 3 at 8 m shows no significant difference in their spectral properties when before and after merged images were compared (Table 2). All bands in the Satolah image at the three resolutions and all bands in Tamassee image at the 1 m resolution significantly changed after being merged with the digital NAPP images. Although the correlations between each band in an image (Tables 3a, b) are relatively higher in the 15 m Satolah image compared to the same image at 8 m and the Tamassee image at 8 and 15 m, the statistical properties of all the bands of the Satolah image are significantly changed at 8 m and at 15 m. In contrast, the statistical properties of the Tamassee image are not changed in band 3 at 8 m and band 4 at 15 m although the correlations at these bands were much lower than the correlations between different bands in the Satolah image at 15 m. This indicates that the correlations between individual bands are not the determining factors for maintaining the spectral properties of images when merged with other images. Because of this observation, another correlation was conducted between the intensity band of the satellite image and the histogram matched image (the resultant image obtained after histogram matching intensity band of the satellite image with the intensity band of the corresponding digital NAPP image). The results of this correlation show a r^2 value of 0% for the Satolah images at both 8 m and 15 m (Table 3c). The lack of correlation contributed to the significant changes in the spectral properties of these images after being merged

with corresponding digital NAPP images. In the Tamassee image, the r^2 value was 0.24% at 8 m, and at 15 m the r^2 value was 1.40%. The statistical properties of the Tamassee image were also significantly changed except for band 3 in the 8 m image and band 4 in the 15 m image. Munechika et al. (1993) reported that a strong correlation between bands of a low resolution image with the high resolution input panchromatic data is necessary for maintaining the radiometric integrity. The results of this study agree with that observation.

CONCLUSIONS

- The intensity hue saturation (IHS) merging technique improved the spatial resolution of the satellite images even when the images had a very low correlation with the scanned NAPP images.
- The 8 m satellite image, obtained by resampling the original satellite image by a factor of almost three, showed the best results compared to the merging at 1 m and 15 m.
- The statistical properties of the satellite images, however, were significantly changed after being merged with NAPP images.
- The topographic shadows in the satellite image were suppressed by merging with the scanned NAPP images at all three resolutions tested.

REFERENCES

Burrough, P. A., W. Van Deursen, and G. Heuvelink. 1988. Linking spatial process models and GIS: A marriage of convenience or a blossoming partnership? *In* GIS/LIS '88 proceedings, San Antonio, Texas, Vol. 2, pp. 598-607.

Catlow, D. R., R. J. Parsell, and B. K. Wyatt. 1984. The integrated use of digital cartographic data and remotely sensed imagery. *In* Integrated approaches in remote sensing proceedings. Guilford, UK, ESA SP. 214, pp 41-45.

Chavez, P. S., Jr. 1986. Digital merging of Landsat-TM and digitized NHAP data for 1:24000 scale image mapping. Photogrammetric Engineering and Remote Sensing 52(10): 1637-1646

Chavez, P. S., Jr., and J. A. Bowell. 1988. Comparison of the spectral information content of Landsat-TM and SPOT for three different sites in the Phoenix, Arizona, region. Photogrammetric Engineering and Remote Sensing 54(12): 1699-1708

ERDAS. 1995. Erdas image processing software, version 7.5. Earth Resource Data Analysis System, Inc., Atlanta, GA.

Grasso, D. N. 1993. Applications of the IHS color transformation for 1:24000 scale geologic mapping: A low cost SPOT alternative. Photogrammetric Engineering and Remote Sensing 59(1): 73-80

Guo, L. J., and J. M. Moore. 1993. Cloud-shadow suppression technique for enhancement of airborne thematic mapper imagery. Photogrammetric Engineering and Remote Sensing 59(8):1287-1291

Janssen, L. L. F., M. N. Jaarsma, and E. T. M. Van der Linden. 1990. Integrating topographic data with remote sensing for land cover classification. Photogrammetric Engineering and Remote Sensing 56(11): 1503-1506.

Jensen, J. R. 1986. Introductory digital image processing: A remote sensing perspective, prentice-Hall, Englewood Cliffs, New Jersey. p 379

Kenk, E. M. Sondheim, and B. Yee. 1988. Methods for improving accuracy of thematic mapper ground cover classification. Canadian J. of Remote Sens. 14 (1): 17-31.

Munechika, C. K., J. S. Warrick., C. Salvaggio, and J. R. Schott, 1993. Resolution enhancement of multispectral image data to improve classification accuracy. Photogrammetric Engineering and Remote Sensing 59(1): 67-72

Sabins, F. F., Jr. 1978. Remote sensing: Principles and interpretation, W. H. Freeman and Company, New York, pp. 258-263

Siegel, B. S., and A. R. Gillespie. 1980. Remote sensing in Geology, John Wiley and Sons, New York, p 536

Schriever, J. R., and R. G. Congalton. 1993. Evaluating seasonal variability as an aid to cover-type mapping from Landsat Thematic Mapper data in the Northeast. Photogrammetric Engineering and Remote Sensing 61(3):321-327

Table 1a. Statistical Properties of 1 meter satolah image before and after merging			
	Band 2	Band 3	Band 4
Minimum-before	000.00	000.00	000.00
Minimum-after	018.00	027.00	021.00
Maximum-before	255.00	255.00	255.00
Maximum-after	249.00	228.00	230.00
Mean-before	115.25	118.90	121.33
Mean-after	127.01	126.94	126.95
Std-before	072.24	072.06	073.48
Std-after	064.81	061.73	064.70
t-values	941.62 *	658.53 *	446.63 *
* Significantly different (P < 0.01)			

Table 1b. Statistical Properties of 8 meter satolah image before and after merging			
	Band 2	Band 3	Band 4
Minimum-before	000.00	000.00	000.00
Minimum-after	014.00	020.00	019.00
Maximum-before	255.00	255.00	255.00
Maximum-after	252.00	235.00	232.00
Mean-before	115.31	118.94	121.38
Mean-after	127.01	126.99	127.00
Std-before	071.24	072.03	073.47
Std-after	063.76	063.65	063.48
t-values	118.92 *	081.42 *	056.24 *
* Significantly different (P < 0.01)			

Table 1c. Statistical Properties of 15 meter satolah image before and after merging			
	Band 2	Band 3	Band 4
Minimum-before	000.00	000.00	000.00
Minimum-after	017.00	021.00	019.00
Maximum-before	255.00	255.00	255.00
Maximum-after	249.00	234.00	231.00
Mean-before	115.32	118.92	121.45
Mean-after	126.99	127.00	126.99
Std-before	071.21	072.07	073.45
Std-after	063.75	063.76	063.76
t-values	063.31 *	043.52 *	029.51 *
* Significantly different (P < 0.01)			

Table 2a. Statistical Properties of 1 meter tamassee image before and after merging

	Band 2	Band 3	Band 4
Minimum-b	000.00	000.00	000.00
Minimum-a	017.00	020.00	018.00
Maximum-b	255.00	255.00	255.00
Maximum-a	249.00	229.00	233.00
Mean-b	126.09	126.95	126.60
Mean-a	126.85	126.69	126.17
Standard deviation-b	073.89	074.14	074.42
Standard deviation-a	063.86	063.79	064.27
t-values	062.11 *	021.22 *	034.98 *

* Significantly different (P < 0.01)

Table 2b. Statistical Properties of 8 meter tamassee image before and after merging

	Band 2	Band 3	Band 4
Minimum-before	000.00	000.00	000.00
Minimum-after	017.00	022.00	018.00
Maximum-before	255.00	255.00	255.00
Maximum-after	251.00	229.00	233.00
Mean-before	126.54	127.05	127.21
Mean-after	127.00	127.00	126.99
Std-before	073.75	074.16	074.30
Std-after	063.89	063.98	063.47
t-values	004.70 *	000.51 NS	002.25 *

* Significantly different (P < 0.01)

Table 2c.Statistical Properties of 15 meter tamassee image before and after merging

	Band 2	Band 3	Band 4
Minimum-before	0	0	0
Minimum-after	17	22	17
Maximum-before	255	255	255
Maximum-after	250	229	234
Mean-before	126	126	127
Mean-after	126.99	127	126.99
Std-before	73.8	74.151	74.302
Std-after	63.77	63.766	63.75
t-values	5.4 *	5.44 *	0.054 NS

* Significantly different (P < 0.01)

Table 3a. Correlations between individual bands in the satellite and NAPP image of Satolah at 8 and 15 meter resolutions

Satellite Band	NAPP Band	r^2 at 8 meter	r^2 at 15 meter
2	1	2.22%	14.65%
3	2	1.45%	15.44%
4	3	0.77%	4.25%

Table 3b. Correlations between individual bands in the satellite and NAPP image of Tamassee at 8 and 15 meter resolutions

Satellite Band	NAPP Band	r^2 at 8 meter	r^2 at 15 meter
2	1	3.16%	3.66%
3	2	6.95%	9.29%
4	3	1.91%	2.41%

Table 3c.Correlations between the intensity band in the satellite image and NAPP image of both Satolah and Tamassee at 8 and 15 meter resolutions

	8 meter	15 meter
	NAPP	NAPP
(Sat.) Satolah	0%	0%
(Sat.) Tamassee	0.24%	1.40%

Plate 1.

The NAPP image of the Satolah quarter quadrangle at 8 m resolution (Bands 3, 2, 1)

Plate 2.

The Landsat-TM image of the Satolah quarter quadrangle at 8 m resolution (Bands 4, 3, 2)

Plate 3.

The Landsat-TM and NAPP merged image of the Satolah quarter quadrangle at 8 m resolution (Bands 4, 3, 2)

Monitoring Chloride Transport using Wick and Gravity-Pan Lysimeter

Robertson, G.C.*, S.A. Aburime, R.W. Taylor
Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS)
Alabama A & M University, Normal, Al. 35762

ABSTRACT

Sampling the vadose zone has proven to be a difficult task, especially in light of existing preferential flow paths. Therefore, the emphasis of this research was placed on soil/water sampling methods (wick pan lysimeter, gravity pan lysimeter) used to study the interaction of water transport phenomenon in a Decatur silt loam soil. The wick and gravity pan lysimeters were located at the 60 cm depth from the soil surface. Chemical composition of soil core samples were determined at the beginning and end of the experiment. Solute breakthrough curves were used to indicate the predominant flow paths sampled by each technique. The predominant flowpath, tracer (chloride) fractions, sampled by each sampling device was analyzed with the classical steady-state convective-dispersion equation, and the mobile-immobile two region transport equations in the CXTFIT program. The spatial variability observed with the wick and gravity lysimeters showed that the tracer was distributed randomly throughout the grids. This behavior represents a non-homogeneous flow that is present throughout the soil system.. [1]The analysis and comparison of the vadose zone sampling methods indicated that the wick lysimeter is able to sample both the unsaturated soil matrix and saturated zone, while the gravity lysimeter measured saturated or preferential flow/transport.

Keywords: Sampling, preferential flow, pesticides, groundwater, macropores.

INTRODUCTION

Modern agriculture is chemical-dependent for quality, pest control and productivity. For example, one billion (10^9) kg of more than 400 different types of pesticides are currently sprayed each year on America's cropland (Congress Watch, 1987). Because of the growing quantity of agricultural chemicals applied to soil, the potential for polluting groundwater with these chemicals has increased.

Agricultural chemicals have been identified as a source of groundwater contamination in many areas of the U.S.. Over 800 of 1,437 counties in the U.S. have been reported to show groundwater pesticide contamination (Lee and Nielsen, 1988). Agricultural chemicals are usually not considered a threat to groundwater quality. The chemicals are suppose to be either adsorbed in the soil particles, be volatilized, undergo microbial, chemical or photochemical transformation, or be taken up by plants before reaching the groundwater.

The environment is our most precious resource. Thus, it is imperative that we protect it and the precious groundwater from which 80% of quality drinking water is extracted. Therefore, basic scientific studies of how these contaminants and mixtures interact with soil, physically, chemically, microbiologically, and hydrogeologically, are required in order to understand the processes that control their fate in soils (e.g., Streile et al., 1990). Also, field-based knowledge of basic hydrological properties of soils is pertinent to understanding the physics of contaminant transport in the vadose zone affecting water quality. However, the presence of macropores and/or unstable wetting fronts, make water often flow through preferential paths and reach deep into the subsoil faster than expected (Bouma et al., 1991). By this route, soluble chemicals and pesticides may be carried deep into the subsoil before uptake or decomposition can occur.

Sampling of the vadose zone has proven to be a difficult task, especially in the light of existing preferential flow paths. Therefore, the method of sampling applied will invariably affect the accuracy of contaminant detection. In addition to determining the objectives of sampling, a thorough understanding of the environment in which the sampling equipment is to be placed is necessary to assure contaminant detection (Barcelona and Morrison, 1988). Therefore, the emphasis in this research was placed on soil/water sampling methods (wick pan lysimeter, gravity pan lysimeter) used to study the interactions of effluent transport phenomena in a North Alabama Decatur silt loam soil. Thus, the specific objectives of this

[1] **Bibliography**

*Garland Carl Robertson, son of the late Ms. Irma Jean Kendrick, was born on October 16, 1970, in Independence, Louisiana. He attended St. Helena Parish Public Schools and graduated from St. Helena High School, in Greensburg Louisiana, in May of 1988. In August, 1988, he entered Southern University and received the degree of Bachelor of Science (Plant Science) in May, 1993. Upon recieving a bachelor's degree, he participated in a Minorities Participating in Graduate Education Program (MPGEP) at Alabama A&M University (AAMU) during the summer of 1993. He then entered graduate school in August of the same year. He completed a Master's of Science (Fall 95) specializing in Soil Physics and is now preparing to enroll in the Ph.D program at AAMU.

research were to evaluate and compare the transport processes using two sampling methods (gravity and wick pan lysimeters) from the vadose zone of a predominant Alabama soil. It is hypothesized that wick lysimeter will sample soil matrix flow while the gravity lysimeter, as well as the wick lysimeter will sample preferential flow.

Materials and Methods

Description of Field Experiment

The research was performed at the Alabama A&M University (AAMU) Winfred Thomas Research Station. In 1992, eight (8) 0.2 ha experimental plots were established. The predominant soil type at the location is Decatur silt loam (Rhodic Paleudult). Each plot was spaced 20 m from each other with a lateral subsurface tile drain installed at 1 m depth. The plots were isolated from the surroundings by perimeter tiles at 1.8 m depths to intercept subsurface flow. The plots are adjacent to an underground sampling station that will not be discussed in this paper which provides an overfall where runoff measurement and sampling devices are located. Plots 1, 2, 7 and 8 will be the only plots discussed in this paper.

Vadose Zone Sampling

Methods available to collect water and solute samples from the unsaturated zone for analysis include soil cores, soil suction lysimeters, gravity pan lysimeters, wick pan lysimeters, piezometers, and tile lines. For our study, we were interested in wick and gravity pan lysimeters.

Pan lysimeters have been used in several studies (Hagerman et al., 1989; Mills and de Vries, 1988) but have been shown to be of limited utility since the soil must be saturated first before the water can flow out. Philip (1988) showed theoretically that the smaller the pan lysimeter, the larger the bypass flow in sandy soils.

Barbee and Brown (1986) introduced a new method using fiberglass wicks to sample water in the vadose zone, which eliminates many of the disadvantages of the other methods. The wick does not require that the soil be saturated before a sample is collected; the sampler is self priming, and a vacuum source is not needed. At Cornell University, a wick lysimeter that uses springs to assure good contact between the soil and wick has been developed (Boll et al., 1991). Most of these monitoring techniques will be incorporated and compared in this study. A typical layout of the plots is given in Figure 1.

Design and Installation of Sampling Devices

Eight fiber-wick and gravity pan lysimeters were installed at a depth of 60 cm as shown in Figure 3.2. A trench was dug approximately 2 m deep, 2 m wide, and 6 m long at the top end of each plot during the late spring of 1993. Horizontal shafts were located at one side of the trench, just large enough to slide the lysimeters in until they were centered under the area to be wetted, leaving approximately 50 cm between the front of the lysimeter and the trench bottom wall. A pan with twenty-five 3 cm-deep by 5 cm^2 flow-through cells in a grid pattern was constructed (Figure 2). A 13 mm hole was punched in the bottom of each cell to which 9.8 mm plastic tubing (i.d. 7.8 mm) was connected. A one meter fiber-glass wick was threaded through the tubing into the cell. A 0.04 m frayed portion of the wick covered the inside of the cell, leaving a 0.98 m segment extended down the plastic tubing (see Figure 2b)

The gravity lysimeters were constructed and installed similarly except that they do not have fiberglass wicks in their cells. In order to obtain good contact with the field soil, a mixture of coarse sand and silica powder was placed on top of the pan.

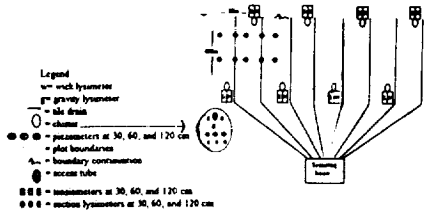


Figure 1. A Typical Instrumented Plot

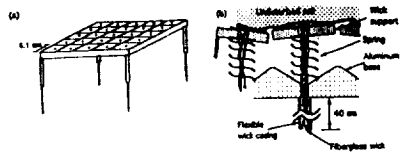


Figure 2. Components of a Typical Fiber-Wick and Gravity Lysimeter

Site Chemical and Physical Characterization

This site was characterized by collecting soil samples (disturbed and undisturbed) randomly at different locations and depths on the plots. These samples were used to determine soil physical, chemical and hydraulic properties such as soil texture,

soil porosity, volumetric water content, bulk density, and hydraulic conductivity. Standard methods as described by Black (1965) and Page (1982) were used. These data are required as input in the computer models which will be used to analyze the suitability of the sampling devices.

Schedule of Pesticide-Sampling

Data was collected more frequently during the initial period after the pesticide application and also after each rainfall event. The eight plots received natural rainfall. Effluent samples were collected with the vadose zone sampling devices according to the following schedule: a) during or one day after precipitation (rainfall), b) three days after precipitation, c) seven days after precipitation, d) ten days after precipitation, e) fourteen days after precipitation, f) or when the collection bottles contained adequate amount of sample for chemical analysis.

Tracer Analysis

Water samples were collected in pre-cleaned glass bottles and stored at 4¡C until analyzed. The soil and water analyses were conducted at the AAMU's Environmental Research Laboratory.

Chloride were chosen to measure the water's velocity because this anion move at approximately the same rate as water through negatively charged clay soils. Chloride concentration were determined with an Ion Chromatography. Diluted effluent solutions were analyzed for chloride by inserting a 5 ml sample into a Dionex series 200 ion chromatograph with an IONPAC AS4 anion exchange column and a model AMMS-1 anion exchange membrane suppressor. A high purity grade helium was used as the gas carrier in this study. Data reduction was accomplished by using the accompanying AI-400 software.

Modeling

The breakthrough curves (BTCs) from each of the sampling method were simulated with the Concentration Distance (x) Time Fit program, CXTFIT, (Parker and van Genuchten, 1984). This model is used to fit velocity, dispersivity, and adsorption partition coefficients for each location sampled. By comparing the observed concentrations with the predicted concentrations at different times, the effectiveness of each vadose zone sampling (VZS) method can be determined.

RESULTS AND DISCUSSION

Spatial Distribution and Variability

The spatial distribution of the total amount of chloride collected from the wick and gravity lysimeters in plot 1 are shown in Figure 3. The wick lysimeter indicated that most of its samples were collected from cell 3D, whereas, in the gravity lysimeter, 3,4AB, dominated the cells by collecting the most water samples. However, the amount of chloride collected with the wick lysimeter is much higher than that collected from the gravity lysimeter

The spatial distribution curves of the chloride collected from the wick and gravity lysimeters in plot 2 are shown Figure 4. Cells 4C and 2B collected the most chloride samples in the wick lysimeter shown in Figure 4. In the gravity lysimeter, the chloride was predominately sampled by cells 1,2 CD, followed by cells 1,2 AB and 1,2,3,4 EF. The lowest peaks occurred in cells 5,6 CD, and 5,6 EF in the gravity samplings.

The tracers distribution curves for plot 7 are shown in Figure 5. Chloride was detected highest from cell 2D in the wick, while and cells 3,4AB, 5,6AB and 1,2AB were larger for the gravity lysimeter.

The spatial distribution curves of the tracers collected from the wick and gravity lysimeters in plot 8 are shown in Figure 6. For the wick lysimeter, chloride was collected the most from cells 2C, 3D, and 3C. Cells 1,2 AB collected the highest for the gravity lysimeter.

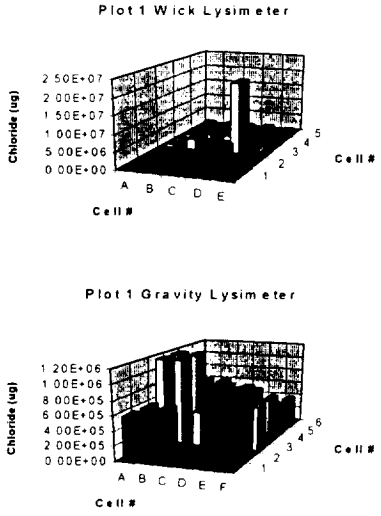
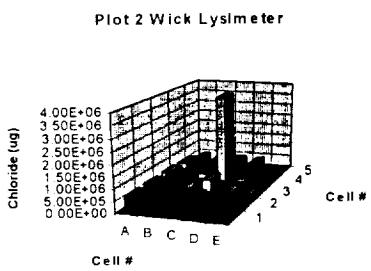
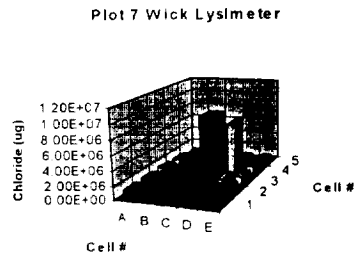
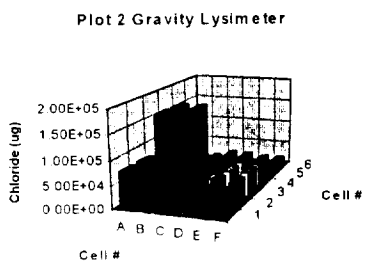


Figure 3. Spatial Distribution of the Total Amount of Chloride Collected from the Wick and Gravity Lysimeters in Plot 1.







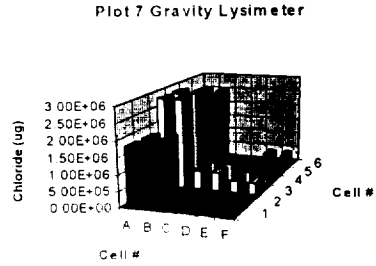


Figure 4 Spatial Distribution of the Total Amount of Chloride Collected from the Wick and Gravity Lysimeters in Plot 2.

Figure 5 Spatial Distribution of the Total Amount of Chloride Collected from the Wick and Gravity Lysimeters in Plot 7

Breakthrough Curves (BTC) of the Lysimeters Wick and Gravity

Natural rainfall was measured with rain gauges, one in each plot, placed in the vicinity of the sampling location

Breakthrough curves of the applied chemical sampled with the various sampling devices are reported and discussed as follows The combined cells chemical breakthrough curves for the wick and gravity lysimeter in plot 1 are shown in Figure 7 Breakthrough curves of the chloride (0.347 g) took place 1 day after 3 3 cm of rain from the wick lysimeter on Day 20 after chemical application (ACA). More chloride (4.99 g) was detected on Day 23 (June 10, 1994) in response to a total rainfall of 10.16 cm on June 6 and 7 Furthermore, the 4.44 cm rainfall of Day 40-42 led to the breakthrough of 22.75 g of chloride. The amount of chloride collected increase to about 1 8 g before decreasing to about zero after a total water application of 36.2 cm in 150 days. The cumulative amount of chloride over a period of 280 day of study was 42 4 g while a total of 42,683 ml (68 3 cm in equivalent depth of water) drainage was

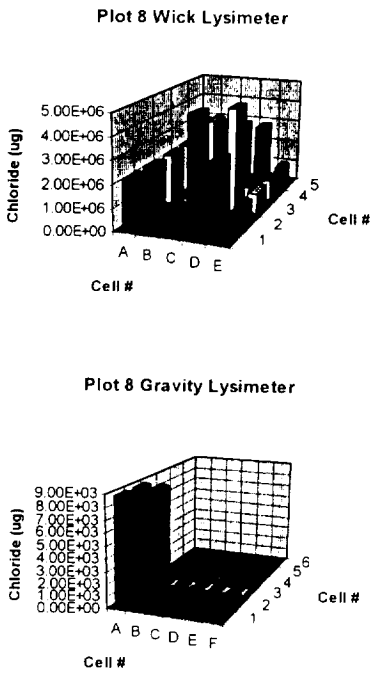


Figure 6 Spatial Distribution of the Total Amount of Chloride Collected from the Wick and Gravity Lysimeters in Plot 8.

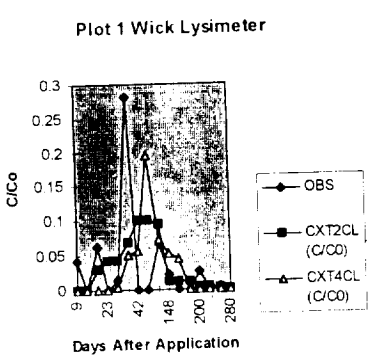
collected from a total of 81.3 cm of rainfall. As for gravity lysimeter, the chloride was measured about one day after the 10.16 cm of rainfall (June 6 & 7, 1994). After the June 6 and 7 high rainfall events, there was not enough rainfall to cause any other breakthrough in the gravity lysimeters.

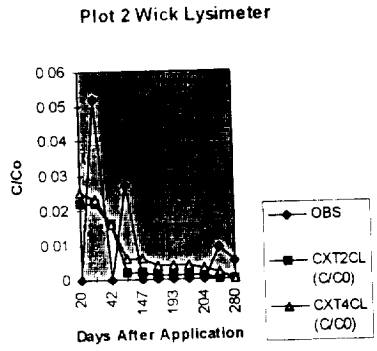
In Plot 2, the chloride in the wick lysimeter increased to 9,791.18 ug/ml on the third day following the rainfall (Day 23 ACA) from a concentration of 48.67 ug/ml measured on day 20 for the wick. Then, in response to a 5.10 cm rainfall on Day 144 there was a breakthrough of 2.21 g of chloride. Finally, on day 241 there was more chloride (0.69 g) effluent from the wick in response to the 2.79 cm rainfall that fell on day 240 ACA (Figure 8). The cumulative amount of chloride collected from the wick lysimeter over a period of 280 days of study was 7.53 g while a total of 18,958 ml (De = 30.3 cm) drainage was also collected.

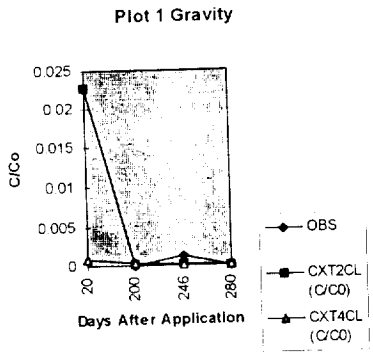
In the gravity lysimeter, the first chloride (342.9 ug/ml) peak occurred on Day 20 followed by six lesser peaks with those on Days 193 (130.2 ug/ml) and 241 (47.33 ug/ml) being more obvious. The cumulative amount of chloride over a period of 280 days of study was 1.88 g while a total of 19,507 ml (De = 21.67 cm) drainage was collected.

In Plot 7, (Figure 9),an exponential reduction of the chloride concentration in the wick lysimeter occurred in Plot 7, with the highest peak on day 23. The cumulative amount of chloride collected during the 280 day of monitoring was 40.21 g in a total of 29,329 ml (46.93 cm) of effluent. In the gravity lysimeter, four peaks of chloride were observed, 11873.47 ug/ml (33.01 g), 149.03 ug/ml (0.458 g), 28.94 ug/ml (0.024 g), and 1,595.07 ug/ml (1.50 g) on days 20, 193, 200, 280, respectively. The cumulative amount of 35 g chloride was collected from the gravity lysimeter during the 280 days of monitoring.

In plot 8, the first breakthrough of chloride (44.42 ug/ml) occurred on day 23 after a cumulative rainfall of 16.7 cm in depth. This was followed by a higher breakthrough of 133.63 ug/ml on day 144 in response to a 5.1 cm rainfall on the same day. Other breakthrough occurred on days 193, 200, 246, and 280 with the highest peak being on day 280 and a concentration of 143.94 ug/ml. the cumulative chloride measured over the 280 days of sampling was 1.24 g after a total of 15,951 ml of effluent drainage.







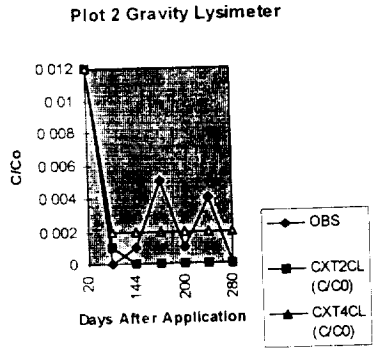


Figure 7. Predicted Chloride (Using CXTFIT Modes 2 and 4) vs. Observed breakthrough Curves for Chloride in Plot 1

Figure 8. Predicted Chloride (Using CXTFIT Modes 2 and 4) vs. Observed breakthrough Curves for Chloride in Plot 2

Comparison of Vadose Zone Sampling Methods

A comparison of sampling methods for the applied chemicals are shown in Table 1. The average chloride recovered in plot 1 was collected by the wick lysimeter with a concentration of 990.52 ug/ml, followed by the gravity lysimeter with 199.01 ug/ml of tracer collected. In plot 2, chloride was sampled best by the wick lysimeter with a concentration of 3099.56 ug/m land then by the gravity lysimeter. In plots 7 and 8, the gravity best sampled chloride followed by the wick of smaller concentrations.

Table 1. Comparison of Vadose Zone Sampling Methods

Plot #	*Chloride Concentration (ug/ml)*	
	Wick	Gravity
P1	9.91E+02	1.99E+02
P2	3.10E+03	1.08E+02
P7	2.05E+03	3.16E+03
P8	1.74E+01	7.76E+01

Soil Core Evaluation

After a total cumulative rainfall of 65.66 cm over a 280 day period, soil core samples were hydraulically augured from each plot in the vicinity of the experimental site. The data from the soil cores showed that most of the tracers still remain in the 60 cm depth of the soil except for plot 8 where chloride was in its highest concentration at the 120 cm depths of the soil. Preferential flow is a possible cause for this phenomenon to happen. In the areas of plots 7 and 8, large rocks were noted during field preparation.

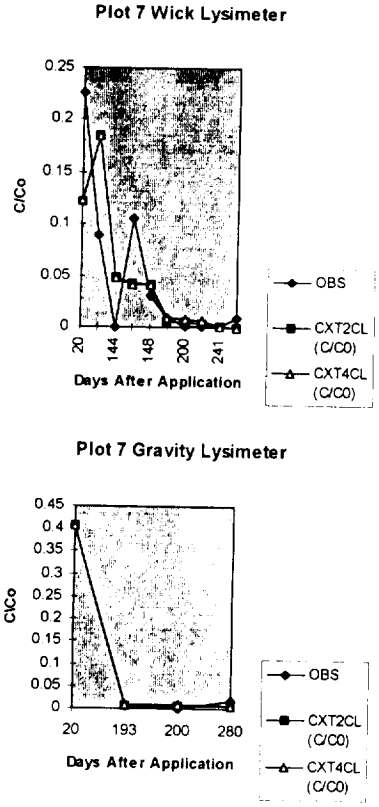


Figure 9. Predicted Chloride (Using CXTFIT Modes 2 and 4) vs. Observed breakthrough Curves for Chloride and Pesticides in Plot 7

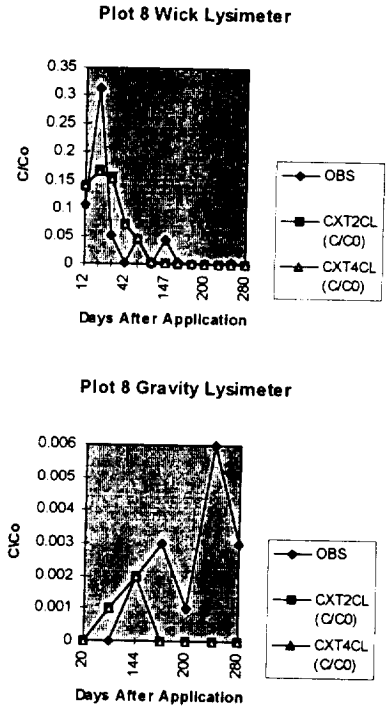


Figure 10. Predicted Chloride (Using CXTFIT Modes 2 and 4) vs. Observed breakthrough Curves for Chloride in Plot 8

Modeling of Chloride BTC's

The CXTFIT modes 2 and 4 (model I and model II, respectively) were used to predict the concentrations of chloride as a function of time and the results of the predicted were compared with the results of the observed. The breakthrough curves have been shown in Figures 7-10 for the wick and gravity lysimeters.

In plot 1 peak concentrations of the observed for the wick lysimeter (Fig. 7) arrived earlier than the predicted concentrations of modes 2 and 4. Overall, the shape of the curves indicates an under prediction of concentration by the model with model II predicting the best because of its closeness of fit. In the gravity lysimeter, model II worked best because its predicted peaks were identical to the observed. Model I obviously missed the first initial breakthrough.

In plot 2, peak arrival time of the observed was identical to the time of the models. However, concentrations were significantly different, with the observed having the highest concentration followed by model II and then by model I. Model II seem to have work the best for the model in this plot because its peaks were closest to the observed peaks (Fig. 8). The same trend was noticed in the gravity as well, except for the initial breakthrough on day 20 which was well predicted by both modes.

In plot 7, the model again worked well in the wick lysimeter (Fig. 9) with both modes slightly under predicting the highest observed concentration on day 23. In the gravity, the model worked well by showing a good to excellent fit of the model to the chloride data.

In plot 8, overall, the model seem to work well for the wick lysimeter (Fig. 10) except for on day 23 and day 144 were it under predicted the observed concentration. As for the gravity lysimeter, the model poorly fits the data from day 200 through 280, as a result, no trends were noticed from the predictions.

CONCLUSIONS

The spatial variability observed with the wick and gravity lysimeters showed that the tracer was distributed randomly throughout the grids. This behavior represents a non-homogeneous flow that is present throughout the soil system.

The analysis and comparison of the vadose zone sampling (VZS) methods indicated that the wick lysimeter is able to sample both the unsaturated soil matrix and saturated zone. Unlike the gravity lysimeter which only responds to saturated or preferential flow/transport, the wick lysimeter seems to sample both soil-water conditions.

The soil cores taking at the end of this study (280 days after application) indicates that most of the tracer that was applied was collected in the 60 cm depth of the soil for most of the plots.

The models used to predict the breakthrough curve of chloride seemed to work well by producing good to excellent fit curves when plotted against the observed for most of the plots. However, model II seem to have worked better than model I in predicting breakthrough of chloride. Model II also seemed to have worked best in predicting matrix and preferential flow observed in the wick lysimeters, while model I best predicted preferential flow which was observed in the gravity lysimeters.

Based on the performance of the models, the observed spatial distribution in the wick and gravity lysimeters, it can be stated that the hypothesis guiding this study was found to be true. The detailed analysis performed showed that the gravity lysimeters sampled water moving in macropores and preferential flow paths only during the periods of high flow, while the wick lysimeter sampled flow through the soil matrix and macropores at all times during both unsaturated (at matrix potential higher than its suction) and saturated flow conditions.

REFERENCES

1. Barbee, G. C. and K. W. Brown. 1986. Comparison Between Suction and Free-drainage Soil Suction Samplers. Soil Sci. 141: 149-154.
2. Barcelona, M.J. and R. D. Morrison. 1988. Sample Collection, Handling and Storage: Water, Soils and Aquifer Solids. In: Methods for Groundwater Quality Studies.
3. Black, C. A. (Editor). 1965. Methods of Soil Analysis. Agron. Mono. No. 9. Am. Soc. of Agron., Madison, WI.
4. Boll, J., J. S. Selker, B. M. Nijssen, T. S. Steenhuis, J. Van Winkle, and E. Jolles. 1991. Water Quality Sampling Under Preferential Flow Conditions. International Symposium on Lysimeters. Honolulu, HA. July 23-25.
5. Hagerman, J., N.B. Pickering, W.F. Ritter, and T.S. Steenhuis. 1989. In Situ Measurement of Preferential Flow. ASCE National Water Conference and Symposium. Newark, DE. July 18-20, 1989. 10 pp.
6. Lee, K.L., and E.G. Nielsen. 1988. Groundwater: Is it safe to drink? Choices Magazine. pp. 23-25
7. Mills, P.C., and M. P. de Vries. 1988. Leachate Movement through Unsaturated Sand at a Low Level Radioactive Waste Disposal Site in Northwestern Illinois. Proceedings of the Tenth Annual DOE Low Level Waste Management Conference. Aug. 30 - Sept. 1, 1988. Denver, Co. Conf. 880839. Sev. III. pp. 54.
8. Page, A. L. (Editor). 1982. Methods of Soil Analysis. Part 2: Agron. Mono. No. 9. Am. Soc. Agron., Madison, WI.
9. Parker, J.C. and M.Th. van Genuchten. 1984. Determining Transport Parameters from Laboratory and Field Tracer Experiments. Virginia Agricultural Experiment Station Bulletin 84-3.
10. Streile, G.P., J.E. Szecsody, J.F. McBride, F.J. Brockman, A.J. Valocchi, and B.E. Rittmann. 1990. Intermediate-Scale Subsurface Transport of Co-Contaminants. PNL Annual Report for 1989 to the DOE Office of Research. Part 2: March. p.39-47.

Spatial Variability of Soil Chemical Properties in Northern Alabama Soils

J. Surrency*, T. L. Coleman, W. Tadesse and A. Manu
Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS)
Alabama A&M University, Dept. of Plant and Soil Science
Normal, Alabama 35762

ABSTRACT

Knowledge of the spatial distribution of soil nutrients at the field level is essential for making reliable soil interpretations and accurate predictions of soil performance as it relates to crop yield or changes to the agroecosystem. The intent in spatial studies of soil is to describe and map soil properties over the landscape from sample data. The objectives of this study were to: (1) quantify the variability of selected soil chemical properties in Northern Alabama soils, and (2) explore the use of geostatistical methods in assessing the scale of soil variability. Soil cores were obtained from a pasture and cultivated field on the Alabama A&M University Winfred Thomas Agricultural Research Station on a 50 m grid. Cores were sampled to a depth of 30 c m and partitioned into 15 cm intervals. The soil samples were analyzed for P, K, Ca, Mg, Cu, Fe, Zn, Mn, Organic Matter, and pH. Spatial variability was investigated using semivariograms and the ratio of nugget to total semivariance expressed as a percentage.

Most of the soil nutrients were moderately spatially dependent. Soil organic matter and pH were strongly spatially dependent. There were only slight differences in the amounts and distribution of soil nutrients at the two sites studied. This could be attributed to the fact that both fields developed from the same uniform parent material and occupy similar geomorphic positions and were subjected to similar management practices. These findings provide a better understanding of the differences that exist within the fields and the adoption of management practices that will result in uniform yield.

Geostatistical techniques offer alternative methods to conventional statistics for the estimation of parameters and their associated variability. Semivariance analysis demonstrated that there were similarities in the patterns of spatial variability for some of the soil nutrients at both fields.

Additional Index Words: spatial variability, geostatistics, heterogeneity, semivariograms, soil classification.

_J. Surrency, Graduate Student, Center For Hydrology, Soil Climatology, And Remote Sensing, Alabama A&M Univ., Normal, AL 35762

Jacques Antonio Maurice Surrency, son of Donald and Lorraine Surrency, was born February 20, 1970 in Fort Valley, Ga. He attended McDuffie County Public Schools and graduated from Thomson High School, Thomson, Ga. in June 1988. He also received a Bachelor of Science degree in Agronomy / Soils in 1993 from Fort Valley State College in Fort Valley, Ga. Currently, Jacques is a graduate student majoring in Soil Science/Remote Sensing at Alabama A&M University. The title of his thesis is entitled "Spatial Variation in Soil Chemical Properties at Different Profile Depths Within Selected Toposequences in Northern Alabama" Jacques' expected graduation date is May 1996.

INTRODUCTION

A knowledge of the variability of soil properties is essential for making reliable soil interpretations and for making accurate predictions of soil performance at any particular landscape. Variability in soil properties is a particularly vexing problem for those attempting to assess either the present status or changes in ecosystems. Studies that attempt to relate soils to other ecosystem attributes, such as tree and crop productivity, must be cognizant of the potentially high variability in estimates of soil properties; this variability can affect both precision of estimates and the ability to detect true underlying relationships (Mader, 1963; Blyth and MacLeod, 1978; Grigal et al., 1991). This variability may be due to several factors, such as variation in geographic location, climate, topography, parent materials, disturbance history and biological, physical and chemical processes within the soil (Beckett and Webster, 1971). Some properties are more variable than others and the data rarely conform to normal distributions.

The integration of spectral and spatial data handling technologies are expected to provide a useful framework for locating and evaluating numerous useful characteristics of soils, such as morphological, physical, chemical, biological and mineralogical properties.

Spatial variation studies are useful tools in delineating soil classes in soil survey research. The intent in spatial studies of soil is to describe and map soil properties over the landscape from sample data (Beckett and Webster, 1971). Spatial variability of soil properties is displayed and used to predict or estimate values at unsampled locations within the regions. The spatial distribution of soils is represented on most maps by defining categorical classes and then delineating bodies of soil, which correspond to the defined classes. A given map unit may vary considerably from the central concept because of the heterogeneity caused by inclusion of other soils within the map unit. Soil classification aims to minimize the map unit variance of pr[illegible]mpared to their variance within the landscape at large (Beckett and Webster, 1971). The precision of estimation depends upon the method used for interpolation and sampling (Wilding and Drees, 1983).

Several studies have examined spatial variations of soil properties (Ball and Williams, 1968, Cameron et al., 1971; Campbell, 1978; Indorante and Jansen, 1981; Burgess and Webster, 1980; Burrough, 1983; Bos et al., 1984) and most of them have relied upon the analysis of variance methods to measure relationships between distance and variance. Spatial data analysis has been carried out using a variety of techniques, which incorporate sample locations to varing degrees in their analysis. Modern analysis of spatial variability demands geostatistics, a relatively new class of statistics designed specifically to address problems of spatial heterogeneity. The objectives of this study are: (1) to quantify the variability of selected soil chemical properties in Northern Alabama soils, and (2) to explore the use of various geostatistical methods in assessing the scale of soil variability.

MATERIALS AND METHODS

The study was conducted on the Winfred Thomas Agricultural Research Station located in Hazel Green, Alabama. Soil sample cores down to a depth of 30 cm using a 120 cm hydraulic probe lined with a plastic tubing were collected at 50 m intervals. Field 1 included 68 samples and field 2 included 33 samples on the 50 m grid. These soil profiles were characterized for their physical properties and then separated into 15 cm increments. Each of the soil profiles were examined for its color code using Munsell color charts, and for its structure and soil consistency under moist conditions. The soil samples were air-dried and sieved through a 2 mm sieve. A complete chemical analysis of the soil samples was performed using standard procedures in Methods of Soil Analysis, Part I, Klute (ed.) 1986 and Part II, A.L. Page et al (ed.) 1982. The chemical properties determined include: Phosphorus (P), Potassium (K), Calcium (Ca), Magnesium (Mg), Copper (Cu), Iron (Fe), Zinc (Zn), Manganese (Mn), Organic matter and pH. Geostatistical software (GS+, Gamma Design Software, St. Plainwell, MI) was used to analyze the spatial structure of the data and to define the semivariograms.

RESULTS AND DISCUSSION

At the 0-15 cm depth, the amounts of Cu, K and Ca extracted using the Mehlich double acid procedure were significantly higher in field 1 than in field 2 (Tables 1 and 2). No significant differences in the content of other nutrients were obtained at the 0-15 cm depth. However, at the 15-30 cm depth, field 1 had significantly higher levels of Cu, Mn, Ca and P than field 2.

Majority of the nutrients were fitted to spherical models at both depths in both fields. Two variables (Fe and Zn) in field 2 at the 0-15 cm depth were defined by Linear/Sill models. Coefficients of the data in Tables 3 and 4 is a summary of isotropic variograms for the different elements evaluated. At the 0-15 cm depth, sills for Ca and Mg were extremely low for both fields. However, K had an extremely high sill in

field 1.

The nugget variance for elements in field 1 were generally lower than those in field 2. The soil nutrients displayed differences in their spatial dependence, as determined by their semivariograms (Figures 2 and 3). The distinct classes of spatial dependence for the soil variables were defined using the criteria of Cambardella et al., 1994. If the ratio was ≤25%, the variable was considered strongly spatially dependent; if the ratio was between 25 and 75%, the variable was considered moderately spatially dependent; and if the ratio was >75%,the variable was considered weakly spatially dependent. In field 2, semivariograms indicated strong spatial dependence for variables, such as Cu, Mn, P and pH. Semivariograms for field 1 indicated strong spatial dependence for Cu. (Tables 3 and 4). Majority of the chemical parameters at both sites exhibited moderate spatial dependency (Nugget percentage of 25-75). Phosphorus in the topsoil in field 2 exhibited a strong spatial dependency (Fig. 7), however, in field 1, the same element exhibited weak spatial dependency (Fig.5). Calcium exhibited a strong spatial dependency in field 1 (Fig. 4) however, in field 2, Ca exhibited weak spatial dependency (Fig. 6).

The range values showed considerable variability among the parameters measured at each field site (Tables 3 and 4). For instance, in field 1, Fe at the 0-15 cm depth had a range of 455.7 m. At the 0-15 cm depth of field 2, Cu had a range of 201.9.

CONCLUSION

This study demonstrates that there are only slight differences in the amounts and distribution (both vertically and horizontally) of soil nutrients at the two sites studied. This could be attributed to the fact that both fields developed from the same uniform parent material and occupy similar geomorphic positions and were subjected to similar management practices. These findings provide a better understanding of the differences that exist within the fields and the adoption of the management practices that will result in uniform yield. Geostatistical techniques offer alternative methods to conventional statistics for the estimation of parameters and their associated variability. The geostatistical software GS+ provided an excellent method of analyzing the spatial structures of the data and defining the semivariograms. Semivariance analysis demonstrated that there were similarities in the patterns of spatial variability for some of the soil nutrients at both fields. This suggests that spatial relationships derived from one set of measurements for one field may have applicability at other fields within similar landscapes.

REFERENCES

Agbu, P.A., K.R. Olson. 1990. Spatial Variability of Soil Properties in Selected Illinois Mollisols. Soil Sci. 150:777-786.

Ball, D.F., and W.M. Williams. 1968. Variability of Soil Chemical Properties in Uncultivated Brown Earths. J. Soil Sci. 19:379-391.

Beckett, P.H.T., and R. Webster. 1971. Soil Variability: A Review. Soils and Fertilizers. 34:1-15.

Blyth, J.F. and D.A. MacLeod. 1978. The significance of soil variability for forest soil studies: Northeast Scotland. Journal of Soil Science. 29:419-430.

Boekhold, A.E., and S.E.A.T.M. VanderZee. 1992. Significance of soil chemical heterogeneity for spatial behavior of cadmium in field soils. Soil Sci. Soc.Am. J. 56:747-754.

Bos, J., M.E. Collins, G.J. Genesheimer and R.B. Brown. 1984. Spatial variability for one type of phosphate mine in Central Florida. Soil Sci. Soc. Am. J. 48:1120-1125.

Burgess, T.M., and R. Webster. 1980. Optimal interpolation and isarithmic mapping of soil properties. I. The semivariogram and punctual kriging. J. Soil Sci. 31:315-333.

Burrough, P.A. 1983. Multiscale sources of spatial variation in soil. I. 1988. Spatial application of fractal comcepts to nested levels of soil variation. J. Soil Sci. 34: 577-597.

Cambardella, C A., T B. Moorman, J.M. Novak, T.B. Parkin, D L. Karlen,R.F. Turco, and A.E. Konopka. 1994. Field-scale Variability of Soil Properties in Central Iowa Soils. Soil Soc. Am. J. 58:1501-1511.

Cameron, D.R., M. Nyborg, J A. Toogood, and D.H. Laverty 1971. Accuracy of field sampling for soil tests. Can. J. Soil Sci. 51:165-175.

Campbell, J.B., 1978. Spatial variation of sand content and pH within single contiguous delineations of two mapping units.

Finke, P.A., J. Bouma, and A. Stein. 1992. Measuring field variability of disturbed soils for simulation purposes. Soil Sci. Soc. Am. J. 56:187-192.

Grigal, D.F., R.E. McRoberts, and L.F. Ohmann. 1991. Spatial variation in chemical properties of forest floor and surface mineral soil in the North Central U.S. Soil Sci. 151:(4):282-290.

Huijbregts, C.H. 1973. 'In' Display and analysis of spatial data. (J.C. Davis and M.J. Mc Cullough editions) pp.38-53. Wiley, NY.

Indorante, S.J., and I.J. Jansen. 1981. Soil Variability on surface mined and undisturbed land in Southern Illinois. Soil Sci. Soc. Am. J. 45:564-568.

Klute, A. 1986. Methods of Soil Analysis. Part I. Physical and mineralogical methods. 2nd ed. Agron. Monogr. 9 ASA and SSSA, Madison, WI.

Lehrsch, G.A., F.D. Whisler and M.J.M. Romkens. 1988. Spatial variation of parameters describing soil surface roughness. Soil Sci. Soc. Am. J. 52:311-319.

Mader, D.L. 1963. Soil Variability- A seriousproblem in soil-site studies in Northeast. Soil Sci. Soc. Am. Proc. 27:707-709.

Matheron, G. 1963. Principles of Geostatistics. Econ. Geol. 58:1246-1266.

Ovalles, F.A., and M.E. Collins, 1988a and 1988b. Evaluation of soil variability in Northwest Florida using geostatistics. Soil Sci. Soc. Am. J. 52:1702-1708.

Page, A.L., R.H. Miller and D.R. Kenney. 1982. Methods of Soil Analysis. Part 2. Chemical and biological properties. 2nd ed. Agron. Monogr. 9 ASA and SSSA. Madison, WI.

Sarma, J.S., V.P. Singh and K.N.S. Sharma. 1988b. Analysis of spatial variability in sodic soils: 2. Point and Block Kriging. Soil Sci. 145:250-256.

Seyfried, M.S., A.G. Hornsby and P.V. Rao. 1992. Partioning variability of soil properties affecting solute movement with soil taxonomy. Soil Sci. Soc. Am. J. 56:207-214.

Tabor, J.A., A.W. Warrick, D.E. Myers and D.A. Pennington. 1985. Spatial variability of nitrate in irrigated cotton: Soil nitrate and correlated variables. Soil Sci. Soc. Am. J. 49:390-394.

Trangmar, B.B., R.S. Yost and G. Uehara. 1985. Application of geostatistics to spatial studies of soil properties. Adv. Agron. Academic Press, New York, NY. 38:45-94.

Van der Zaag, P., R.L Fox, R.S. Yost,B.B.Trangmar, K. Hayashi and G. Uehara. Spatial variability in selected properties of Rwanda soils. Proc. Int. Soil Class Workshop. 4th Kigali, Rwanda. 2-21 June 1981.

Warrick, A.W., D.E. Myers and D.R.Nielson. 1986. Geostatistical methods applied to soil science. pp.53-82. 'In' A. Klute (ed.) Methods of Soil analysis. Part 1. 2nd edition Agron. Monogr. 9 ASA and SSSA. Madison, WI.

Webster, R. and T.M. Burgess. 1983.Spatial variation in soil and the role of kriging. April. Water. Mgt. 6:111-122.

Webster, R. 1985. Quantitative Spatial Analysis of Soil in the field. Springer-Verlag, New York.

Wilding, L.P., and L.R. Drees. 1978.Spatial variability: A pedologists viewpoint. In diversity of soils in the tropics. Soil Sci. Soc. Am. Spec. Publ. 34:1-12.

Xu, J. and R. Webster. 1984. A geostatistical study of topsoil properties in Zhangwu County, China. Catena. 11:13-26.

Yost, R.S., G. Uehara and R.L. Fox. 1982. Geostatistical analysis of soil chemical properties of large land areas. I. Semivariograms. Soil Sci. Soc. Am. J. 46:1028-1032.

Zhang, R., D.E. Myers and A.W. Warrick. 1992. Estimation of spatial distribution of soil chemicals using pseudo-cross- variograms. Soil Sci. Soc. Am. J. 56:1444-1452.

Table 1. Mean, range, standard deviation and number of samples for soil variables at different depths for Field 1.

Variable	Depth (cm)	Mean	Range	Standard Deviation	Number of samples
Cu	0-15	0.66	0.27 - 1.86	0.10	68
	15-30	0.72	0.10 - 5.08	0.63	68
Fe	0-15	10.62	5.09 - 22.16	2.75	68
	15-30	11.23	4.49 - 27.76	11.23	68
Mn	0-15	125.8	62.30 - 306.9	40.7	68
	15-30	70.8	22.20 - 128.0	26.1	68
Zn	0-15	2.00	0.23 - 6.76	0.98	68
	15-30	1.35	0.11 - 8.44	1.38	68
K	0-15	176.2	66.11 - 483.6	89.6	65
	15-30	86.1	29.00 - 339.7	53	65
Ca	0-15	2072	318.0 - 6587	1169	68
	15-30	2088	25.0 - 8015	1281	68
Mg	0-15	251.2	47.50 - 720.0	123.9	68
	15-30	221.2	0.00 - 810.0	130.1	68
P	0-15	18.7	3.41 - 47.84	9.14	68
	15-30	8.7	0.08 - 21.78	4.99	68
pH	0-15	5.92	4.96 - 6.79	0.41	68
	15-30	5.98	4.93 - 6.78	0.44	68
OM%	0-15	2.98	2.01 - 4.07	0.52	68
	15-30	2.6	0.67 - 4.02	0.71	68

Table 2. Mean, range, standard deviation and number of samples for soil variables at different depths for Field 2.

Variable	Depth (cm)	Mean*	Range	Standard Deviation	Number of samples
Cu	0-15	0.54	0.0 - 1.50	0.27	33
	15-30	0.45	0.16 - 0.90	0.17	33
Fe	0-15	10.92	4.70 - 39.70	6.05	33
	15-30	10.47	3.96 - 23.16	3.95	33
Mn	0-15	120.60	47.30 - 201.7	39.00	33
	15-30	57.80	5.00 - 118.60	30.80	33
Zn	0-15	1.77	0.20 - 9.58	1.69	33
	15-30	3.45	0.02 - 84.0	14.28	33
K	0-15	112.80	56.8 - 177.0	48.40	30
	15-30	81.80	27.6 - 146.3	28.90	30
Ca	0-15	1472	92.0 - 5053	955	33
	15-30	1144	109.0 - 3001	668	33
Mg	0-15	246	0.0-1444	255	33
	15-30	192	63.0-1018	190	33
P	0-15	12.57	0.48 - 48.13	11.96	33
	15-30	3.64	0.08 - 16.56	4.22	33
pH	0-15	5.26	4.43 - 6.50	0.49	33
	15-30	5.25	4.30 - 6.58	0.59	33
OM%	0-15	3.17	1.83 - 4.26	0.59	33
	15-30	2.64	1.52 - 3.65	0.59	33

Mean* -Cu, Fe, Mn , Zn, K, Ca, Mg and concentrations in (mg/kg^{-1})

Table 3. Parameters for variogram models for Field 1.

Variable	Depth (cm)	Nugget Variance	Sill	% Nugget	Spatial Class	Model+	r^2
Cu	0-15	0.0124	.0783	16	S	S	0.21
	15-30	0.124	0.364	34	M	S	0.01
Fe	0-15	4.17	9.18	45	M	S	0.95
	15-30	7.03	12.10	58	M	S	0.20
Mn	0-15	749.0	1323	57	M	S	0.30
	15-30	316.0	698.8	45	M	S	0.68
Zn	0-15	0.46	0.75	61	M	S	0.01
	15-30	1.71	2.15	80	W	S	0.15
K	0-15	3510.0	7060	50	M	S	0.02
	15-30	1410	2372	60	M	S	0.08
Ca	0-15	7.14×10^{-5}	1.16×10^{-4}	--	R	S	0.03
	15-30	7.21×10^{-5}	1.29×10^{-4}	--	R	S	0.01
Mg	15-30	1.34×10^{-4}	2.15×10^{-4}	62	M	S	0.6
P	0-15	82.80	91.9	90	W	G	0.12
	15-30	11.48	20.11	57	M	S	0.08
pH	0-15	0.080	0.129	62	M	S	0.49
	15-30	0.187	0.224	83	W	S	0.79
OM	0-15	0.252	0.302	83	W	G	0.27
	15-30	0.118	.544	22	S	S	0.56

% Nugget=(Nugget variance/Sill) x100
S=Strong spatial dependency(<25%), M=Moderately spatial dependency(Between 25and 75%), W=Weak spatial dependency(>75%), R=Random, S=Spherical, L/S=Linear/Sill, G=Gaussian, Models are all isotropic.

Table 4. Parameters for variogram models for Field 2.

Variable	Depth (cm)	Nugget Variance	Sill	% Nugget	Spatial Class	Model+	r^2
Cu	0-15	0.020	0.105	19	S	S	0.97
	15-30	0.006	0.034	18	S	S	0.98
Fe	0-15	23.40	38.25	61	M	L/S	0.69
	15-30	7.82	24.88	31	M	S	0.91
Mn	0-15	1.00	1739	.06	S	S	0.92
	15-30	1.00	1068	.09	S	S	0.85
Zn	0-15	1.41	2.70	52	M	L/S	0.99
	15-30	106.4	159.6	67	M	S	0.60
K	15-30	424	750	57	M	S	0.11
Ca	0-15	4.89×10^{-3}	7.48×10^{-3}	65	M	S	0.17
	15-30	2.88×10^{-3}	4.74×10^{-3}	61	M	S	0.02
Mg	0-15	3.47×10^{-5}	4.13×10^{-4}	6	S	S	0.62
P	0-15	0.10	169.7	.06	S	S	0.91
	15-30	8.61	22.4	38	M	S	0.96
pH	0-15	0.02	0.295	7	S	S	0.97
	15-30	0.12	0.41	29	M	S	0.95
OM	0-15	0.09	0.35	26	M	S	0.91
	15-30	0.31	0.35	89	W	S	0.02

%Nugget= (Nugget Variance/ Sill) x 100
S=Strong spatial dependency (<25%), M=Moderate spatial dependency (Between 25 and 75%),
W=Weak spatial dependency (>75), R=Random, S=Spherical, L/S=Linear/sill, Models are all isotropic.

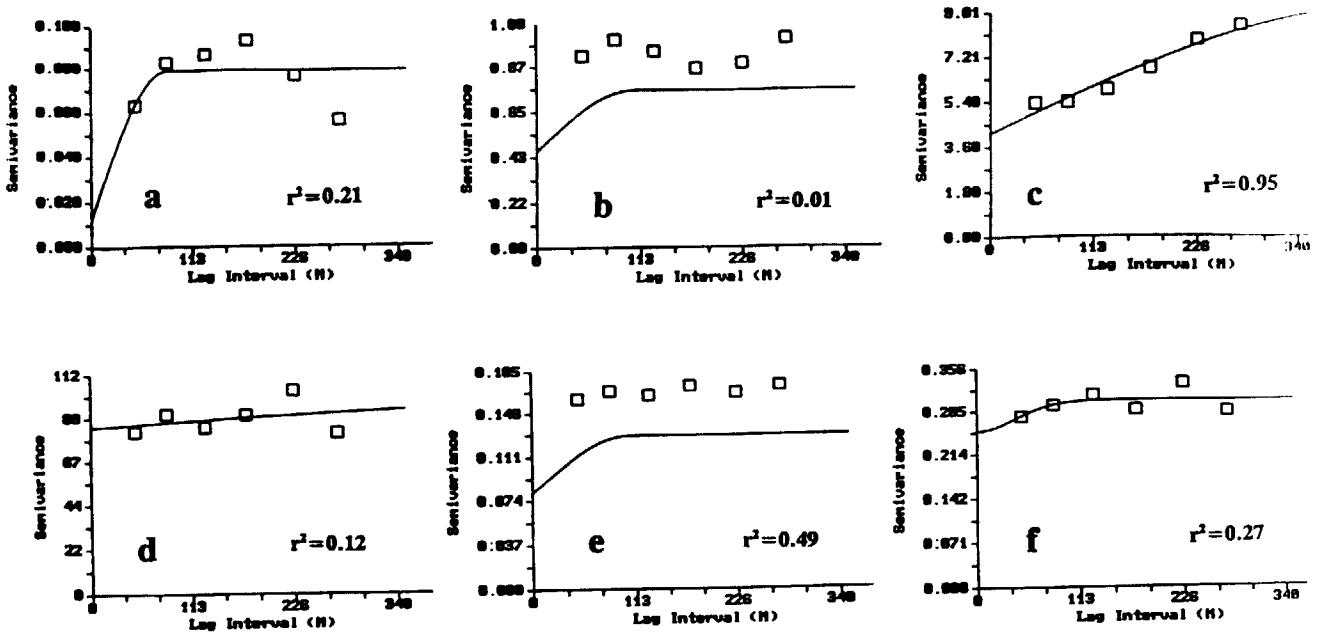


Figure 2. Isotropic semivariogram for soil variables (a) Cu, (b) Zn, (c) Fe, (d) P, (e) pH and (f) OM at Field 1 (0-15 cm depth).

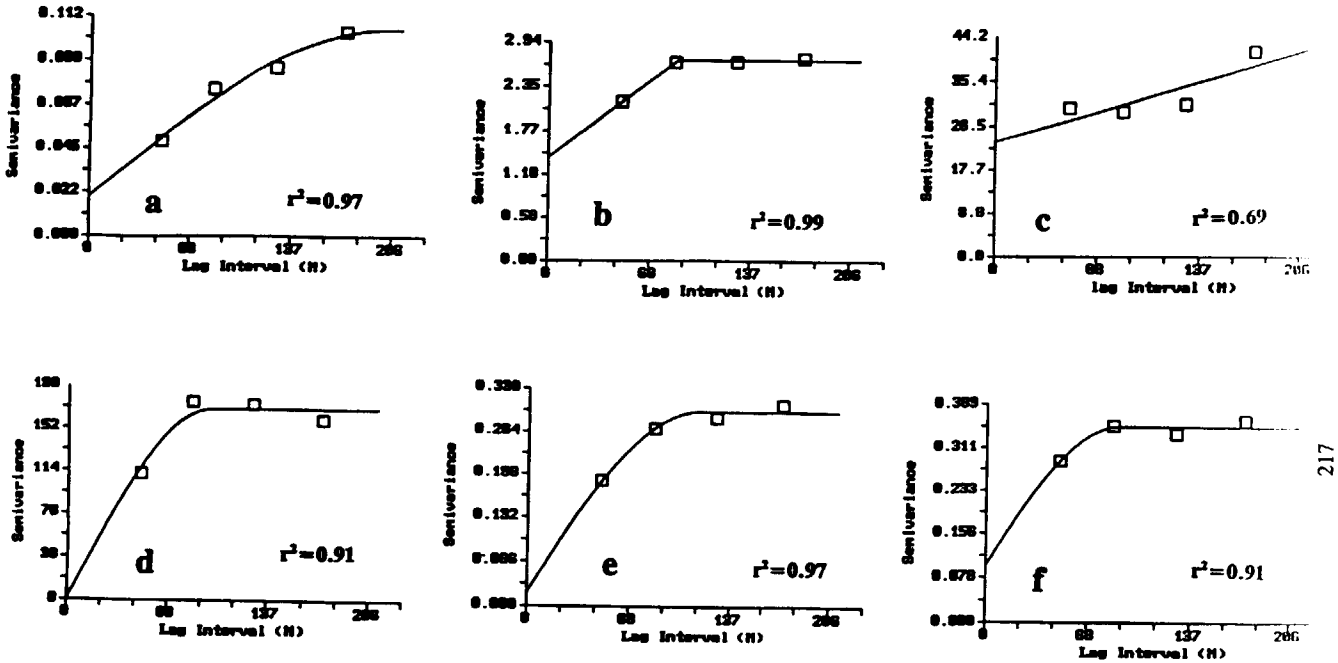


Figure 3. Isotropic semivariogram for soil variables (a) Cu, (b) Zn, (c) Fe, (d) P, (e) pH and (f) OM at Field 2 (0-15 cm depth).

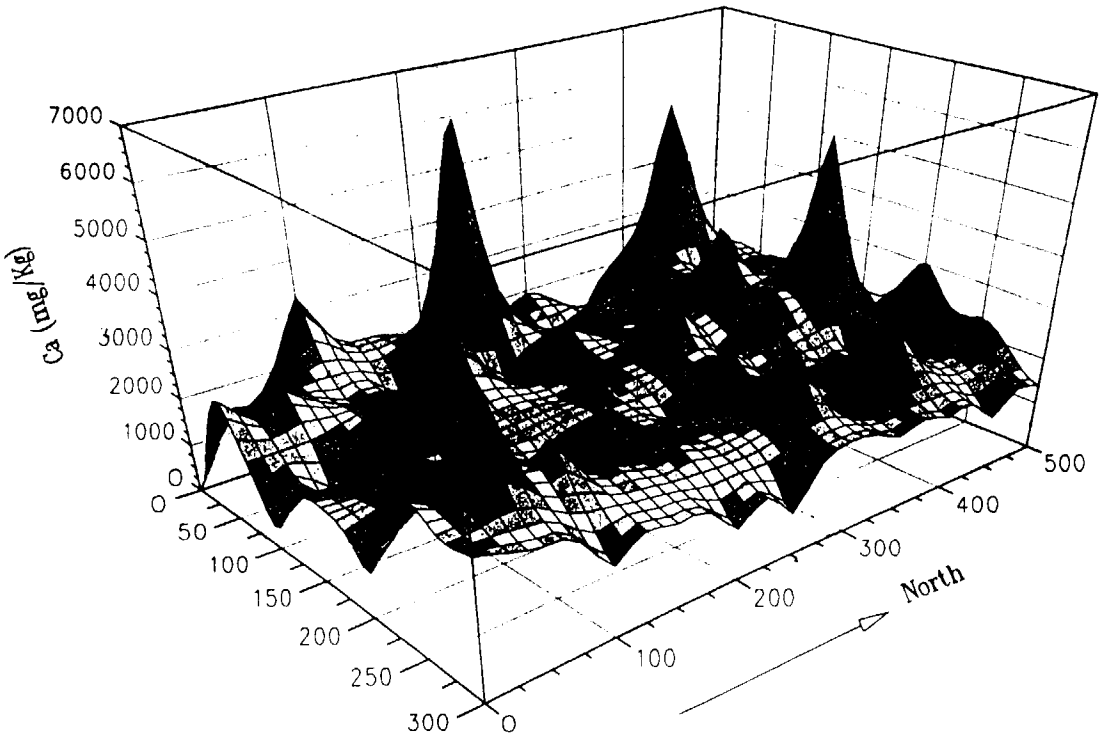


Figure 4.Three dimensional diagram of available Ca at Field 1 (0−15 cm depth).

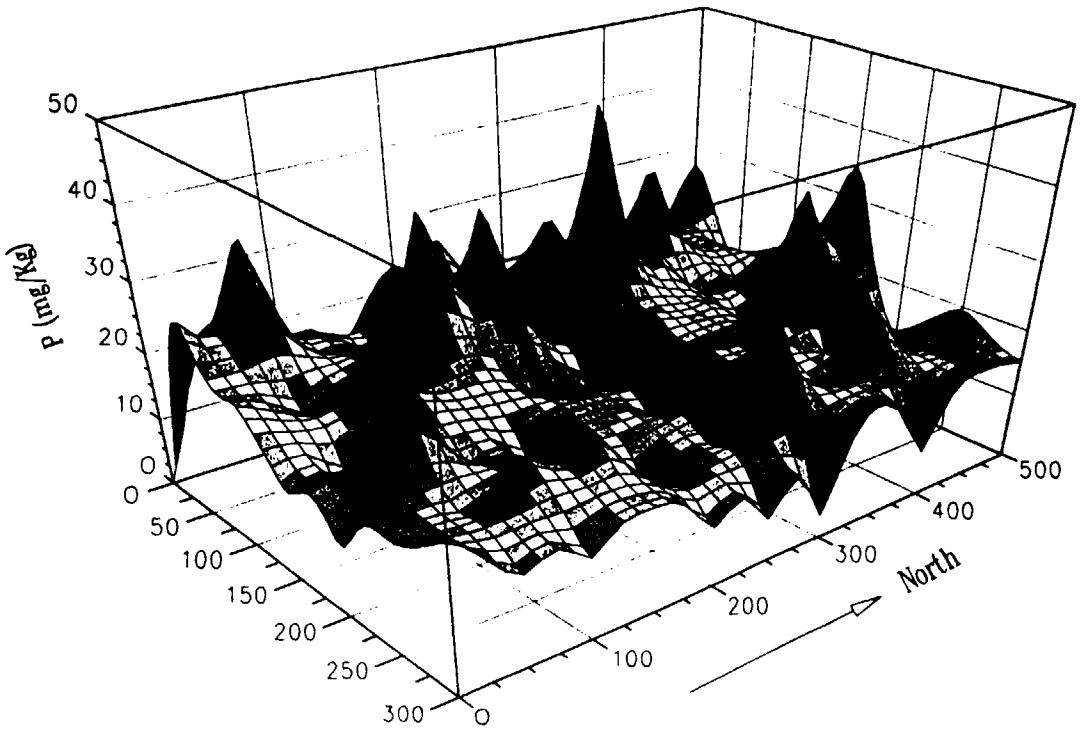


Figure 5.Three dimensional diagram of available P at Field 1 (0−15 cm depth).

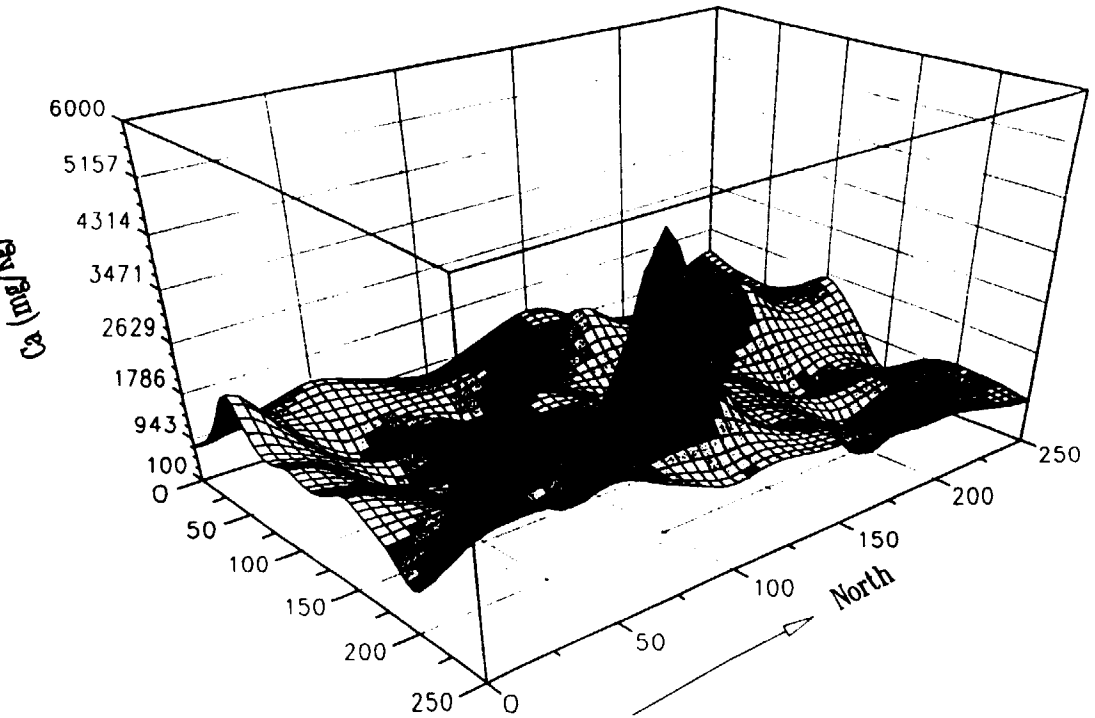


Figure 6.Three dimensional diagram of available Ca at Field 2 (0−15 cm depth).

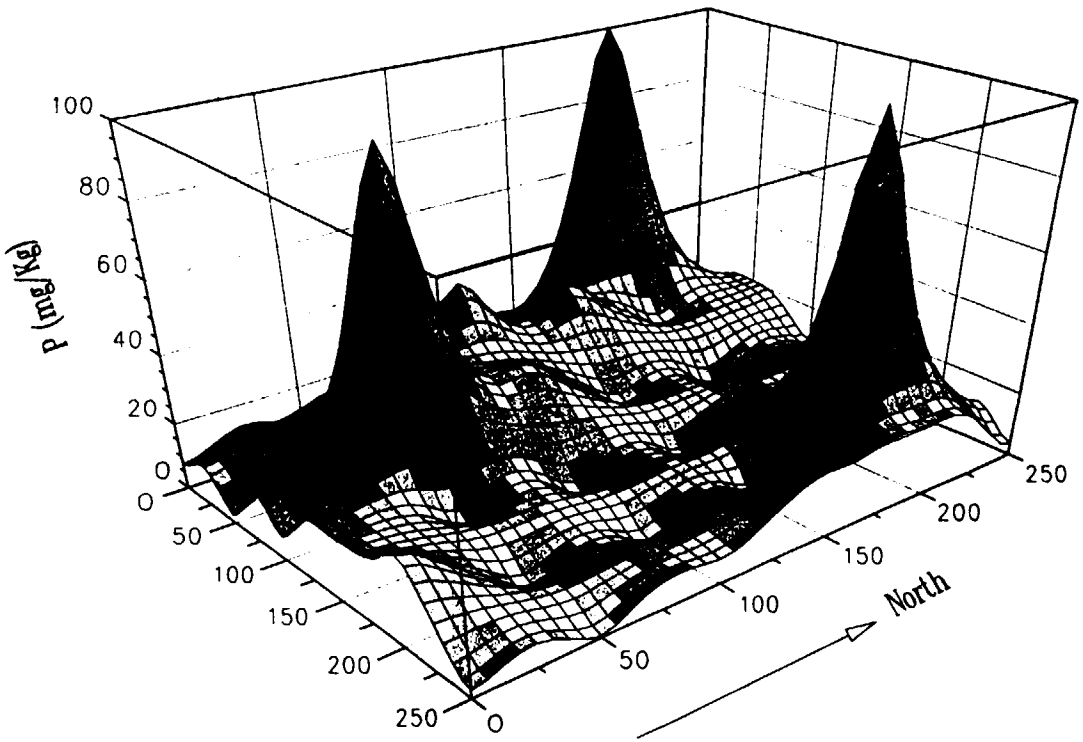


Figure 7.Three dimensional diagram of available p at Field 2 (0−15 cm depth).

Evaluation of GIS Technology in Assessing and Modeling Land Management Practices

ARCHER F.*, T.L. COLEMAN, A. MANU, W. TADESSE, AND G. LIU
CENTER FOR HYDROLOGY, SOIL CLIMATOLOGY, AND REMOTE SENSING (HSCaRS)
ALABAMA A&M UNIVERSITY, DEPARTMENT OF PLANT & SOIL SCIENCE

ABSTRACT

There is an increasing concern of land owners to protect and maintain healthy and sustainable agroecosystems through the implementation of best management practices (BMP). The objectives of this study were: (i) To develop and evaluate the use of a Geographic Information System (GIS) technology for enhancing field-scale management practices; (ii) evaluate the use of 2-dimensional displays of the landscape and (iii) define spatial classes of variables from interpretation of geostatistical parameters. Soil samples were collected to a depth of 2 m at 15 cm increments. Existing data from topographic, land use, and soil survey maps of the Winfred Thomas Agricultural Research Station were converted to digital format. Additional soils data which included texture, pH, and organic matter were also generated. The digitized parameters were used to create a multi-layered field-scale GIS. Two dimensional (2-D) displays of the parameters were generated using the ARC/INFO software. The spatial distribution of the parameters evaluated in both fields were similar which could be attributed to the similarity in vegetation and surface elevation. The ratio of the nugget to total semivariance, expressed as a percentage, was used to assess the degree of spatial variability. The results indicated that most of the parameters were moderate spatially dependent. Biophysical constraint maps were generated from the database layers, and used in multiple combination to visualize results of the BMP. Understanding the spatial relationships of physical and chemical parameters that exists within a field should enable land managers to more effectively implement BMP to ensure a safe and sustainable environment.

Additional Index Words: geostatistics, soil variability, spatial variability, BMP, conservation.

F. Archer, Graduate Student, Center For Hydrology, Soil Climatology, And Remote Sensing, Alabama A&M Univ., Normal, Al. 35762

INTRODUCTION

An understanding of the distribution of soil properties at the field and watershed scale is important for making reliable soil interpretations and assessing the effects of agriculture on environmental quality. The level of variability in soil properties is an increasing concern of land owners who are attempting to protect and maintain healthy and sustainable agroecosystems. This is also of practical importance to researchers who are investigating the complex interrelationships between soil properties. The variability may be due to several causes including variation in geographic location, climate, topography, parent materials, land use history, and the biological, physical and chemical processes within the soil (Beckett and Webster, 1971).

Spatial variation studies are fundamental to the perception of the order within the spatial distribution of soil properties (Wilding and Drees, 1978) and can be used to facilitate reasonably accurate soil boundary delineations in soil surveys. A major relevance of studies of soil variables is to describe and map soil properties over the landscape from sample data (Beckett and Webster, 1971). Properties of soils vary from place to place both laterally and vertically. The vertical variation, profile, has been the concern of pedologist for many years. It has been described conventionally by recognizing layers, horizons, and then treating each of these separately. Lateral variation has been treated similarly. Soil surveyors recognize where the soil changes in a relatively abrupt manner and draw boundaries there to separate the soil into classes. They describe each class separately from sampling points within them. Average or typical values within classes are then used as predictors for those classes (Webster and Burgess, 1983).

Geographical Information Systems(GIS), in combination with geostatistics, can be effectively used in solving many management problems (Stein, 1994). Geostatistics has afforded scientists the capability to study the spatial dependency of various soil properties. It has been reported that soil

properties are often spatially correlated either isotopically (Burgess and Webster, 1980a; 1980b) or anisotropically (Boss et al., 1984; McBratney and Webster, 1983). Geostatistical methods have been used to study the spatial dependence of soil salinity (Hajrasuliha, 1984), bulk densities (Entz and Chang, 1991), and electrical conductivities (Chery et al., 1994) within given field situations. This method has also been used to compute and display semi-variograms for soil texture and pH of soil derived from loess and glacial till (Cambell, 1978). Semi-variograms have been used to show the spatial correlation of soil properties such as phosphate-phosphorus and potassium over a range of greater than 100 m (Yost et al., 1992). Petiole nitrate content of cotton has been shown to be closely related to soil clay content and not soil nitrate (Tabor et al., 1985). Sharma and Karr (1994) showed that the high spatial variabilities of soil water and nitrogen fertilizer at the subsurface depths of an irrigated lateritic soil were affected by the high variability of clay and bulk density. It has also been shown that the average soil test potassium (K) values may be misleading if the spatial variation of K is not considered (Ndiaye and Yost, 1989).

Interest in spatial pattern of soil properties on the landscape continues to grow and it is of practical importance to both researcher and producers in making land use decisions. Although land owners have always sought better ways to manage information, the GIS technology has not been within reach for many potential users. Hardware and software are gradually becoming affordable, with the new wave of personal computers and stand-alone workstation. Most potential users now see a GIS as inevitable because the system will help them do their jobs better and faster. The hypothesis established for this study states that there is no difference in properties of soils as a result of management practices as assessed using field-scaled GIS techniques.

The objectives of this study were: (i) To develop and evaluate the use of Geographic Information System (GIS) technology for enhancing field-scale management practices; (ii) to evaluate the use of three-dimensional modeling techniques to visualize changes due to different management practices and (iii) to define spatial classes of variables from geostatistical analysis.

MATERIAL AND METHODS

Data Base Analysis and Decision-Making

A 50-m grid system was established at the Alambama A&M University Winfred Thomas Agricultural Research Station located in Hazel Green, Alabama. Soil mapping units and interpretation were extracted from the soil survey report of Madison county, Alabama (Swenson et al., 1958). Soil cores were extracted down to a depth of 30 cm, characterized using standard soil horizon terminology (Soil Survey staff, 1985) and separated into 15 cm increments. Soil physical properties (particle size distribution, organic matter, and pH) were determined for each sample using standard procedures in Methods of Soil Analysis, Part I (Klute, 1986). Land use and best management practice (BMP), depth of the A-horizon, depth to the B-horizon, drainage, infiltration rate, permeability, and water holding capacity of each soil mapping unit were coded according to the terminology set forth in the National Soils Handbook (Soil Survey Staff, 1985) used by the USDA-Natural Resource Conservation Service. The GIS analysis and modeling were done using the UNIX version of ARC/INFO and Arcview software (ESRI, 1995) on a SunSPARCstation 10 platform. The Motorla Global Positioning System by Geolink with base station was used to determine the geographic coordinates of the study area. The resulting map was then used to spatially locate areas of interest for further investigation.

Geostatistical Data Analysis

The spatial structure of the soil properties was determined using standard geostatistical techniques. Semivariograms were computed omnidirectionally at each sampling depth and at 4 angles (0, 45, 90, 120) to test for anisotropy. There were 83 samples for fileld 1 and 68 samples for field 2. The semivariograms were computed for a maximum distance of 50 m. Variance was graphed as a function of sample separation distance. The model coefficients (sill, nugget effect and range) were calculated. All geostatistical computations were performed using geostatistical software (GS+, Gramma Design Software, St. Plainwell, MI).

Table 1. Summary statistics of selected soil physical properties at different depths for field 1 and field 2.

Parameters	Depth (cm)	Mean	Minimum	Maximum	Standard deviation	# of Samples
Field 1						
Sand, %	0-15	12.99	3.68	67.68	10.74	83
	15-30	16.2	3.68	57.68	9.74	83
Silt, %	0-15	43.83	0.00	62.88	10.26	83
	15-30	38.97	0.00	72	11.03	83
Clay, %	0-15	42.88	1.44	60.32	7.86	83
	15-30	45.14	8.32	62.88	9.04	83
Org. Matter, %	0-15	1.34	0.74	1.79	0.24	83
	15-30	0.82	0.23	1.52	0.32	83
pH in water	0-15	6.23	4.74	7.29	0.45	83
	15-30	6.2	4.87	7.3	0.54	83
Field 2						
Sand, %	0-15	14.61	0.80	32.80	8.29	68
	15-30	12.78	0.80	32.80	8.56	68
Silt, %	0-15	40.59	12.00	60.00	8.56	68
	15-30	37.97	8.00	60.00	9.88	68
Clay, %	0-15	44.80	7.20	71.20	9.94	68
	15-30	49.28	31.20	71.20	8.74	68
Org. Matter, %	0-15	2.95	2.01	4.02	0.53	68
	15-30	2.57	0.67	4.02	0.71	68
pH in water	0-15	5.93	4.96	6.79	0.40	68
	15-30	5.98	4.93	6.78	0.44	68

Table 2. Semivariograms of selected soil physical properties at different depths for field 1 and field 2.

Parameters	Depth (cm)	Nugget variance	Sill	Range	Isotropic* model	Nugget %	R^2	Spatial** Class
Field 1								
Sand, %	0-15	27.2	137.2	97.5	S	19.83	0.40	S
	15-30	49.4	80.2	109.9	E	61.59	0.07	M
Silt, %	0-15	54.0	121.0	128.7	S	44.62	0.43	M
	15-30	123.0	142.6	664.8	E	86.25	0.15	W
Clay, %	0-15	36.10	67.27	17.2	E	53.66	0.07	M
	15-30	64.9	80.7	43.5	E	80.42	0.36	W
Org. Matter, %	0-15	0.04	0.07	810.9	S	57.14	0.68	M
	15-30	0.09	0.12	791.7	S	75.00	0.59	M
pH in water	0-15	0.12	0.22	190.4	S	54.54	0.72	M
	15-30	0.25	0.31	227.8	S	80.64	0.72	W
Field 2								
Sand, %	0-15	45.50	76.50	121.3	E	59.47	0.97	M
	15-30	64.00	85.00	259.5	S	75.29	0.38	W
Silt, %	0-15	48.49	72.73	105.0	L\S	66.67	0.28	M
	15-30	56.7	85.1	105.0	S	66.62	0.68	M
Clay, %	0-15	53.0	107.6	252.3	S	49.25	0.87	M
	15-30	52.10	86.35	179.0	S	60.33	0.90	M
Org. Matter, %	0-15	0.27	0.41	652.0	E	65.8	0.41	M
	15-30	0.36	0.52	131.7	S	69.23	0.24	M
pH in water	0-15	0.12	0.16	85.2	S	31.25	0.41	M
	15-30	0.08	0.20	85.3	S	40.00	0.28	M

* L=Linear
L/S=Linear/sill
S=Spherical
E=Exponential
** S=Strong (% nugget < 25)
M=Moderate (% nugget between 25 and 75)
W=Weak (% nugget > 75)

RESULTS AND DISCUSSION

Multi-layered field-scale GIS

The GIS framework described herein is a first step toward developing a more comprehensive management system for assessing and modeling land management practices. The study addresses the use of 2-D displays while providing detailed spatial data needed for the implementation of best management practices (BMP).

The thickness of the A-horizon ranged from 5 to 20 cm. Water table depths exceeded the 2 m sampling depth; however, in low areas evidence of the presence of a high water table for intermittent periods during the year was observed. Infiltration rates ranged from 0.5 cm / hr to 7.5 cm / hr and permeability ranged from 0.06 cm / hr to 50.1 cm / hr. Slope percentages ranged from 0.1 to 20 percent. The advantage of viewing information contained in the database in 2-D is that it provides a more realistic and simplistic view to the user as conditions are being evaluated.

The GIS database allows one to make any number of comparisons or speculative analyses of the entire land area or any specified field. It provides the land manager an opportunity to quickly perform analyses comparing previous management practices with current practices. The potential usefulness of this technology in evaluating crop performance due to soil variablilty as result of soil physical properties and spatial variability has been demonstrated. The Two and three-dimensional displays have been proven to be useful as decision aid for land managers. Conversely, two dimensional displays affords the farm manager an opportunity to view and model landscape conditions that are consider to be potential problems. The systems will enable decision-makers to develop a better management plan and maximize their inputs.

Variation of Soil Physical Properties

Table 1. shows that the clay content ranged from 1.44 to 60.32 in field 1 at the 0-15 cm depth. The range in clay at the same depth in field 2 was narrower (7.20-71.20%). Mean clay content in the 15-30 cm depth was significantly higher in field 2 than in field 1, however, similar amounts were obtained in the topsoil of both fields.

Silt content was higher in the topsoil and subsoil in field 1 than in field 2.

Mean of sand, organic matter, and pH in the top soil were significantly different at both sites. The same relationships were obtained in the subsoil.

Semivariograms

Semivariograms were used to determine the spatial dependence of soil physical parameters at both sites. The attributes of the semivariogram investigated were the sill, which is directly related to total sample variance, the range which is the lag distance at which the variance levels off and nugget variance which represents random and sampling error. Another attribute was the nugget semivariance expressed as a percentage of the total variance. This ratio was used to define distinct classes of spatial dependence for soil physical parameters (Cambardella et al., 1994). Majority of the parameters in field 1 were fitted to spherical models (Table, 2) spherical models were also defined for most of the same parameter in field 2 except sand content at 15-30 cm depth.

With the exception of organic matter and silt nugget variable of the parameters were higher in field 2 than in field 1 at corresponding depths (Table, 2). Similar relationships were obtained for the sill and range. Organic matter had an unusually large sill at both sites. All of the parameters in field 2 exhibited moderate spatial dependency (nugget percentage 25-75%), except sand content at 15-30 cm depth. Sand in the 0-15 cm depth were strongly spatially dependency (nugget percent < 25%) in field 1. Silt, clay and ph in the subsoil exhibited weak spatial dependency (nugget percent >75%) in field 1.

CONCLUSION

The study illustrates that there are only slight differences in the amounts and distribution (both vertically and horizontally) of soil physical properties at two sites studied. It was ascertained through the construction of semivariograms that there were similarities in spatial variability patterns for most of the soil physical parameters evaluated. It must be emphasized that the two fields studied were not chosen to represent any specific different physical-chemical conditions. The use of this technology will become more acceptable as a management tool for assessing and modeling land management practices.

REFERENCES

Beckett, P.H.T., and R. Webster. 1971. Soil Variability: A Review. Soils and Fertilizers. 34: 1-15.

Burgess, T.M., and R. Webster. 1980. Optimal interpolation and isarithmic mapping of soil properties. I. The semivariogram and punctual kriging.J. Soil Sci. 31: 315-333.

Campbell, J.B. 1978. Spatial variation of sand content and pH within single contiguous delineation of two mapping units. Soil Sci. Soc. Am. J. 42: 460-464.

Chery, P.,M. Dabas, A.Bruand, and D. Arrouays. 1994.Kriging electrical conductivity determinations for detailed study of soil thickness variability. A case study using GIS in Beucearea (France) : Trans. 15th Int. World Soil Sci Cong. (Acapulco, MX; 10-16 July 1994) vol 6b: 302-303. 1994.

Entz, T. and Chang, C. 1991. Evaluation of soil sampling schemes for geostatistical analyses: A case study for soil bulk density. Can. J. Soil Sci. 71: 165-176.

Gee, G.W., and J.W. Bauder. 1986. Particle size analysis. p. 383-412. In A. Klute (ed.) Methods of soil analysis Part 1. 2nd ed. Agron. Monogr. 9 ASA and SSSA, Madison, WI.

Hajrasuliha, S., N. Banialbassi, J. Mettey, and D.R. Nielsen. 1980. Variability in soil sampling for salinity studies in South-West Iran. Irrig. Sci. 1: 197-208.

Klute, A. 1986. Methods of soil analysis. Part 1. Physical and mineralogical methods. 2nd ed. Agron. Monogr. 9 ASA and SSSA, Madison, WI.

McBratney, A.B. and R. Webster, 1981. Spatial dependence and classification of the soil along a transect in Northeast Scotland, Geoderma. 26: 63-82.

Ndiaye, J.P. and R.S. Yost. 1989. Influence of fertilizer application nonuniformity on crop responce. Soil Sci. Soc. Am J. 53: 1872-1878.

Sarma, J.S., V.P. Singh and K.N.S. Sharma. 1988b. Analysis of spatial variability in sodic soils: 2. Point and block kriging. Soil Sci. 145: 250-256.

Sarma, J.S., K.N.S. Sharma and N.K. Tyagi. 1988a. Analysis of spatial variability in sodic soils: 1. Structural analysis. Soil Sci. 145: 180-187.

Sharma, S.K. and S. Kar. 1994. Spatial variability of fertilizer nitrogen distribution in lateritic soil under rainfed rice and irrigated wheat. Trans. 15th Int. World Soil Sci. Cong. (Acapulco, MX, 10-16 July, 1994). vol 6b: 308-309. (1994).

Soil Survey Staff. 1975. Soil Taxonomy: A basic system of soil classification for making and interpreting soil surveys. USDA-SCS. Agri. Handb. 436. U.S. Government Printing Office, Washington, DC.

Stein, A. 1994. Regionalized variable theory and geostatistics for modelling and representing soil spatial variability in GIS. Trans. 15th Int. World Soil Sci. Cong. (Acapulco, MX, July 10-16, 1994) vol 6a: 699-717. 1994.

Tabor, J.A., A.W. Warrick, D.A. Pennington, and D.E. Myers 1984. Spatial variability of nitrate in irrigated cotton: I Petioles. Soil Sci. Soc. Am. J. 48: 602-607.

Trangmar, B.B., R.S. Yost and G. Uehara. 1985. Application of geostatistics to spatial studies of soil properties. Adv. Agron. Academic Press, New York, N.Y.

Trangmar, B.B., R.S. Yost and G. Uehara 1986a. Spatial dependence and interpolation of soil properties in West Sumatra, Indonesia. I. Anisotrophic variation. Soil Sci. Soc. Am. J. 50: 1391-1395.

Webster, R. and T.M. Burgess. 1983. Spatial variation in soil and the role of kriging. April. Water. Mgt. 6: 11-122.

Wilding, L.P., and L.R. Drees. 1978. Spatial variability: A pedologists viewpoint. In diversity of soils in the tropics. Soil Sci. Soc. Am. Spec. Publ. 34: 1-12.

Yost, R.S., G. Uehara and R.L. Fox. 1982. Geostatistical analysis of soil chemical properties of large land areas. I. Semivariograms. Soil Sci. Soc. Am. J. 46: 1028-1032.

ALABAMA A&M UNIVERSITY
WINFRED THOMAS AGRICULTURAL RESEARCH STATION

SOIL MAPPING UNITS

- Abernathy
- Baxter undulating phase
- Baxter eroded undulating phase
- Baxter eroded rolling phase
- Baxter severely eroded rolling phase
- Cookeville undulating phase
- Cookeville eroded undulating phase
- Decatur and Cumberland undulating phase
- Decatur and Cumberland severely eroded undulating phase
- Decatur and Cumberland severely eroded rolling phase
- Decatur and Cumberland gullied phase
- Decatur and Cumberland eroded undulating phase
- Dewey severely eroded rolling phase
- Dickson eroded undulating phase
- Dickson level phase
- Dickson undulating phase
- Greendale cherty silt loam
- Greendale silt loam
- Guthrie
- Hermitage eroded undulating phase
- Huntington
- Lawrence
- Lee-Lobelville

500 0 500 1000 Meters

Map Projection: UTM

228

Center for Hydrology, Soil Climatology, and Remote Sensing

Applying the Experimental Method to Programming Language Development

Richard Duran[1]
Department of Computer Science
The University of Texas at El Paso 79968
rduran@cs.utep.edu

Abstract

The experimental method can be applied to programming language development. Given a language that is syntactically and semantically defined, an interpreter is built. Once completed, it may be used to provide feedback on the language definition.

We use this feedback to check the language definition, and make changes to the language and interpreter as necessary This process will be repeated as often as necessary, until we have achieved the definition that produces the best results.

It can be shown that this method of language development is modeled after the experimental method. Here, the definition serves as the hypothesis, and the interpreter serves as a means of testing and modifying the hypothesis.

Our last step will be the realization of a production version of the language. The final goal is to have a high-level language that can serve as a data mining tool used in exploring satellite telemetry data.

1. INTRODUCTION

The purpose of the paper is to present an example of how the experimental method can be used in the process of programming language development. The experimental method involves taking a given hypothesis, and testing that hypothesis via some convincing number of experiments. The results of these experiments are then compared against the hypothesis. Any discrepancies that arise force the experimenter to revise the hypothesis, and begin testing anew.

The cycle of revision and retesting is repeated as many times as needed, until the experimenter has arrived at a hypothesis that is true in all test cases. This is a scientific approach commonly used in many fields of physical science. We would like to show an additional field where this approach is applied: programming language development.

It should be briefly mentioned that it is typical to require that any such hypothesis possess certain desirable theoretical features. Otherwise, there may be no point in performing any tests. In the case of programming languages, one such requirement would be Turing Computability.

For our purposes, programming language development is the process of designing, and implementing, a language for use by some community of programmers. Every phase of this process is included, from specifying the syntax and semantics of the language, to providing some means for a programmer to actually run programs written in that language. The former shows what the language does, the latter shows how it is done.

During the course of this paper, we will

- present the approach we are taking in the development of our language,
- show how our work implements the approach used in the experimental method,
- and summarize what we are hoping to achieve as a result of this work.

[1]The student is presently working towards a Master of Science in Computer Science. Expected graduation date is May 1996. Research interests include software engineering, programming languages, and computer graphics.

2. DEVELOPING THE LANGUAGE

According to Daniel Cooke, the designer of the language being developed, a language is a theory about how one might best approach problem solving. New languages are motivated, oftentimes, by the need/desire for a different paradigm.

The motivation for the research in our language is the need/desire to explore satellite telemetry data more efficiently than is currently technically possible. Presently, only one percent of all data has been explored. We believe that the reason for this lies in limitations of the tools being used.

2.1 Defining the Language

Preliminary philosophical consideration aside, one of the first things a language designer must do is answer two questions:

- How are instructions stated in this language?
- What will those statements mean?

The answers to these can be specified via syntactic and semantic definitions, respectively. When referring to the language definition, we imply the combination of both of these definitions.

2.2 Implementing the Definition

We can now set about constructing a means of running programs from a variety of problem solving domains written in our language. Practical considerations led to our decision to build an interpreter to prototype the language. We use the language definition as the foundation upon which we build an interpreter that operates accordingly. Once completed, we are able to test our initial definition.

2.3 Testing the Definition

We test our language definition by running programs written in the language through the interpreter, and using the results to provide feedback on the definition. For our work, it proved to be very useful for the implementer to be able to readily discuss these results with the language designer.

It is during these discussions that the implementer can obtain the answers to any questions about the language, it's definition, and the test results. In addition, it becomes possible to point out potential problems, and suggest possible improvements to the language definition.

Sometimes, as a result of these discussions, it becomes necessary to modify the language definition and/or the implementation. If this is the case, the appropriate changes are made. Once this has been done, we may again run our test programs, and use the feedback from the subsequent results to determine whether or not we are now achieving the desired results.

We will repeat the cycle of testing-discussing-adapting as many times as we need to until we have achieved the language definition that produces the best results. Once this is completed, we can have a high degree of confidence in the fact that a program written in the language, as defined, will produce the results that the language designer intended.

3. RELATING THIS WORK TO THE EXPERIMENTAL METHOD

We would now like to address the matter of how the experimental method can be applied to programming language development. For this, we look at our general description of the experimental method, and our specific description of the process of developing a programming language, as done in our work. It should be fairly apparent that a correlation between the two exists.

In our description of the experimental method we discussed a hypothesis, and experiments performed to test that hypothesis. If we look at our efforts, it can be seen that the language definition serves as our hypothesis, while the interpreter itself, together with the programs we run, serve as the experiments we are performing in order to test, and modify, our hypothesis.

4. GOALS

There are a number of goals of the work described in this paper. Our immediate goal is to develop stable versions of the language definition and the interpreter. A long term goal is to realize a production version of the defined language. Our final goal will be met when we have a high-level language that can serve as a data mining tool that will be used in the exploration of satellite telemetry data.

5. SUMMARY

This paper presents the process currently being followed in the development of a language interpreter. It presents this process as an example of how the experimental method can be applied in the field of programming language design.

6. BIBLIOGRAPHY

(1) Daniel E. Cooke, Programming Languages, (in preparation).

7. KEYWORDS

Software, Software Engineering, Language Prototyping, Language Development, Programming Languages

Monitoring Properties of Programs

Francisco G. Fernandez Ann Q. Gates*
Department of Computer Science
The University of Texas at El Paso 79968
ffernand@cs.utep.edu agates@cs.utep.edu

Abstract

Development of complex systems requires interaction between a large group of people at various levels of software development, including the communication of properties of the system and the data to be manipulated. A natural idea is to maintain a centralized database of properties of the system to which all members of the development group have access, and to automate the process of checking for violations against this database. The focus of this paper is to discuss such an automated process, called integrity constraint checking. The paper defines the notion of an integrity constraint and discusses considerations for adding an automated checker to a programming language compiler or interpreter. Current work on the implementation of integrity constraint checking in a very high-level language called SequenceL is discussed, and future work in developing a similar checker in an imperative language is outlined.

1. Introduction

Developing large scale programs is a difficult task because of the vast amount of information that must be managed and communicated to personnel with a variety of backgrounds in each phase of the software development cycle. When programs require the expertise of highly trained people from multiple disciplines, the need to communicate knowledge of the system that is being developed becomes more critical. One approach to capturing some of this knowledge is through integrity constraints that state the conditions that data or objects must satisfy. The approach called *context monitoring* [Gat94; Gat95] requires specification of integrity constraints and the automatic checking for satisfiability of those constraints during execution of the program. The constraints, then, will ensure that the developed system does not violate the specified constraints.

Section 2 provides the reader with a background of the work on integrity constraints. Section 3 presents the implementation of an integrity constraint satisfiability mechanism that was. implemented in a very high-level language (VHLL) called SequenceL. This section also gives the motivation for using this language. Section 4 presents the future research direction , and the paper concludes with a summary in section 5.

2. Background

In specification languages such as VDM, Spec or Larch, although integrity constraints can be specified, no automatic facility is provided to ensure that the constraints are not violated [Cog96]. A global constraint normally becomes embedded in pre- or post-conditions either implicitly or explicitly, making it difficult to decipher the constraint without annotations.

Integrity constraint checking during prototyping, program development, and testing is useful for determining the consistency of information generated by a prototype, or a program with respect to the constraints identified during the requirements, design and coding phases. Because different groups may identify and specify constraints, it is conceivable that conflicts may exist between constraints as well as in the program itself. Because constraints are merged into a single constraint program, they provide an additional layer of communication. Furthermore, the collection of constraints can be examined to determine conflicts or potential conflicts within a constraint program.

By including the specification of constraints, the specifier is directed into considering knowledge about the operating environment of the program and not just the functions that have a direct effect on the services of the system. The specifier creates separate specifications of the functions and the constraints of the system that, ultimately, are combined into a complete specification or program of the system. The extended program provides a bridge between the problem and the program by capturing the knowledge that may invalidate the mapping or the representation.

* This work was sponsored by NASA under contract NAG-1012 and NCCW-0089

2.1 The Role of Integrity Constraints

The integrity constraints serve several purposes during the software's life cycle [Gat95]. Because a repository of constraints is maintained, it becomes possible to analyze the constraints to determine if potential conflicts exist. The earlier in the life cycle that potential conflicts are detected and resolved, the less impact they will have on the system being developed.

During the coding and testing phases, the constraints can be used to determine if properties and relationships of data in the program are maintained by the program dynamically by testing the program on numerous data sets. The process of formalizing the relationships and properties of data can be of great benefit in developing a testing strategy that will trap errors as early as possible.

There are several factors that initiate the need for maintenance:

- an independent program corrupts existing data because of faulty software
- an unpredictable sequence of events results in an inconsistent database,
- the environment in which the program is run changes and introduces different types of data, reclassifies data or requires calculations that were not considered in the requirements,
- conflicts exist between independent programs that operate on the same data

Constraints that monitor a software system dynamically can be used to identify when these situations arise. Again, with this approach the constraints are not embedded in the program, but are maintained independently. Separating the constraints ensures that changes in the code do not inadvertently change a constraint and that new constraints do not violate existing ones. Through constraint satisfiability, it is possible to determine if added or changed code violates existing constraints and to examine the effect of change on the system.

2.2 A Class of Integrity Constraints

Many categories of constraints exist (see [Gat96]). In this paper, we are concerned with the broad category of constraints that deal with the notion of "when" a constraint should be checked with respect to the state of the machine. We define *state* as a set of program-variable--value pairs that capture a snapshot of memory at different points in execution of a program.

The first type of constraint holds over every state and is called an *immediate* constraint. Immediate constraints are those in which the check occurs after a change in the state of the machine. An example of this type of constraint is one that checks the range for a particular variable. For instance, assume a program that models a network of computers and determines the shortest path between any two computers. An immediate constraint would be as follows:

> The cost associated with each connection must be less than X amount.

Intermediate constraints are checked after a sequence of state changes have occurred. This category of constraints is typically defined on more than one program-variable and require that each are updated in order for the constraint to hold. Consider the following intermediate constraint:

> For every connection between computer A and computer B, there exists a connection between computer B and computer A.

Note that the constraint may be temporarily violated in one or more states. This particular category of constraints is difficult to model *without knowledge of how the program is implemented.* It appears that dynamically checking this type of constraint requires a trigger other than merely the notion of state change. *Delayed* constraints are checked only in the final state.

3. Implementation in a VHLL

The purpose behind embedding a constraint violation monitoring mechanism in a language is to prevent the introduction of faulty data into program execution. Consider the network modelling problem mentioned in the previous section. Assume that the shortest path algorithm is used in the program and that as a result, a constraint is specified stating that only positive edge costs are allowed. If a change to the program corrupts or circumvents the embedded check that ensures all costs are positive, a negative cost may accidentally be entered. As a result, the shortest path of the network may be calculated incorrectly, since the algrorithm used to solve the problem assumed only positive costs. As a result of this faulty calculation, the system may make bad decisions, such as sending a message between two machines using the wrong path, which may in turn cause the message to not be recceived by a particular deadline, and so on. Even in this simple example, it is easy to see that the introduction of a single piece of faulty data leads to a series incorrect calculations and the introduction of even more faulty data to the program execution. By expanding a language to include constraint monitoring, lack of embedded

violation checks such as the one described above can be caught and handled during the development cycle.

One of the goals of the research is to implement an integrity constraint monitoring mechanism in the compiler or interpreter of different programming languages. The initial attempt was made in an interpreter for a VHLL called SequenceL. In this section, we consider the implementation of immediate and intermediate constraints in this VHLL.

3.1 Overview of SequenceL

SequenceL is a language developed by Dr. Daniel Cooke of the University of Texas at El Paso [Coo96]. It was designed as a language for processing nonscalar data using a single data structure. The language abstracts the interaction between the data structure and the algorithm. As Cooke explains, SequenceL is "based upon a strategy for problem solving wherein one solves a problem by describing data structures strictly in terms of their form and content, rather than also having to describe the iterative/recursive detail to produce and/or process algorithmically the data structures" [Coo96].

The only data structure in SequenceL is the ordered sequence. Sequences may be single or multi-dimensional, and may be composed of basic data types (such as integers, floating point, boolean, or character strings) or other sequences. These sequences can be used to build any data structure.

In addition, SequenceL provides several types of program constructs for processing data: regular, irregular, and generative. Each of these constructs involves the evaluation of a guarded command to an input sequence (domain) in order to produce an output sequence (range).

In regular constructs, the command is applied to all elements of the domain sequence. For example, it is possible to compute the sum of a sequence of elements by applying the addition operator to the sequence. The result is the summation of all elements of the sequence.

For irregular constructs, only selected elements of the sequence, based on position or value, are processed. For example, it is possible to count the number of items in a sequence, whose values are larger than 5 (selection based on value), or to sum the elements in even numbered positions in the sequence (selection based on position).

The generative construct is used to generate new sequences based on an expansion formula and on the contents of existing sequences. For example, the generative construct can take a sequence of a single element, [60], and generate the fibonacci sequence, [0,1,1,2,3,5,8,13,21,34,55].

Data in the program is represented by an unordered list of variable-sequence pairing, called the *universe*. A SequenceL program is applied on the universe.

Functions in SequenceL are defined with the following information:

1. Domain variables - The list of variables required to initiate firing the function. All variables in the list must be paired with a sequence in the universe in order for the function to be eligible to fire.

2. Range variables - The list of variables to be returned by the function.

3. Function body - The operations to be applied to the domain variables in order to produce the range variables.

Program execution itself is an event-driven process based on data availability. Program execution begins by traversing the list of functions (in non-deterministic order) and finding the first eligible function. A function is considered to be eligible when all arguments in its domain variable list are paired with sequences in the universe. When an eligible function is found, the domain variables are consumed from the universe, and the function body is executed. Upon completion, the variables in the range variable list (paired with their values) are added to the universe, and the process of finding the next eligible function begins. When no more functions are eligible to be fired, program execution ends. Note that because SequenceL does not provide explicit constructs for input/output, any variable-value pairings remaining in the universe at the end of program execution are considered output.

The current implementation of the SequenceL interpreter was developed in a language called Rigal. Rigal is itself a language that was designed expressly for compiler writing [Aug90]. This tool supports parsing, code generation, code optimization, and language rapid prototyping. It has built-in facilities for many compiler principles, including pattern matching with formal grammars, syntactic analysis, and tree manipulation.

3.2 Motivation for Using SequenceL

There were two primary reasons for selecting SequenceL as the language for the initial attempt. First, SequenceL provides easy access to variables

through the universe. This makes it easy to locate and check variables against integrity constraints. Second, variables are only modified by consumption from and/or production into the universe. These two processes occur at very specific points in program execution, making it easy to determine when checks against integrity constraints should be made. A sequence is considered to be a single object, even though it may be composed of many pieces of data. When a change is made to a sequence, the entire sequence is updated at one time. In fact because of this, immediate and intermediate constraints with respect to closely related objects need not be differentiated because the code that modifies the state is easily encapsulated. We make this assumption in the paper. It is important to note, however, that intermediate constraints between distinct objects is still an issue. By choosing a language where these considerations are abstracted out, we are able to focus on the interaction between constraints and programs.

3.3 A SequenceL Implementation

A program in SequenceL consists of functions that specify domain and range terms and the operations on them and the program's initial universe. Upon completion of a function's execution, the range term is paired with the sequence generated by the function and added to the universe. The SequenceL function is of the general form:

f(**Dom**($dterm_1$,...,$dterm_n$),**Ran**($rterm$))=
$term_1$[**when** relation$_1$ else]...
[$term_m$ **otherwise**],

where the brackets indicate optional items, $dterm_1$,... $dterm_n$ represent terms that are in the domain of function f $rterm$ is a term which represents the range of function f, and $term_1$, ..., $term_m$ represent terms that form the body of the function definition along with the *when* clauses. The *when* clause gives the precondition to the corresponding term.

We define the semantics of the basic and constraint function using axiomatic definitions (condensed) following the approach and notation presented in [Dij75]. Let *BasicFunction* denote

f(**Dom**(D),**Ran**(N))=T_1 **when** B **else**
T_2 **otherwise**

The semantics for a basic SequenceL function is defined as,

wp(*BasicFunction*, R **AND** P) =
x = N **AND**
$(\forall y \in \{D\}: \langle y, \_ \rangle \in U)$ **AND**
$(M(B,U) \rightarrow R^x{}_{T1}$ **AND** $P^x{}_{T1})$ **AND**
$(\sim(M(B,U)) \rightarrow R^x{}_{T2}$ **AND** $P^x{}_{T2})$,

where *wp* is a predicate transformer, T denotes a simple term in SequenceL, B denotes a boolean expression, U denotes the state of the system, "_" can match anything as an anonymous variable, $M(X,S)$ is a semantic function which yields the meaning of X in state S, R is a post-condition on the state, $R^x{}_T$ denotes a copy of R where each occurrence of x is replaced by the meaning of T, and $P^x{}_T$ denotes a copy of $U \supseteq P, \langle x, \_ \rangle$, in which the value paired with x in the universe is replaced by the meaning of T

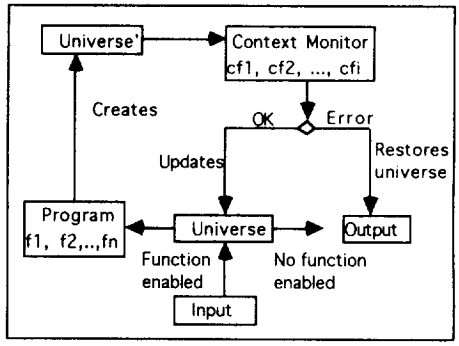


Figure 1 Execution Model.

The extended version of SequenceL permits the specification of constraint functions through which constraints are defined at a global level. Let *ConstraintFunction* denote

cf(**Dom**(D),**Ran**(N))= CT_1 **when** B **else**
CT_2 **otherwise**

The semantics of the constraint function is defined as:

wp(*ConstraintFunction*,R) =
U=N **AND**
$\forall x \in \{D\}: \langle x, \_ \rangle \in U_{S2})$ **AND**
$\exists x \in \{D\}: M(x, U_{S1}) \neq M(x, U_{S2}))$ **AND**
$M(B, U_S) \rightarrow R^u{}_{CT1}$ **AND**
$\sim(M(B, U_S)) \rightarrow R^u{}_{CT2}$

where cf is a constraint function, CT denotes a term defining a universe, U_{S1} and U_{S2} denote the state space prior to and after, respectively, the execution of some main function in SequenceL, U_S is the super universe, and $R^{U_{CT}}$ denotes a copy of the predicate defining R in which universe U is replaced by CT. Note that a constraint function fires only when at least one domain variable changes from one state to the next.

Figure 1 shows that all updates to the universe filter through the context monitor.

The process of adding an integrity constraint monitoring mechanism to the current implementation of the SequenceL interpreter involved locating the points in the interpreter where the domain variables are consumed from the universe and range variables are produced in the universe, i.e.., the function firing mechanism. In order to monitor integrity constraints, the function firing mechanism was modified to include an embedded monitoring mechanism at a point after function body execution, but before the updating of the universe with range variables. After a function is fired, the range sequences are checked against the integrity constraints for possible violations. If no violations have occurred, the range sequences are updated in the universe. However, if a violation has occurred, the range sequences are not produced in the universe. Instead, error messages are updated to the universe. This allows the interpreter to prevent the introduction of faulty data to the program while also informing the programmers where violations occurred, so that closer examination of those areas of the program can be made to determine how the violations can be prevented. Note that if an integrity constraint violation occurs, domain sequences are still consumed, but range sequences are not produced in the universe. This means that in many cases the program execution may end, because data necessary for the firing of a particular function is not available for consumption.

4. Future Research

Future research in the area of context monitoring includes many different goals. The creation of an integrity constraint reasoning mechanism that will detect possible constraint inconsistencies and contradictions. A mechanism for translating constraints expressed in first-order logic statements into constraint language statements is also an area to be studied. Finally, a major goal for the future is to implement a monitoring mechanism in an imperative language, C.

In an imperative language such as C, variables may be modified at almost any point in the program execution (within the main program, within function bodies, at the end of function execution, and so forth. This complicates integrity constraint checking by making it difficult to decide when violation checks should be made. Consider an array in C. It is possible to process a single item in the array, selective sections of the array, or the entire array through the use of loops. The difficulty lies in determining when a constraint on the array should be processed. Should it be immediately after an assignment to any part of the array? After processing the loop? How is the end of processing of a particular array determined? Because the sequence in SequenceL is an encapsulated object, this is not an issue in that language.

A more general issue concerns the trigger for constraint checking. Should checks be made at the end of a function call, after each assignment, or after some other event? Others areas of investigation include distinguishing static constraints, i.e., constraints that examine variables in a single state, and transitional constraints, i.e. constraints that examine variables over consecutive states. Finally, the possibility of revising the language to include constructs that can be used to distinguish immediate, intermediate and delayed constraints.

5. Summary

As the complexities of software systems continue to increase, the properties of objects in the system become more important and at the same time more difficult to maintain. Adding an integrity constraint monitoring mechanism to a programming language is one proposed method for managing these properties. The mechanism automates the process of checking program code against a set of integrity constraints for violations. Current efforts in this area are aimed at studying the performance of such a mechanism in the SequenceL language interpreter. The robustness of the constraint specification language and the interaction between constraints and programs will provide insight to the feasibility of implementing a similar mechanism in a compiler for the C programming language. Implementing such a mechanism in C requires investigation of the following issues: capturing the notion of state, identifying intermediate constraints and distinguishing static and transitional constraints.

References

[Aug90] Auguston, M., "Programming Language Rigal as a Compiler Writing Tool," *ACM SIGPLAN Notices*, December 1990, **25**(12), pp. 61-69.

[Cog96] Cooke, D.E., Gates, A., Demirors, E., Demirors, O., Murat, M.T., and Kramer, B., "Languages for the Specification of Software," to appear in *Journal of Systems and Software*, 1996.

[Coo96] Cooke, D.E. "An Introduction to SequenceL: A Language to Experiment with Constructs for Processing Nonscalars," to appear in *Software Practice and Experience*, 1996.

[Dij74] Dijkstra, E., "Guarded Commands, Nondeterminacy and Formal Derivation of Programs," *Communications of the ACM*, **18**(8), pp. 453-457, 1975.

[Gat96] Gates, A., "On Defining a Class of Integrity Constraints," submitted to the *International Conference on Software Engineering and Knowledge Engineering*, 1996.

[Gat95] Gates, A. and Cooke, D.E., "The Use of Integrity Constraints in Software Engineering," SEKE '95 Proceedings Software Engineering Knowledge Engineering, Rockville, MD, 1995. Skokie, IL: Knowledge Systems Institute, 1995, pp. 383-390.

[Gat94] Gates, A., *Context Monitoring with Integrity Constraints*. Las Cruces, NM: New Mexico State University, 1994 (Ph.D. Dissertation).

Towards a Methodology for Identifying Program Constraints During Requirements Analysis*

Lilly Romo, Ann Q. Gates,
Connie Kubo Della-Piana

Department of Computer Science
The University of Texas at El Paso 79968
lromo,agates,connie@utep.edu

Abstract

Requirements analysis is the activity that involves determining the needs of the customer, identifying the services that the software system should provide and understanding the constraints on the solution. The result of this activity is a natural language document, typically referred to as the requirements definition document. Some of the problems that exist in defining requirements in large scale software projects includes synthesizing knowledge from various domain experts and communicating this information across multiple levels of personnel. One approach that addresses part of this problem is called context monitoring and involves identifying the properties of and relationships between objects that the system will manipulate. This paper examines several software development methodologies, discusses the support that each provide for eliciting such information from experts and specifying the information, and suggests refinements to these methodologies.

I. Introduction

Requirements analysis continues to be one of the most challenging phases of the software development process. The purpose of requirements analysis is to determine the needs of the customer, to identify the services that the software system should provide and to determine the constraints on its development [Som96, Dav93]. The process requires communication across different groups of personnel[1] and involves the synthesis of user knowledge, management objectives, domain expert knowledge, and software engineering expertise. Clearly, communication is a crucial component of the process. The process of requirements analysis involves and results in the following [Som96, Med94]:

- Requirements elicitation: the process of gathering information concerning real world objects and the functionalities of the system from a number of sources;
- Requirements representation: the organization and representation of the gathered information using diagrams or other notations;
- Requirements specification: structured documentation of the functionality of the system[2]
- Software specification: a formal description of the software system that can be used for design and implementation.
- Software validation: the process of ensuring that the system being developed is what is actually wanted; this includes validation of the requirements specification and the software specification.

Although not necessarily linear, the various aspects of requirements identified above represent a stage of the development process. Requirements analysis traditionally centers around identifying the functionality of the system being developed, which includes the specification of functional properties such as inputs, outputs, operations and exceptions; dependability properties such as safety and security; performance constraints; and expected changes [Med94]. Another approach supports a social--technical view, i.e., it emphasizes the properties of a system that are essential for it to run successfully in its intended environment [Gog94].

An approach called *context monitoring* [Gat94, Gat95] supports the latter view. This approach promotes the notion that the requirements process, in addition to identifying the functionality of the proposed system, should elicit knowledge about the objects being

*This work was sponsored by NASA under contract NAG-1012 and NCCW-0089.

[1]This includes the customer, the user, the domain experts, the analysts, the designers and the coders.

[2]This is also referred to as the requirements document.

modeled by the software system and that this knowledge, called integrity constraints, should be maintained in a repository that can later be used to monitor the behavior of the program. Integrity constraints reflect the properties, relationships, restrictions and limitations of the objects as well as properties of the environment in which the software will run. Context monitoring should not be viewed as replacing any of the existing methodologies to requirements analysis, but rather as a supplemental method.

Constraints arise not only from the requirements process, but also from the design, coding and the maintenance phases of software development. In this paper, we examine the process of identifying constraints from requirements analysis. For example, consider a pharmacy prescription management system. Constraints that could arise from the requirements phase would be as follows: the number of refills allowed on a prescription is determined by the type of the drug dispensed; narcotic medication will have no refills, and over-the-counter drugs may have refills up to a year from the original date of prescription.

A main focus of the current research effort is to investigate the following steps that are necessary during the requirements phase to specify constraints: identification of knowledge about the types of constraints that are useful for monitoring the software, and an approach for eliciting relevant constraints from the customer and domain experts.

This paper presents a preliminary approach to eliciting such constraints. The next section of the paper reviews the process of requirements analysis in particular the relevant approaches to elicitation of requirements with respect to the constraints. Section 3 outlines a methodology for clarifying and eliciting constraints. We conclude with a summary in section 4.

2. Background

In this section, we present an overview of the requirements process and examples of the tools and techniques that are available for integrity constraint identification and specification.

2.1 Requirements Elicitation

When a system is being developed that requires deep domain knowledge and a broad understanding of the job, the process of elicitation requires the gathering of information that identifies and clarifies the requirements of the system with respect to integrity constraints for context monitoring [Gat96]. Communication is crucial to the success of the process, from initial stages of formulating the parameters of the project and eliciting requirements to formulating integrity constraints and satisfying the customer's needs. Successful completion of the project and in particular, the elicitation process, is dependent upon the ability of the requirements analyst to develop professional relationships with key personnel. The emphasis on building and maintaining professional relationships suggests that the development of dense communication networks facilitates the identification of direct/indirect customer-developer links [Kei95]. In addition, the requirements analyst must act as a translator, presenting information gathered from one source to other key personnel in a form they readily understand. A communication approach to requirements elicitation emphasizes the social embeddedness, highly interactive, and interpersonal nature of gathering information from diverse sources and making it accessible to all stakeholders. The social embeddedness of requirements elicitation suggests that the analyst must take into account larger social situations, rather than relying solely on gathering information and drawing inferences from informants in formal dyadic communication situations [Gog93, Mar]. That is, the analyst must become acquainted with and utilize different types of techniques which allow him/her to answer the questions: What has to happen? And why? An array of techniques for requirements elicitation exist, such as interviews, task analysis, focus groups, observation and participant observation, ethnomethodology, conversational analysis, protocol analysis and apprenticeships [Bey95, Gog94, Jir94, Kei95]. Each has its strengths and weaknesses to gather the information needed to identify and clarify constraints on the system.

1. Structured and unstructured interviews: dyadic communication with a predetermined and serious purpose involving the asking and answering of questions [Cas91];

2. Task analysis: a method for developing a theory of the job and the procedure associated with that theory [Car86];

3. Focus groups: a carefully planned discussion in which the perceptions of participants are solicited concerning a specific area of interest [Kru88];

4. Observation and participant observation: a blend of techniques which allows an analyst to participate and observe the everyday practices of informants [Ham83];

5. Ethnomethodology: a set of procedures for collecting and analyzing information on how competent members of a group organize behavior

[Gar67];

6. Conversational/discourse analysis: a method which allows an analyst to collect and analyze natural language situations in which competent members of a group accomplish a task or set of tasks [Gog93];

7. Protocol analysis: a set of procedures for collecting and analyzing people's accounts of what they are doing and why they are doing it [Eri80];

8. Apprenticeships: a relationship in which an analyst/designer works with a customer to learn how and what a customer does [Bey95].

Based on the argument that requirements elicitation is essentially grounded in establishing and maintaining effective professional relationships, the interview is the *sine qua non* of requirements elicitation. Interview questions must be developed which draw on the analyst's knowledge of software engineering and a lay person's common understanding of the theory of the job; however, we agree with Goguen (1993) that the problem of tacit knowledge, the inability of the interviewee to articulate one's knowledge about how to do something and why, is a continual challenge for the interviewer. Techniques, such as observation, ethnomethodology and apprenticeships offer ways to meet the challenge. They allow the analyst to ask questions of key personnel as they accomplish the task. Combining the actual "doing" of the task with questions that clarify what and why they are doing their task provide the analyst with information from which to infer constraints on the system. That is, by playing the role of the naive observer or apprentice, the analyst is able to ask questions of the "expert/master" which probe for a deeper understanding of the domain.

There are many techniques from which the requirements analyst may draw. Selection of a particular technique must be based on the criterion of usefulness. That is, which technique among the many available will provide the analyst with the information from which he/she can identify and clarify constraints on the system.

2.2 Requirements Representation

The information that is gathered from the elicitation stage is typically organized and represented using diagrams or other notations. This stage of requirements analysis is referred to as the requirements representation stage. Some guidelines for organizing the information include [Ber91]:

- Define the objectives and performance constraints of the customer organization.
- Identify the types of objects in the problem domain.
- Identify undefined concepts mentioned in the objectives and develop definitions for them.
- Identify interfaces that are required for managing instances of all types controlled by the system or maintained by users of the system.

Many of the tools available for this stage of requirements analysis use an object-oriented view. In this sense, an object represents any real-world entity that has a well-defined boundary and that is essential to the discussion of the problem domain [Dav93]. The attributes of the object, the functions performed by the object (the behavior of the object) and the possible states of the object are characteristics that should be represented. Numerous tools are available to aid the analyst. We provide an overview of some available tools [Dav93].

- Structured Analysis and Design Technique (SADT) supports the notion that natural language must be used to describe a problem and its solution, and uses a hierarchy of diagrams annotated with natural language to remove any ambiguity that may exist. The context diagram provides an abstract overview of the problem showing all inputs and outputs. Continual refinements are made to the diagram resulting in a hierarchy of diagrams. Other graphical diagrams may model the interrelationship between objects, functions or states.
- Coad's Object-Oriented Analysis (OOA) consists of specifying objects, attributes, structures, services and subjects in the problem domain. This approach provides explicit guidance to determine appropriate objects.
- Jackson System Development methodology constructs a model of the real world using entities, relationships and actions performed by the entities. The entities must map to real world objects in the problem domain.
- Data-flow diagrams show the movement and transformation of data through an organization, software system, hardware system or a combination of these. To manage the information about the data items being modeled, data dictionaries may be used. Along with the name of the item, the repository includes such things as a description, the purpose, related data items and range of values. Information such as the data structure associated with the item would be added

at a later stage.

- Structured Requirements Definition (SRD) examines inputs and outputs to learn about the context of the application. This approach represents inputs and outputs using data-flow diagrams for each unique task in the system using information elicited through interviews. The individual diagrams are merged to provide a user-level data-flow diagram. Last, the external inputs and outputs are isolated by collapsing that part of the diagram that represent the system being analyzed, and the events associated with the inputs and outputs are ordered.

Requirements techniques serve to facilitate communication and provide means of defining the system boundary, one of the goals of requirements analysis. Techniques such as SADT and OOA identify related objects. The goal of the analyst during to constraint elicitation and specification is to probe deeper into these relationships. The data dictionary manages information critical for integrity constraint checking such as range of values and related data items.

2.3 Requirements Specification

The requirements specification is typically a natural language document, supported by diagrams, that serves as the communication link between the customer and the developer. The purpose of a requirements specification is to determine the customer's needs in "sufficient detail" to plan the construction of a software system [Ber91]. This is accomplished by defining what the "software will do externally without describing how it will function internally" [Dav93, p.2]. The written document describes the requirements at the levels of the designers and developers. The document provides an understanding of the problem to be solved and identifies the goals of the system to be built. The types of information needed to specify the system according to [Dav93] include: define the objects, functions or state; control the actions associated with an object, function or state; and define relationships between an object, function or state. We present a brief overview of a few of the standards and guidelines used for requirements specification.

The Department of Defense has a software development standard called DOD-STD-2167A. The section that is comparable to the requirements specification is called the Data Item Descriptions (DID). The document consists of general background information, support documents, constraints, behavioral requirements (inputs, outputs and the relationship between them) and non-behavioral requirements (portability, reliability, efficiency, human engineering, testability, understandability and modifiability).

The NASA data item description is similar to DOD's version. Although the organization of the document if different, it also provides guidance on developing the requirements specification and uses many of the same major headings.

Another approach [Ber91] uses a formal language called Spec to ensure that the requirements are thoroughly understood before development begins and provides a systematic procedure for validating the requirements. Guidelines for development of the requirements specification includes:

- Build a formal model of the environment, called the environment model, using Spec and based on the concepts appearing in the initial problem statement.
- Express the goals from the initial problem statement in terms of the environment model.
- Express constraints such as implementation, performance and resource constraints.
- Refine each goal into subgoals until you identify the functions that the proposed system must perform.
- Identify testable goals that each function must satisfy.

The last two step leads the analyst into the software specification. Because many of the objects and relationships in an environment model are common within a domain application, Spec supports a library of such objects and relationships that can be reused and refined for the specific application.

The requirements specifications records many of the constraints needed for context monitoring. The constraints, however, are typically embedded within the document. Integrity constraint specification maintains the constraints separately.

2.4 Software Specification

The purpose of the software specification is to describe what the software will do externally without describing how it will function internally [Dav93, Ber90]. A survey of a wide variety of software specification languages [Coo96, Gat94] revealed that non-executable languages such as VDM, Spec or Larch provide constructs for specifying constraints;

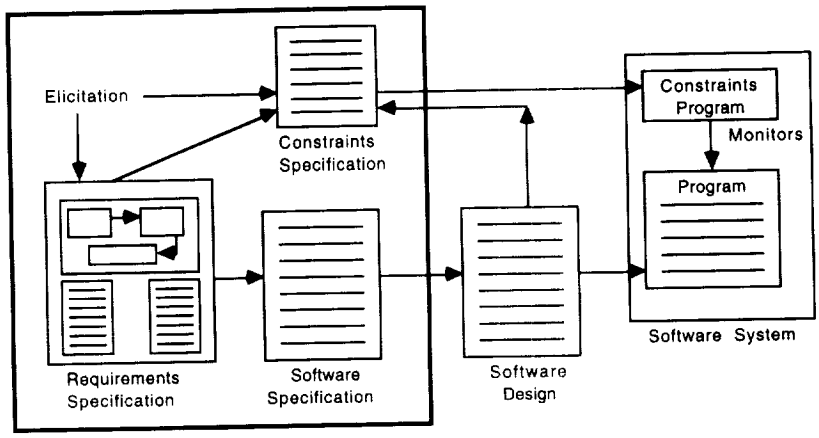


Figure 1. High-level view of the requirements process.

typically, however, the constraints must also be reflected in parts of the pre- and post-conditions of the pertinent operations or components of the specification, either implicitly or explicitly, in order to be included in the implementation of the software system. In general, integrity constraints tend to be embedded in the code of executable specification languages, and these languages do not facilitate verification that the software upholds the constraints when specified separately. Again, if formal methods are not used, the constraints remain buried in the requirements specification.

3. An Approach to Constraint Specification

In this section, we examine some of the main issues involved in eliciting integrity constraints and outline a preliminary elicitation method for formally including an approach to constraint specification. Fig. 1 illustrates the interaction between elicitation, the requirements specification, the software specification and the constraint specification in the development of the software system. The outlined box encapsulates the requirements process. This figure presents a high-level view and does not show the evolutionary nature of requirements analysis. In this section, we are concerned with the process that results in the constraint specification, i.e., the two arrows entering the box labeled "constraint specification."

To illustrate the steps that we outline for specifying constraints, we provide examples from a pharmaceutical software system that fills prescriptions and maintains a financial and medical history of the patient. We assume that elicitation and the software requirements specification has been completed. As a result, the following (selected) data items have been identified: date of prescription, date prescription filled, prescription number, patient name, patient social security number, drug code, NDC number, original quantity prescribed, quantity dispensed, number of refills allowed and number of refills left.

An outline of the constraint specification process follows:

1. **Understand the significance of integrity constraints** with respect to context monitoring. A mental model lies between human expertise and an implemented program which

results in multiple sources of ambiguity in a problem. The constraints capture the properties of objects (including domain and restrictions) and relationships between objects, and are used to ensure that the program does not violate the specified constraints. The focus of integrity constraint specification is not on the problem solution but on the properties that must be maintained by the program if it is to work correctly. Furthermore, the amount of knowledge needed to understand a program application and the diversity of this knowledge provides serious obstacles in designing current programming systems, and is thinly spread throughout the software development team [Cur88]. This complexity can ultimately lead to violations of constraints. Second, by isolating the constraints from the requirements specification, it is possible to reason about the constraints and determine potential inconsistencies. Capturing constraints from multiple sources will aid in determining when certain conflicts in the requirements exists. Last, the constraints monitor changes to the requirements during the development process and during maintenance of the system after it has been in use.

2. **Become familiar with the types of constraints** that can be specified on a software system. A large classification of integrity constraints can be found in [Gat96].

3. **Review the requirements document** to retrieve constraints that have been specified from the initial elicitation process. The constraints identified at this step include those that deal with the domain of the data item and the range of valid values. For instance, the following constraints may be culled from the requirements document:

- The drug code ranges from integer values 0-5.
- The NDC number assigned to a drug is a unique number.

4. **Identify domain experts and compile a list of directed questions** for each expert. For instance, for the above example the domain experts could include a member of the State Board of Pharmacy and one or more pharmacists. The types of questions that are directed to the state board member would likely concern relationships between objects, domain of objects, and laws. A possible question would be, "Is there a relationship between the date a prescription is prescribed and the date it is filled?" The following constraint may result:

- "Date filled" must be within 7 days of "date prescribed" for drugs in category X.

The above constraint may trigger the following:

- The value stored in "date filled" must be consistent with the system clock.

The questions directed at the pharmacist would target technical knowledge as well as questions concerning how this expert verifies information. During the interview process, the specifier extracts the methods, approaches, or "tricks" which an expert uses to verify that the information generated by the program is accurate and that calculations are correct. For example,

- The number of refills should always decrease based on the original quantity prescribed and the quantity dispensed.

5. **Identify representative users** and compete a list of directed questions for each user. For instance, a pharmacist's technician will use the system to fill a prescription. Therefore, the questions may probe verification of the information output by the software. An example constraint is:
- The quantity dispensed should be consistent with the amount billed the patient with respect to price per unit.

6. **Use of techniques such as apprenticeship and ethnomethodology** to identify further properties and relationships and to clarify constraints. The use of these techniques will help the analyst uncover tacit knowledge that is key to understanding and clarifying domain specifications.

4. Summary

Context monitoring is an approach that requires specification of integrity constraints and automatic checking for satisfiability of those constraints during execution of the program. The constraints are collected from multiple sources during the different phases of software development. The paper focuses on the requirements analysis phase. Techniques such as interviews, apprenticeship and ethnomethodology were discussed as ways to gather and identify information concerning the properties and relationships of objects that are modeled by the program. Developing reliable, useful software is one of great challenges of today and specifying constraints is one approach toward meeting this challenge.

REFERENCES

[Ber91] Berzins. V., and Luqi, *Software Engineering with Abstractions*. Menlo Park, CA: Addison-Wesley Publishing Company, 1991.

[Bey95] Beyer, H.R. and Holtzblaatt, K., "Apprenticing with the Customer," *Communication of the ACM*, **38**(5):45-52, 1995.

[Car86] Carlisle, K.E., *Analyzing Jobs and Tasks*. Englewood Cliffs, N.J. Educational Technology Publications, 1986.

[Cas91] Stewart, C.J. and Cash, W.B. Jr., *Interviewing: Principles and Practices*. Dubuque, IA: Wm. C. Brown Publishers, 1991.

[Coo96] Cooke, D.E., Gates, A., Demirors, E., Demirors, O., Murat, M.T., and Kramer, B., "Languages for the Specification of Software," to appear in *Journal Systems Software*, 1996.

[Dav93] Davis, A.M., *Software Requirements: Objects, Functions and States*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[Eri80] Ericcson K.A. and Simon, H.A., "Verbal reports as Data," *Psychological Review*, **87**(3):215-251, 1980.

[Gar67] Garflinkel, H., *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall, 1967.

[Gat94] Gates, A., *Context Monitoring With Integrity Constraints*. Las Cruces, NM: New Mexico State University, 1994 (Ph.D. Dissertation).

[Gat95] Gates, A. and Cooke, D.E., "The Use of Integrity Constraints in Software Engineering," SEKE '95 Proceedings Software Engineering Knowledge Engineering, Rockville, MD, 1995. Skokie, IL: Knowledge Systems Institute, 1995, PP. 383-390.

[Gat96] Gates, A., "On Defining a Class of Integrity Constraints," to be submitted at the International Conference on Software Engineering and Knowledge Engineering, 1996.

[Gog94] Goguen, J.A., "Requirements Engineering as the Reconciliation of Social and Technical Issues," in Jirotka, M. and Goguen, J.A., (eds.), *Requirements Engineering Social and Technical Issues*. San Diego, CA: Academic Press Limited, 1994, pp. 165-200.

[Ham83] Hammersley, M. and Atkinson, P., *Ethnography: Principles in Practice*. London: Tavistock Publications, 1983.

[Jir94] Jirotka, M. and Goguen, J.A., *Requirements Engineering Social and Technical Issues*. San Diego, CA: Academic Press Limited, 1994.

[Kei95] Keil, M. and Carmel, E., "Customer-developer Links in Software Development," *Communication of the ACM*, **38**(5):33-44, 1995.

[Kru88] Krueger, R.A., *Focus Groups: A Practical Guide for Applied Research*. Newberry Park, CA: Sage Publications, 1988.

[Mar] Maritzen, L.M., *Identifying Customer rRequirements*. A NASA Workshop handbook on file at the Computer Science Department, University of Texas at El Paso.

[Mcd94] McDermid, J.A., "Requirements Analysis: Orthodoxy, Fundamentalism and Heresy," in Jirotka, M. and Goguen, J.A., (eds.), *Requirements Engineering Social and Technical Issues*. San Diego, CA: Academic Press Limited, 1994, pp. 17-40.

[Som96] Sommerville, I., *Software Engineering*. Menlo Park, CA: Addison-Wesley Publishing Company, 1996.

Integrated Geoscience Analysis of Groundwater Resources: El Paso, Texas Region

C.L. Gillespie [1]
Pan American Center for Earth & Environmental Studies
The University of Texas at El Paso (UTEP)
El Paso, Texas 79902

G.R. Keller
Pan American Center for Earth & Environmental Studies
The University of Texas at El Paso (UTEP)
El Paso, Texas 79902

Abstract

El Paso, Texas, has a population of about 600,000. Its sister city, Juarez, Mexico, has a population of about 1,000,000. These arid communities receive most of their water from the aquifer in the Hueco Bolson. The increase in population and industrialization in the region continues to escalate the rate of water extraction from the known aquifers. The Hueco Bolson is a complexly faulted, asymmetric graben and is one of a series of basins due to Cenozoic age extension in the Rio Grande Rift. The rift extends north into Leadville, Colorado and as far south as Chihuahua City, Mexico. Though these basins are large in areal extent, recent studies suggest fresh water aquifers occur only as small lenses. Our initial results show the subsurface structures to be more complex than the surface expressions of the basins. The purpose of this study is to integrate subsurface and surface data (e.g., geologic maps, gravity, land satellite imagery, and drill hole and well logs) to evaluate the groundwater resources.

Introduction

Water is the most important natural resource in the El Paso region, Texas region. This increasingly populated area depends on a shared underground aquifer (Hueco Bolson) for the majority of its water. The supply of water from this resource has been rapidly declining for a number of years due to the increased demand from all three borders (Texas, New Mexico and Mexico). This demand for fresh water warrants exploration activity. Information about basin fill thickness and subsurface structures is needed to understand the occurrence of ground water distribution within the Hueco Bolson aquifer. To attack this problem gravity data were used because it provided information about the geometry of the basin. An overlay of this data with land satellite images allowed us to examine the complexity of the subsurface features relative to the "simple" surface expression. The combination of these two tools also aided in evaluating the large areal extent of the basins in the region to comparatively small fresh water lenses within the basins.

1Cindy Lorraine Gillespie obtained her B. S. Degree from Texas Christian University (Fort Worth, Texas), and her M.S. degree from the University of Texas at El Paso (UTEP). She has spent six summers with ARCO Oil and GAS Co. and Mobil as a research and exploration geologist. Presently, she attends UTEP and is pursuing her PhD in Geoscience. The PhD research will incorporate geophysical tools (gravity, seismic, land-satellite imagery, and well log data) to understand geologic problems.

Regional Geology

At the beginning of the Cenozoic era, forces associated with the Laramide orogeny produced uplift, thrusting and faulting of large masses of rock. During the Oligocene and Neogene times, an extensional regime associated with magmatic intrusions began affecting the area that culminated in the development of the Rio Grande Rift (Keller et al., 1990) which formed a series of basins (Hueco, Tularosa, Mesilla) and faulted troughs along the Rio Grande River. The Rio Grande rift extends 950 km from north Chihuahua through New Mexico into Leadville, Colorado (Ramberg et al.,1978) (Fig. 1).

The Hueco Basin is characterized by broad depressions and high-angle normal faults which have down-dropped basins in relation to the adjacent mountain masses.

It has a general north-south trend north of El Paso and northwest to southeast trend southeast of El Paso. It extends north into New Mexico as the Tularosa Basin near Oro Grande, New Mexico, and as far south as Chihuahua in Mexico. Within Texas and southern New Mexico, it is bounded by the Franklin and Organ Mountains on the west and by the Hueco Mountains on the east. To the south in Mexico, the basin is flanked by the Sierra de Guadalupe and the Sierra de San Ignacio.

Instead of being associated with one continuous basin which changes strike from north-south to northwest-southeast at El Paso, the Hueco Bolson is composed of three sub-basins. North-south trending sub-basins contains most of the water supplied to the El Paso and Juarez regions.

The majority of the Hueco Bolson is filled with the Fort Hancock Formation which is composed of alternating layers of clay, siltstone and minor sandstone. These layers were deposited in a fluvial environment (Strain, 1966) and the Camp Rice Formation (limestone conglomerate and igneous boulders cemented by caliche deposited in an alluvial environment). Most of the formations are covered with Quaternary alluvium, terrace gravel, fan gravel, calcretes or eolian sand.

Previous Work

Previous geophysical data in the region indicate the presence of sub basins within the Hueco Bolson. The northernmost basin is an elongated north-south directional basin and extends from Texas into Chihuahua, with an estimated fill in the area of Cuidad Juarez of 2.0 km. This basin is represented by a prominent north-south gravity low in the residual Bouger anomaly map. The second sub-basin lies southeast of El Paso, in the area of Fabens. Here the gravity data have been used to obtain an estimated basin fill thickness of 2.6 km. North of the area of Fabens there is a small third sub-basin with an estimated 1.3 km thickness of unconsolidated deposits.

Using low Bouger gravity anomaly and well data, Wen (1983) concluded that the northernmost sub-basin has a maximum sediment thickness of 3,000 m. Gravity profiles across the Hueco Bolson show that the deepest part of the basin is associated with the steep gravity gradient flanking the mountain ranges to the west and that the fault systems are parallel to the frontal zone of the Mountain ranges (Hadi 1991).

Data Collection and Processing

Gravity data from the UTEP database were used to evaluate hydrologic conditions. Standard corrections that account for the flattening of the earth, latitude and elevations were applied to the data. Terrain corrections were performed, and 2.67g/cm^3 was used for the crustal density.

Hydrogeologic Conditions

The Hueco Bolson is located in the southern part of the Rio Grande rift which is

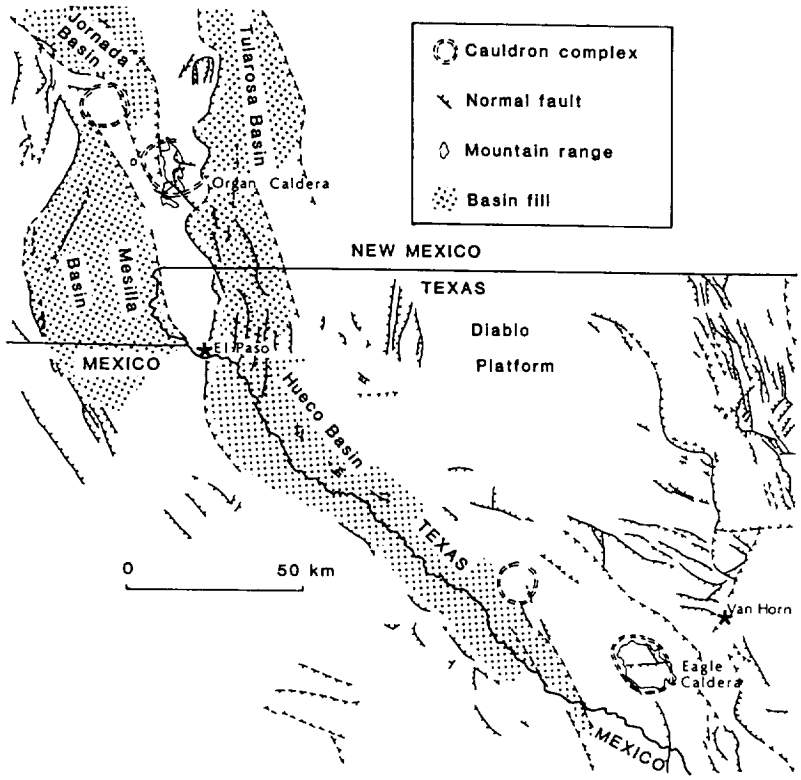


Figure 1
El Paso Texas region. Mountain ranges that flank the Hueco Bolson are the Organ, Franklin and Sierre de Juarez, to the west and the Hueco Mountainsto the east. In the south, the basin is bounded by the Sierras de Guadalupe and San Ignacio, and by the Quitriaun, Malone and Finlay Mountains. The basins are large in areal extent, but only contain small lenses of fresh water aquifers located adjacent to the mountains.

characterized by closed intermontane basins. The mountains surrounding the basin are the primary source of material that filled the basin during the Tertniary and Quaternary. Recharge to the aquifer also occurs along the main mountain ranges surrounding the aquifer in the region and their associated fault zones.

Fresh water within the Hueco Bolson is restricted to approximately the 300 m of upper basin fill deposits, mainly the Camp Rice Formation and uppermost sediments of the Fort Hancock Formation. Below this zone there is a smaller lens that contains water of varying salinity (800 to 2500 TDS). The water in the remainder of the basin is very salty with a TDS content reaching approximately 40,000 (Cliett, (1969).

Availability of Groundwater

It is estimated that by the year 2010 there will be more than 3 million people in this region. Ashworth (1990) estimated that as a consequence of this increase in population, there will be a 70% increase in the water consumption in the El Paso area, with the majority being withdrawn from the Hueco Bolson.

White(1983) estimated that the U.S. side of the Hueco Bolson had over 10 million acre feet (MAF) (3,048 acre km) of recoverable fresh water remaining in storage. Forecasts of the rate of depletion (calculations based on dividing fresh water in the storage by the yearly usage) suggests drinkable water in the Hueco Bolson will be exhausted before the year 2050. Since this estimate does not consider Mexico, the year of expiration is probably sooner than the above prediction.

Summary and Conclusions

The area mountains are primary source of the materials that filled the basin during the Ternary and Quaternary. The primary role of the surrounding mountains is that they serve as a way for infiltration and recharge to the aquifer, as well as a barrier to the movement of groundwater and the source of the dissolved solids found in it. Gravity lows exist over the basins and gravity highs over the mountains. These gravity lows are a result of the accumulation of the less dense material in the down-faulted areas. High-density areas are a result of dense sedimentary, igneous and metamorphic rock.

The basins have thick sediment packages and are large in areal extent, but only in areas adjacent to the mountains are fresh water aquifers found. At the present rate of extraction, the aquifers will be exhausted in about 50 years. The increased population growth of the region will hasten the depletion.

More work is needed to understand the nature of faults in this area and how fault systems serve as pathways for water recharge, water movement, and control of where fresh water occurs.

Key Words

LANDSAT Images;Groundwater; Gravity Data

References

Ashworth, J.B., 1990, Evaluation of groundwater resources in El Paso County, Texas: Texas Water Development Board Report 324.

Cliett, T.,1969, Groundwater occurrences in El Paso and its related geology: New Mexico Geological Society, 20th Field Conference Guidebook, p. 202-214.

Hadi, Julfi, 1991, A study of the structure and subsurface geometry of the Hueco Bolson: Unpubl. M.S. Thesis, University of Texas, El Paso, 88 p.

Keller, G.R., Morgan, P. And Seager, W.R.,1990, Crustal structure, gravity anomalies and heat flow in the southern Rio Grande rift and their relationship to extensional tectonics: Tectonophyics, v.174, p. 21-37.

Ramberg I.B., Cook, F.A. and Smithson, S.B., 1978, Structure of the Rio Grande rift in southern New Mexico and west Texas based on gravity interpretation: Geological Society of America Bulletin, vol.89, p107-123 .

Strain, W.S., 1980, Pleistocene rocks in the El Paso and Hudspeth Counties, Texas, adjacent to interstate Highway 10, *in* Guidebook to the Trans-Pecos Region: New Mexico Geological Society, p179-182.

Wen, C.L., 1983, A study of Bolson fill thickness in the southern New Mexico, west Texas and northern Chihuahua: Unpubl. M.S. Thesis, University of Texas, El Paso,74 p.

White, D.E., 1983, Summary of hydrologic information in the El Paso, Texas area, USGS Open File Report, p. 83-775-77.

Mass Loading Characteristics of Crystal Clock Oscillators

J. B. Cobb[ac], V. R. Morris[bc], and A. N. Thorpe[ac]

[a]Department of Physics and Astronomy
[b]Department of Chemistry
[c]Center for the Study of Terrestrial and Extraterrestrial Atmospheres
Howard University, Washington, D.C. 20059

Abstract

The 10-MHz piezoelectric quartz-crystal microbalance (QCM) has been used extensively for stratospheric aerosol sampling (1). We have undertaken laboratory studies of the QCM response to mass loading by trace gases. However, this device requires dual oscillator circuitry and the mass sensitivity can often be affected by the electronics. The coatings on the quartz crystals are sometimes difficult to remove after they have reacted with a particular gas and a disposable crystal system would be desirable. The cost of the dual oscillator-based QCM makes this a prohibitive option. Since our goal is to develop a cost-effective microbalance system with stable electronics we have begun testing of crystal clock oscillators, which are assembled with their own circuitry. We have been using chemically specific coatings for ozone to determine if the sensitivity and mass-frequency ratios are comparable to that of the 10-MHz QCM.

Introduction

The Center for the Study of Terrestrial and Extraterrestrial Atmospheres (CSTEA) at Howard University is investigating the effects of Supersonic Transports (SST's) in the upper stratosphere. The exhaust plumes of the SST may be a contributing factor to the number of aerosols and trace gases that affect the overall chemical cycles for ozone and other important stratospheric gases. In order to determine the actual effect of these aircraft, a highly sensitive detector is needed to detect gases down to pptv levels and aerosols in sub-microgram amounts. The original QCM was made up of a ten-stage cascade impactor with piezoelectric substrates, which is used only for aerosol detection. Modifications have been made to the stack to include four additional stages for gas detection.

Each stage contains two quartz crystal oscillators. The first crystal is used for aerosol impaction, and the lower crystal compensates for frequency shifts due to temperature changes. The beat

frequency, which is the difference the reference crystal and the sampling crystal, is monitored to give near-real time measurements of the mass of deposited aerosols (2).

Piezoelectric quartz crystals have frequency and time standard accuracy of 1 part in 10^8 or better (3), which leads to the extensive application of these crystals as highly sensitive detectors for gas chromatography and dosimetry (4). The crystals vibrate with minimal energy dissipation which make them perfect mass balances. When an electric field is applied across a piezoelectric material, it induces a stress in the material that changes its resonant frequency. This causes a vibrational, or oscillatory, motion in the quartz crystal, resulting in the generation of standing acoustic waves. The electromechanical coupling and stresses (resulting from the electric field) depend on crystal symmetry, angle cut of the crystal substrate, and configuration of electrodes used to apply the electric field across the crystal (5). The AT-cut quartz crystal, are wafers of quartz cut at approximately 35^o from the z-axis which are often used as gas detectors (5). These crystals have two excitation electrodes on opposite sides of the crystal which result in shear vibration in the direction parallel to the electric field and the propagation of a transverse shear wave through the crystal in the thickness direction. When the quartz thickness is increased, there is a decrease in the fundamental frequency. The electrodes used for previous and current cascade impactors was composed of a thin layer of silver, chrome, and nickel. This combination was chosen to protect the silver which is the electrode. The coatings used for selected gases tend to react with the silver electrodes. Nickel is less reactive, but does not bond easily with the silver. Hence chrome is used to bond the silver to the nickel. Even though the electrodes do eventually erode over a period of time with the reaction of the gases and coatings, you can prolong their lifetime by the choice of substrates.

The oscillation frequency of a quartz crystal depends on the total mass of the crystal, and that of any coating layers (or electrodes) on the crystal surfaces. Metal and many other solids do not greatly affect the crystal ability to vibrate. When liquids are deposited on the quartz surface, the

ability to vibrate is often impaired. The relation between the frequency change ΔF and the crystal thickness variation Δz is given by

$$\frac{\Delta F}{F} = \frac{-\Delta z}{z} \qquad (1)$$

Equation (2) can also be written as

$$\frac{\Delta F}{F} = \frac{\Delta M}{A\rho z} \qquad (2)$$

where F is the resonance frequency, ΔM is the change of the mass of the quartz crystal due to the absorption of the gas, A is its surface area and ρ is the density of the quartz (6).

There is a great concern about the overall performance and operation of the QCM design. During the course of a typical flight experiment the quartz crystals are subjected to temperature and pressure changes. We have observed that the circuitry of the sensing device is not always stable in each stage over the range of active temperature and pressure. In order to eliminate the problems that exist with the original design, we have decided to search for a sensing device that is more stable. Clock oscillators are widely used in many different types of computers, but we decided to expand the actual usage. Since clock oscillators are developed with wide range quartz crystals, we decided to replace the dual oscillator impactor with a single impactor. Crystal clock oscillators are inexpensive, and are developed with their own circuitry, which allows for consistency. There are no current references pertaining to the use of clock oscillators as quantitative mass sensing devices. Questions concerning the operation of the CCO's addressed herein are: (1) frequency stability of

the CCOs with metal covering and without it; (2) sensitivity and mass-frequency ratios; and (3) QCM mass-frequency comparison. If CCOs have a better frequency stability and higher mass sensitivity, then the QCM flight stack will be redesigned for field testing with the CCOs.

Experimental Methodology

The 10-MHz, 18.43-MHz, and 24-MHz clock oscillators are all being testing for stability. The procedure for each test consists of finding out the baseline frequency of the crystals with the metal encasing and without the metal encasing. Once the metal covering is removed, nitrogen gas is passed over the clock oscillator in the laboratory stack, which consists of a base, one dummy stage, and a test stage. The experimental setup consists of the following: (1) a power supply (Hewlett Packard 6236B) for a potential drop of 5V; (2) a flowmeter (Matheson Model, FM-1050) for measuring the flow through the ozone generating system; (3) an oiless pump (Thomas Industries, Model No. 14TFEL) with a standard micrometering valve for the air flow (1.5 LPM); (4) nitrogen gas (99.995%, Potomac Airgas) for stability testing and coatings; (5) an ozone generator (Troline, Model 115c-1) as a source for variable ozone concentrations. The rate of ozone produced can be controlled by varying the flow rate of air through the reaction tube or by varying the intensity of the radiation from the tube. A constant concentration of ozone was employed for these studies (ca. 80 ppm). A 1.0% solution of polybutadiene (PBD) in toluene is used as the coating for ozone. A 8"x 11 1/2" board with an opening as large as the substrate is used to cover the clock oscillator. This coating is applied to the substrate by an airbrush technique. We determine the amount of the coating deposited on the surface by the change in frequency. For 10-MHz clock oscillator, the frequency is dropped by 3,000 Hz. The frequency drop is monitored by Universal Counter (Hewlett Packard, 5313A), which is interfaced with a PC to store data, and a 100-MHz oscilloscope (Tenma, 72-6025). For the 18.4-MHz and 24-MHz a ratio is determined by dividing their frequencies with the frequency of 10-MHz crystal. Since the sensitivity increases with the increase in frequency, the ratio found for each crystal is multiplied by 3-KHz drop. This process is used to determine the amount of coating for each crystal.

Summary & Future Directions

Each graph displays the clock oscillator frequency as a function of time. For the 10-MHz crystal a 4-hour run was taken with the metal covering still intact. After a minimum period of two hours, the crystal oscillator frequencies have stabilized to ±2 Hz. A second 4-hour run is taken with the metal covering removed. Finally, a 4-hour run is taken to determine the effects of 80 ppm of ozone on a coated clock oscillator for 4 hours. This procedure is repeated for each set of crystals (10-MHz, 18.4-MHz, 24-MHz).

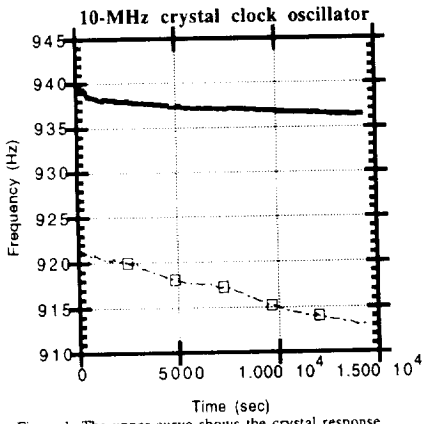


Figure.1 The upper curve shows the crystal response with metal encasing still intact. The lower curve shows the crystal response after the metal encasing was removed with no gas flow.

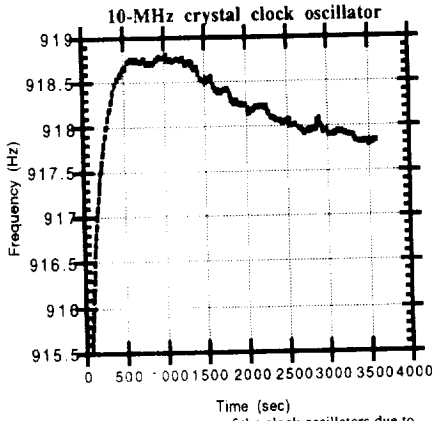


Figure.2. Frequency response of the clock oscillators due to the effects of 1.5 LPM of nitrogen gas across the crystal surface.

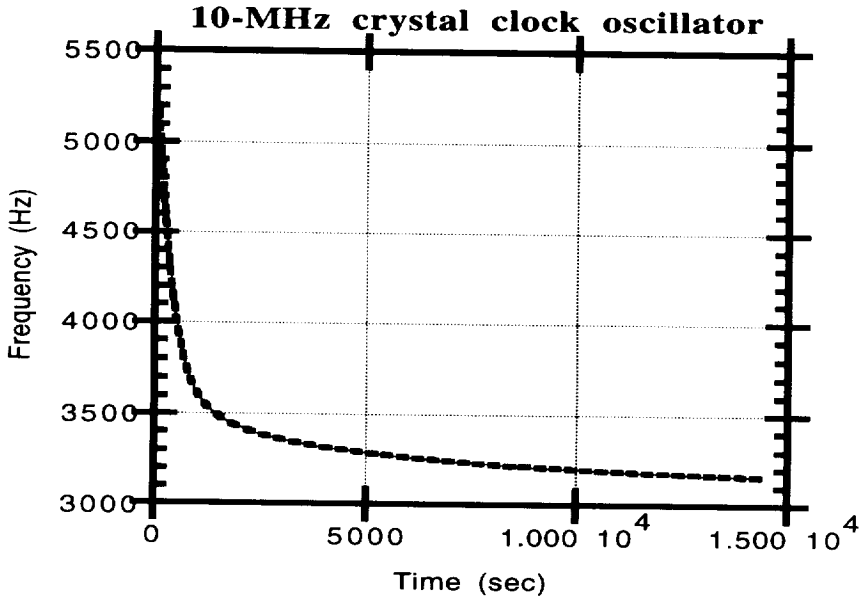


Figure.3. This curve displays the frequency response of a crystal coated with 1% PBD and exposed to a 1.5 LPM flow of 80 ppm O_3 in air for 4 hours.

At this point, we have only taken data for the 10-MHz crystal. We are still in the process of collecting data.

The clock oscillators show a strong sensitivity to the 1.0% solution of polybutadiene in toluene. The original frequency of the 10-MHz crystal with metal covering still intact was 9.999940 MHz and began to decrease over time. In Figure 2, after the top was removed, the frequency was 9.999913 MHz and was still decreasing. When the test first began, we assumed that the

frequencies were being taken at room temperature. However, there are some temperature gradients within the lab we need to be concerned with. Our next set of data will include acquisition of simultaneous temperature and frequencies data.

Acknowledgements

The authors would like to thank Dr. M. Ware and Mr. D. Davis for being key components in this research. This research is supported by NASA grant NAGW-2950.

Biography

Janel Cobb is a second year graduate student in the Physics Department at Howard University. She received her B.S. Degree in Physics from Grambling State University in 1994. She is currently a member of the American Meteorological Society (AMS), and the Graduate Physics Student Body (GPBS). Primary research interests involve the study of atmospheric modelling and remote sensing of atmospheric processes.

References

1. R. Chuan, *Rapid Measurement of Particle Size Distribution*, 764-774, Chapter 5.
2. S. V. Hering, *Aerosol Science and Technology* 7, 257-274 (1987).
3. W. H. King, Jr., *Analytical Chemistry* 36, 1735-1739 (1964).
4. G. G. Guilbault, J. Kristoff, and D. Owen, *Analytical Chemistry* 57, 1754-1756 (1985).
5. M. D. Ward and D. A. Buttry, *Science* 249, 1000-1007 (1990).
6. C. Christofides and A. Mandelis, *J. Appl. Phys.* 66, No.1, 3986-3992 (1989).

OSCILLATOR STRENGTHS OF Ar I FROM 86.6nm TO 89.5nm, AND THE 104.8 AND 106.7 LINES

I.U. Anya, C. K. Kumar
Howard University Center for Studies of Terrestrial & Extraterrestrial Atmosphere
Harry D. Morgan
National Institute of Standard & Technology, Gaithersburg, MD.

Subject headings: Photoabsorbtion; Cross section; Oscillator strength; Ultraviolet; Spectra; Argon.

1. INTRODUCTION

Accurate oscillator strengths for the spectral lines of Ar. I in the discrete and continuum regions provide valuable quantitative information for further understanding the electronic structure of matter and its interaction with energetic electromagnetic radiation. This information is of importance in areas of application such as radiation-induced decomposition, plasma physics, astronomy, biophysics, and the testing and development of theoretical methods. Column densities of argon atoms in diffuse interstellar clouds can be determined from measured equivalent widths only if accurate oscillator strength (f-value) data are available.

Experimentally, photoabsorption and photoionization cross-section measurements in the ionization continuum regions of Ar I have been widely performed using Beer-Lambert law photoabsorption and the double-ion-chamber methods. Line emitting light sources have most commonly been used. This gives the cross section at discrete wavelengths. The Hopfield continuum [7], generated by repetitive, condensed discharge through helium, provides a useful continuum source in the energy region 11.3-21.4 eV. With the advance of synchrotron radiation an intense and continuous light source has become available for measuring the photoionization cross sections of Ar I up to 29.6 eV.

Several theoretical oscillator strength calculations of Ar. I have been reported. Cooper [1], employing a one-electron central potential model, Amus'ya [2], applying the random phase approximation with exchange (RPAE) method, and Weiss [3] using transition probabilities of life time measurements, and employing the Einstein and Frank- Cordon coefficients (critical compilation), have calculated the oscillator strengths for the transitions from the ground states to the $3s^23p^5(^2P_{3/2,1/2})4s$ and 4d states. The photoionization cross-section maxima of argon are shifted to energies above the ionization threshold, showing significant departure from the hydrogenic model. In addition, minima (sometimes called the cooper minima) have been observed in the photoionization cross section of argon. Instead of using a pure coulomb nuclear potential, Cooper [1], employing a more realistic potential similar to the

Hatree-Fock potential for the outer subshell of each atom. Similar methods by Manson and Cooper [4] and McGuire [5], starting with Herman-Skillman central potentials, have been able to theoretically reproduce the maximum above the threshold and also the existence of the cooper minima in the photoionization cross sections starting from one-electron approximations. However, the above calculations give narrower peaks shifted in energy relative to the experimental cross sections, with the cross sections at the peak maxima two or three times higher than the experimental values. This paper presents measured precise photoabsorption oscillator strengths of Ar I lines from 86.6 to 89.5nm, and the 104.8 & 106.7 mm lines. There were measured at a resolution of 0.9 pico meter using the Very High Resolution Spectrometer (VHRS) at the Synchrotron Ultraviolet Radiation Facility (SURF II) of the National Institute of Standard and Technology. Measurements of the 104.8 nm & 106.7 nm were taken for comparison, since there is extensive literature on these two argon lines.

EXPERIMENTAL APPARATUS AND PROCEDURE

The VHRS which is a 6.65 meter off-plane Eagle mount spectrometer that can also be used as a monochromator connected to SURF II. This experiment was performed in the monochromatic mode.The SURF II utilizes an electron storage ring as its light source. The VHRS provides scanning of the vacuum ultraviolet spectrum with photoelectric detectors. Online photoelectric recording at intervals of 0.1 pico meter in first order of the spectrum can be made. A resolution of 0.8 pico meter [6] has been obtained at 80 nm. The detector is a photomultiplier with a sodium salicylate scintillator.

The NIST spectrometer uses a 4800 lines/mm gold-coated spherical grating blazed for maximum efficiency at 90 nm. The rulings and entrance and exit slits are oriented horizontally.

The beam line leading to the spectrometer entrance slit includes three cylindrical glass focusing mirrors at about 80, 70, and 80 degrees of incident respectively, to provide both vertical and horizontal convergence of the beam at the f-number of the instrument (~f/60). The radiation is primarily horizontal polarized. The exit slit is the entrance to a 11.43 cm long absorption cell. With entrance and exit slit widths set at ~20 um and the exit slit length masked to 0.5 mm, the absorption detector records about 10E4 cps.

Constant pressure is maintained in the cell against gas loss through the exit slit into the spectrometer by a pressure control valve. Two detectors are used, one is used as a reference. Each of these utilizes a glass envelope photomultiplier tube (PMT). Each PMT is potted in an epoxy with its preamplifier to form a sealed unit. The exposed ends of the PMT'S

are the sodium salicylate scintillators. The cell PMT forms the rear wall of the 11.43 cm long absorption cell. At the front of the cell is the 20 um exit slit of the spectrometer. The reference PMT is mounted 1.27 cm behind a 355 um pinhole. The pinhole and exit slit are 6 cm apart on the Rowland circle. To reduce stray light, masks are mounted in front of the cell and the reference detector. The entire light path is windowless from the storage ring to the photomultiplier detectors.

Data acquisition and equipment control is achieved by a stepping motor and an IBM-AT compatible microcomputer with a CAMAC powered crate. The microcomputer and crate use the GBIP (IEEE-488) bus interface platform. PMT counts, sample temperature and pressure, beam current, elapsed time, and exit slit position are digitized by CAMAC modules in real time and read into the computer. The development of the spectral features is observed in the progress of the scan.

Absolute cross sections, using the post exit slit cell, are obtained by normalizing the absorption cell counts to the vacuum intensity counts at each wavelength position of the scan where data are taken. The vacuum intensity is measured on the absorption cell channel before and after each scan for this purpose.

The cutoff wavelengths of krypton, neon, and argon are utilized to provide estimates of the remaining scattered light and of second order contributions of the light. The gas pressures are chosen to maintain absorption

intensities in the linear portion of Beers' law and is treated as Lorentzian profile. The instrumental linewidth is treated as Doppler profile. The absorption intensities are therefore treated by a curve growth technique assuming Voight profiles. To avoid line-saturation problems all data were taken at very low pressure (< 30 milliTorr). This also eliminates the problem of pressure broadening. Contributions from stray light and higher order radiation have to be carefully assessed and the measurements appropriately corrected. These Beer-Lambert photoionization measurements give good agreement in terms of the shape of the continua [6]. Masking the exit slit, to eliminate stray light, and correcting by subtracting higher-order radiation there is good agreement with theoretical results.

Cross sections are determined by the expression:

$$\sigma = [\ \ln[(I_0 - I_D)/(I - I_D - I_V - I_S)]]/N$$

where I_0 is the incident intensity, I_D is back ground intensity, I_V is vacuum intensity, I_S is the scattered intensity, which is 20% of I_0, I is the attenuated intensity and N is the column density of Ar. In the sample cell.

The optical oscillator strengths have been computed using the relationship

$$f = (112.96/\lambda^2)\int \sigma(\lambda)d\lambda$$

where $\sigma(\lambda)$ and λ are measured in units of 10^{-17} cm^2, and nm, respectively.

The following table shows the result.

λ(nm)	f (Theoretical)	f (This paper)
106.67	0.061	0.062
104.82	0.254	0.223
89.43	No available data	0.112
87.99	0.0268	0.031
87.61	0.093	0.124
86.98	0.0119	0.027
86.68	0.106	0.147

REFERENCES

[1] J.W. Cooper, phys. Rev. 128, 681 (1962).

[2] M. Ya. Amus'ya, N. A. Cherepkov, and L.V. Chernysheva, zh. Eksp. Teor. Fiz 60, 160 (1971) [Sov. Phys. JETP 33, 90 (1970)].

[3] W. L. Wiese, M. W. Smith and B. M. Miles, Atomic Transition Probabilities: Sodium through Calcium, Natl. Stand. Ref. Data ser., Natl. Bur. Stand. (U.S.)Circ. No. 22 (U.S. GPO, Wash., D.C. , 1969), Vol. II.

[4] S. T. Manson and J. w. Copper, Phys. Rev. 165, 126 (1968)

[5] E. J. McGuire, Phys. Rev. 175, 20 (1968).

[6] H. D. Morgan, et al., J. Geophysical Res. Vol. 98, No. A5, (1993)

[7] P. H. Metzger and G.R. Cook, J. Opt. Soc. Am. 55, 516 (1965).

Characteristic of Nitron for Use as a Chemical Sensor in Studies of the Upper Atmosphere

K. Meadows*, C.K. Wright, S.C. Sims, V.R. Morris

Department of Chemistry
Howard University
The Center for the Study of Terrestrial and Extraterrestrial Atmospheres
Washington, DC 20059

Abstract

We are investigating the use of nitron as a potential chemical sensor for nitric acid and other electron deficient nitrogen oxides. Solutions of nitron in 1-propanol, toluene, and chloroform have been tested for use on a piezoelectric quartz crystal microbalance. We are testing various solvents and metal cations which can maximize the lifetime and reaction specificity of nitron so that they may be used as chemical coatings for stratospheric measurement of trace gases. Results of the work to date will be shown, and future direction discussed.

Introduction

The Stratospheric Wakes Analysis Project (SWAP) is a research campaign conducted by scientists at the Center for the Study of Terrestrial and Extraterrestrial Atmospheres (CSTEA) at Howard University. The mission of SWAP is to develop a scientific basis for the assessment of the atmospheric impact of subsonic and supersonic commercial aircraft emissions. We are using a Quartz Crystal Microbalance/ Surface Acoustic Wave (QCM/SAW) cascade impactor mounted in a belly pod on an ER-2 high altitude airplane to make in-situ measurements of emission particulate and to measure concentrations of several trace gases: among them O_3, SO_2, HCL, and HNO_3.

The balance between the production and destruction of ozone in the remote

stratosphere is largely dependent upon the local concentration of NO. In polluted regions the local concentrations of stratospheric chlorine species(ClO_x) and oxygen (HO_x) also play significant roles. One of the chief measurement concerns is the partitions between active and reservoir species in the stratosphere and how anthropogenic injection may affect this balance. Nitric acid is the principal reservoir for stratospheric NO_x and HO_x. This research specifically focuses on the use of nitron for the detection of gas phase nitric acid.

Experimental

Nitric acid gas engenders severe experimental handling problems at low concentrations due to its large sticking coefficient and its affinity for water. This problem was reduced by threading Teflon tubing through the instrument to reduce the effective surface area and by using extra dry diluent gas.

A 0.1% nitron in propanol coating was originally chosen for the detection of HNO_3.[1] However, it decomposed too quickly to be of practical use in our applications. Typical half-lives of less than an hour dictated the rejection of this coating as a possible candidate for nitric acid detection. The search for a more stable nitron solution, that is tuned for reactivity with nitric acid, is presently under investigation. The possible solvents chosen for this study are hexane, 1-propanol, toluene, benzene, chloroform, and carbon tetrachloride.

A spectrophotometer(Bausch&Lomb,Spectronic-20) was used to generate spectral calibration curves. The wavelength of maximum absorbance was determined as an indicator of the stability of nitron in each solvent, and was used as an indicator of the extent of photoxidation.

Each experiment consisted of three concentrations of nitron solutions(0.05%, 0.1%,

0.5%). The first solvents tested were propanol and toluene. Each solution was tested three times to have an average percent transmittance of the solution for each time tested. In order to get a true representation of the life time, the solutions were tested immediately after preparation. Experiments took from two to four hours, depending on the response of the solution with regards to time. Absorbance values were recorded every ten minutes.

FTIR studies (Perkin-Elmer, 1600 Series) were performed to determine the spectra of nitron and its reaction products after exposure to HNO_3. Background FTIR spectra of the solvents were acquired, then subtracted from the spectra of the nitron solutions. Time-dependent studies were carried out to estimate the half-life of nitron in each solvent.

The effects of adding metals to 0.5% nitron in toluene solutions were also studied. The metal used were iron, magnesium, zinc, and cadmium. The purpose of adding these metals was to see if metal complexes could be created to extend the lifetime of nitron and tailor the reactivity of nitron in toluene for specific nitrogen oxides.

In these experiments, 0.03 g of each metal was added to 2 ml of a 0.5% solution of nitron in toluene. The solution was heated for 10 minutes and allowed to cool. The FTIR spectrum of each metal containing solution was recorded and compared to the original spectrum of the 0.5% nitron solution. Evidence for complex formation was observed in the solution only after the addition of metallic cadmium. The addition of nitric acid to this combination resulted in the formation of a brown gelatinous precipitate. There were no detectable complexes formed by the other metals.

The next step was the determination of the effects that the addition of the metals would have on the ability of nitron to determine particular NO_x species. These tests were performed by adding 0.25 ml of 0.5 M HNO_3, separating the supernatant from the different

precipitates that formed, and obtaining the FTIR spectrum of the supernatant. The next step will be to bubble HNO_3 gas through a solution of the 0.5% nitron in toluene with metals added to see if the same reaction takes place.

We are now investigating the use of chloroform and carbon tetrachloride as possible solvents. These solvents are being considered because of the high solubility of nitron, the solvent volatility, and their low reactivity. The principal FTIR absorptions for nitron have been identified by using a 2.0% solution of nitron in chloroform.[2] The strongest absorptions were observed at 1616, 1583, 1541, and 1501 cm^{-1} with an instrument resolution of 4.0 cm^{-1}.[2]

Summary

At this stage of testing, the following results have been obtained.

*Nitron, which is light sensitive, decomposes rapidly in propanol with a half-life on the order of 30 minutes. Nitron does not as rapidly decompose in toluene. We observe an increase in the transmittance as a function of time. The accelerated photoxidation of nitron in propanol may be due to the presence of OH^- species.

*The four IR peaks that characterize nitron lie in the region of about 1500-1625 cm-1.

*Concentrated HNO_3(16M) would not nitrate the ring as efficiently as hydrated forms of HNO_3. This is because the reaction requires a water in order for the acid to completely disassociate.

*The formation of a metal complex between nitron and cadmium has been observed.

Acknowledgements

The authors would like to acknowledge NASA grant number NAGW-2950. I would also like to give a special thanks to Dr. Vernon Morris for all of his wisdom and guidance.

References

1. Merch Index (1995).

2. E.J. Beckman, AIChe Symposium No. 309, 91, (1995).

Biography

Kapres Meadows is a chemistry graduate student at Howard University. He graduated from Fort Valley State College in 1995 with a BS in Chemistry. His research experience includes working for the Geochemistry section of Phillips Petroleum Co., in Bartlesville, OK, for the last two years and currently determining suitable solvents for nitron as a chemical sensor. Mr. Meadows is currently pursuing his Masters degree in Chemistry under the direction of Dr. Vernon Morris.

Calibration of the QCM/SAW Cascade Impactor For Measurement of Ozone In The Stratosphere

Cassandra K. Wright, S.C. Sims, C.B. Peterson, and V.R. Morris

Department of Chemistry
Center for the Study of Terrestrial and Extraterrestrial Atmospheres
Howard University
Washington, D.C. 20059

Abstract

The Quartz Crystal Microbalance Surface Acoustic Wave (QCM/SAW) cascade impactor collects size-fractionated distributions of aerosols on a series of 10 MHz quartz crystals and employs SAW devices coated with chemical sensors for gas detection. Presently, we are calibrating the ER-2 certified QCM/SAW cascade impactor in the laboratory for the detection of ozone. Experiments have been performed to characterize the QCM and SAW mass loading, saturation limits, mass frequency relationships, and sensitivity. We are also characterizing sampling efficiency by measuring the loss of ozone on different materials. There are parallel experiments underway to measure the variations in the sensitivity and response of the QCM/SAW crystals as a function of temperature and pressure. Results of the work to date will be shown.

Introduction

The concern regarding the photochemical destruction of ozone in the Earth's upper atmosphere stems from its ability to shield Earth's surface from potentially harmful solar ultraviolet (UV) radiation. Ozone produced in the stratosphere arises from Ultraviolet dissociation of molecular oxygen to form atomic oxygen. Although the UV radiation destroys ozone, a steady state is achieved because atomic oxygen and molecular oxygen rapidly recombine to form ozone. The ozone balance also involves reactions with families of naturally and anthropogenically produced radicals that contain chlorine(ClO_x), nitrogen(NO_x), hydrogen(HO_x), and oxygen atoms. The nitrogen oxides are principal components of aircraft

exhaust and hence lie behind concern over the effects that supersonic stratospheric aircraft may have on ozone. Models predict that increased NO_x can significantly accelerate, perhaps accelerating ozone depletion.[1]

The Stratospheric Wakes Analysis Project (SWAP) is a research campaign conducted by scientists at the Center for the Study of Terrestrial and Extraterrestrial Atmospheres (CSTEA) at Howard University to determine the effects of the stratospheric aircraft exhaust on ozone. The mission of SWAP is to provide data inputs for current models to assess the atmospheric impact of stratospheric supersonic and subsonic aviation. SWAP has completed several successful test and data-taking flights using a quartz crystal microbalance surface acoustic wave cascade impactor (QCM/SAW) The NASA/ER-2 high altitude research aircraft was flown and in-situ measurements of emission particulate and measure concentrations of several trace gases; hydrochloric acid, ozone, and nitric acid. The QCM/SAW is an analytical instrument which is capable of measuring eight size-fractionated distributions of aerosols and the concentrations of trace gases in the Earth's stratosphere down to the parts per trillion (pptv) range simultaneously.[2] It is a compact, lightweight, and sturdy device that is one of the principal instruments used for the characterization of atmospheric aerosols.

The piezoelectric sensor used in the QCM stages is a slab of quartz with an electric field applied normal to the surface that causes it oscillate. When mass is added to the crystal, its frequency drops. The eight QCM stages measure the concentration of aerosols while the four Surface Acoustic Wave (SAW) stages measure the concentrations of chemical vapors. The SAW stages contain piezoelectric quartz crystals with interdigital electrodes deposited on their surfaces. The applied electric field sets up a surface acoustic wave whose wavelength is

determined by the spacing of the electrodes. The frequency of the surface acoustic wave is approximately 200 MHz, with the SAW stages having 400 times the sensitivity of the QCM stages. The complete instrument is controlled by an embedded microcontroller and the only command issued by the pilot is to start the preset sampling sequence set by the user prior to flight.

Experimental

Normal operation consists of a sampling pump accelerating a gas stream through the first nozzle of the instrument with a high speed jet. The QCM is positioned close to the jet's exit and normal to the direction of the flow. The jet of air is forced to curve sharply around the crystal and the inertial motion of the aerosols in the air stream impact against the crystal. Particles which do not have a sufficient inertia to move forward and impact upon the sensing crystal are carried by the air stream to the following stage through a smaller nozzle. The inertial separation process is repeated with each stage capturing aerosols down to a certain size. The velocity of the aerosol particles within the flow are dependent on the aerosol mass and nozzle orifice.

The air sample enters the gas detection section of the QCM/SAW after the aerosol stages. This section of the detector contains four individual stages. The stages designed to detect the four gases, O_3, HNO_3, SO_2 and HCl are constructed using the SAW piezoelectric microbalance crystals. Each gas sensor has a specific polymer coating applied to its surface that increases its response to a specific gas. The detectors are capable of detecting these gases down to the low ppb level corresponding to a one minute exposure. Long exposure times can dramatically reduce the lower limit of detection.[3]

The use of piezoelectric quartz crystals as stable gas detectors was first demonstrated by King in 1964.[4] King showed by using a coating with a specific reactivity towards one compound PQC's could be used to detect and measure gas-phase concentrations to the parts-per-million (ppm) range.[4] More recently, SAW devices have been used for the detection of biological materials, antibodies, warfare, and small aerosol particulate.[5] We are using SAW devices to detect and trace stratospheric gases. The primary objective of this study is to determine which coatings are stable and selective enough for stratospheric deployment.

A typical laboratory experiment consists of using a crystal coated with a dilute (0.1%) solution of coating on the crystal. For the ozone gas, polybutadiene is used as the coating. A stream of a gas with a well defined concentration is pumped through the instrument for a certain length of time. A Thermoenvironmental (Model 49 PS) Ozone Calibrator is used to supply the ozone to the QCM/SAW. For the ozone loss experiments a Thermoenvironmental (Model 49 C) Ozone Analyzer was attached on the outlet port of the QCM/SAW stack to quantify the loss of ozone through the instrument. Both instruments have been certified and are traceable to NIST standards. The flow rate is controlled by a micrometering valve and monitored by a volume (Matheson Model FM 1050) flow meter and exhausted with the sampling pump. Unless actual gas sampling is taking place, an N_2 gas stream is directed across each of the crystals consecutively to stabilize the beat frequency between experiments. The change in the beat frequency is monitored electronically and stored in a computer for later analysis. Slopes from a series of response curves at a given flow rate and temperature are combined to generate the calibration curve of change in beat frequency vs. time. A typical mass loading curve showing the response of the SAW to different concentrations of ozone is illustrated in Figure 1. The slope is

seen to change dramatically and nearly linear with each change in the ozone concentration. Exposure times ranged from five minutes to three hours depending on the concentration of ozone. Figure 2 shows a typical mass loading at a single concentration with a theoretical fit superimposed on the data.

Summary and Future Directions

We have found that polybutadiene is adequate for the detection of O_3, to concentrations in the low ppbv range. This coating yields a linear response in beat frequency as a function of the concentration under standard operating conditions. In the future we plan to simulate flight conditions and optimize coating concentrations.

Acknowledgment

This research was supported by NASA funding under grant number NAGW-2950. The authors acknowledge the invaluable direction and support of their advisor, Dr. Vernon R. Morris and The Center for Terrestrial and Extraterrestrial Atmospheres director, Dr. Arthur N. Thorpe.

References

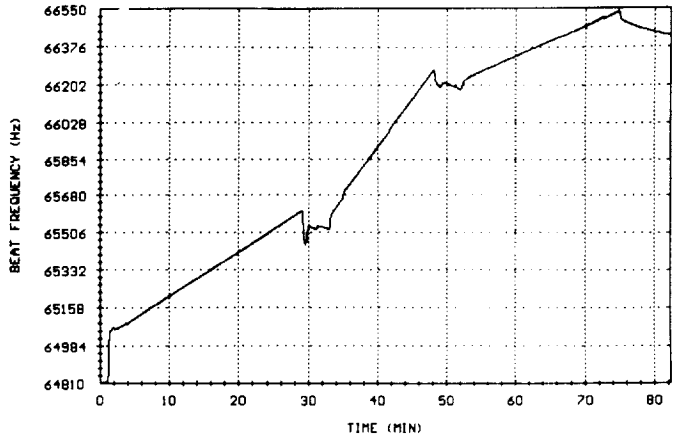
1. Zurer, P., *C & EN* 8 (1993)
2. Chuan, R. L., *J. Aerosol Sci.* 1, 111 (1969).
3. Andersen, H. H., and Bay, H. L., *J. Appl. Phys.* 46,1919 (1975).
4. King, W. D., Jr., *Analytical Chemistry.* 9, 1735 (1964).
5. Guibault, G. G., Tomita, Y., and Kolesar, E. S. Jr., *United Sates Air Force Technical* TR80-21, 1980.

Biography

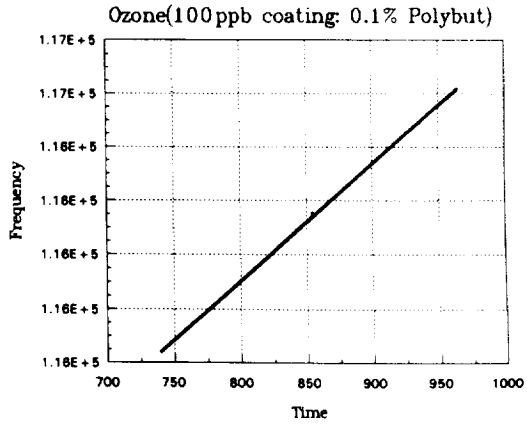
Cassandra K. Wright, a native of Indianapolis, Indiana, is a graduate student in the Chemistry Department at Howard University. She earned her B.S. degree, *Cum Laude*, in chemistry in 1994 from Alcorn State University. Currently, she is completing work for a Master's degree in chemistry. The title of her master's thesis is entitled, **"Calibration of the QCM/SAW for Detection of Ozone in the Upper Atmosphere"**.

Figures

(1) Response curve of PBD with ozone for changing concentrations



(2) Data and exponential fit of ozone with 100 ppb PBD
eqn: a + b(exp(-cx) -1)



DO PORPHYRINS AND METALLOPORPHYRINS OCCUR IN INTERSTELLAR DUST ?

Sabrina Cromer, Quintence Mays and Peter Hambright
Chemistry Department, Howard University, Washington, D. C., 20059

Sabrina Cromer from Mt. Vernon, NY, has a B. S. from Hampton University and is a second year graduate student. Quintence Mays from Dayton, OH, is a undergraduate senior chemistry major. Both do research with Dr. Hambright, under the auspices of the NASA sponsored Center for the Study of Terrestrial and Extraterrestrial Atmospheres at Howard University.

Introduction

Although it has been over sixty years since their discovery, the species responsible for the 127 Diffuse Interstellar Bands (DIBs) from the absorption spectra of reddened stars which lie behind clouds of interstellar dust (which absorbs, scatters, and polarizes radiation) are unknown entities [1]. The bands are termed diffuse because they are broader than the simultaneously observed atomic lines of Na and Ca. These DIBs are in the visible and near IR regions, falling between 4406 and 13175 A. Suggestions have been made that the DIBs carriers may be polycyclic aromatic hydrocarbons, polycarbon chains or or the polyhedron form of carbon, C_{60}, in their neutral, cationic or anionic forms. The DIBs have been divided into families and we focus on the 4428, 5780 and 6284 grouping, because many porphyrin types absorb near these wavelengths.

Johnson [2] has claimed that the molecules $(py)_2$Mg(II)-tetrabenzporphyrin (TBP) and its metal free analog H_2-TBP are responsible for 76 of the DIBs. Miles and Sarre [3] noted that the 6284.5 band observed in supersonic expansion work on the simplest reduced dihydro-porphyrin, the chlorin of porphin, is the source of the DIB at 6283.86. Since porphyrins are the basis of life on earth (hemoglobin, chlorophyll, Vitamin B-12), it would be of extreme interest if such cyclic conjugated tetrapyrrole pigments occur in interstellar dust. We have synthesized and obtained the spectra of a variety of porphyrins and metalloporphyrins in frozen glasses at 77 K in order to determine if their wavelengths and relative intensities approximate DIBs features.

Results and Discussion

Absorption spectra of the macro-cyclic pigments at uM concentrations were measured in clear glasses of 2-methyltetrahydrofuran, or 3:7 1-butanol / 3-methylpentane or 2:5:5 EtOH / iso-pentane / Et_2O at 77 K. No evidence of Shpolskii multiplets were noted in these studies, and such phenomena are usually found only at nanomolar levels in fluorescence studies [4]. When going from room temperature to liquid N2, the bands in general sharpen, occasionally split, and depending on the compound, their maxima are red shifted by 10 to 300 nm.

With our 3:7 mixture, the chlorin of porphin ($C_{20}H_{16}N_4$) was at 6335 compared with 6284.5 in the jet, giving a matrix shift of ~ -51 A. For this peak, our FWHM was ~69 A and ~3 A in the LIF jet. The strongest band of chlorin at 77 K is the Soret found at 3960, and since no DIBs lie in this range, this molecule is probably not a component of interstellar dust [1].

In benzene solutions containing pyridine, MgTBP exists as a mixture of MgTPB, py-MgTBP and $(py)_2$-MgTBP. At 298 K, the formation constant K_1 of the monopyridinate is 3160 M^{-1}, and

K_2 = 35.5 M^{-1} for the mono to dipyridine transformation. The enthalpy (-39 kJ / mol for K_1 and -32 kJ / mol for K_2) and entropy (-64 eu and -78 eu) values for both formation constants are negative, and thus the dipyridinate is favored at low temperatures in solutions containing moderate amounts of pyridine.

This porphyrin can be crystallized as the solid dipyridine adduct. When redissolved in solutions lacking pyridine, the coordinated ligands are labile and partially dissociate, and this free pyridine in dictates the relative concentrations of the aforementioned three species. As such, this explains the results of Donn and Khanna [5] who found that MgTBP in nujol mulls containing excess pyridine show blue spectral shifts (4360 to 4150) with time (days) at room temperature. The gradual evaporation of pyridine from the mull leads to the transformation of the monopyridinate into the unligated complex. The solid dipyridinate is stable to pyridine loss at room temperature [5]. Differential thermal analysis showed that one ligand was lost at 165 C and the second at ~ 500 C.

One of the most characteristic DIBs is at 4428, and in agreement with both Johnson [6] and Donn & Khanna, we find for the bis-pyridinate 4425 in 2-MeTHF, 4420 in 3:7 and 4380 in 2:5:5. As noted in Table 1, while seven other bands are also near interstellar lines, the relative absorbances do not particularly match the DIBs intensities. The metal-free H_2-TBP has strong Sorets at 4140 and 4290, while only a 4180 DIB is in this region. Johnson ascribes his 6175, 6591 and 6597 bands (frozen paraffin matrices) to H_2-TBP, while we observe 6030, 6105 and 6580 in this range under our conditions. Pyridine has yet to be found in space. Many of Johnson's 77 lines are of fluorescence origin, requiring porphyrins in excited states. Due to the limits of resolution of our spectro-photometer, we cannot experimentally duplicate his low concentration polycrystalline Shpolskii matrices, where the porphyrins lie in several spectro-scopically distinct sites.

The new porphyrins studied are listed in Table 1. They involve the basic porphyrin nucleus with four phenyls in the *meso* positions, para-substituted with -H, -OMe,-Me or 2,6-dichloro functions. All are derivatives of tetraphenylporphyrin (TPP). The di, tri and tetrabromo compounds are TPP with bromo groups on the beta-pyrrole positions. The tetramethylchlorin has four methyl groups in place of the phenyls, and is reduced at one beta pyrrole ring. The chlorin (of porphin) listed has four hydrogens substituted for these four methyl groups. Some of these porphyrins are metal-free with two protons in the center, or di-cations with four central protons, where the dipositive charge is balanced by two trifluoroacetate groups. Others are metallo derivatives with coordinated Mg(II), Cr(III), Fe(III), P(V) or O=W(V) atoms. Only listed are compounds that we found to have Soret bands in the 4400-4500 range. Hundreds of porphyrins and metallo-porphyrins have their Sorets between 4000 and 4200, and are not further considered.

Since 127 DIBs lines are known to date, almost any molecule with a number of visible and near IR bands will probably show DIBs features purely by coincidence under the conditions of this study using frozen glasses which produce large band width spectra. The results in Table 1 show that in addition to Johnson's $(py)_2$MgTBP, many porphyrin types absorb near 4428. In the most recent review of this area, Herbig [1] notes that the postulation of C_{60}, various PAHs, polycarbon chains or clean and dirty dust grains as the sources of the DIB lines are "either not supported by observations or await better data." In general, the results in Table 1 neither prove nor disprove that porphyrins are interstellar species. Supersonic expansion experiments on some of the compounds in Table 1 are

TABLE 1. SPECTRA OF PORPHYRINS AND METALLOPORPHYRINS, 77K

$(Py)_2$- Mg(II)TBP	
OBS	DIB
4165 (0.43)	4180 (1.2)
4295 (0.24)	------
4425 (5.2)	4428 (6.5)
5795 (0.06)	5795 (1.6)
5890 (0.1)	5849 (0.7)
6020 (0.03)	6010 (1.7)
6300 (1.0)	6307 (1.0)
6330 (1.0)	6310 (1.0)

$Cl_2P(V)TPP$	
OBS	DIB
4430 (29)	4428 (6.5)
5255 (0.2)	5359 (0.6)
5670 (1.3)	5705 (0.3)
6130 (1.0)	6113 (1.0)

Fe(III)-T(2,6-Cl)Br_4	
OBS	DIB
4205 (26)	4180 (0.2)
4445 (61)	4428 (11)
5735 (2.0)	5762 (0.7)
6045 (3.3)	6010 (2.9)
6370 (1.0)	6367 (1.0)

$O=W(V)TPP(OCH_3)_2$	
OBS	DIB
4460 (31)	4428(0.4)
5785 (1.0)	5780 (1.0)
6205 (0.8)	6203 (0.3)

ClCr(III)TPP	
OBS	DIB
3295 (3.0)	4066 (0.6)
4180 (1.5)	4180 (0.9)
4475 (22)	4502 (26)
5565 (1.0)	5545 (1.0)

H_4-T(4-CH_3)PP^{2+}	
OBS	DIB
4410 (7.6)	4428 (0.6)
6630 (1.0)	6614 (1.0)

H_4-T(4-OCH_3)PP^{2+}	
OBS	DIB
4495 (4.3)	4501 (2.9)
6930 (1.0)	6940 (1.0)

H_2-$TPPBr_4$	
OBS	DIB
4150 (35)	4180 (1.3)
4445 (84)	4428 (7.0)
5445 (4.0)	5450 (1.3)
5805 (3.5)	5809 (1.1)
6815 (1.0)	6811 (1.0)

H_2-$TPPBr_3$	
OBS	DIB
4170 (50)	4180 (1.3)
4435 (72)	4428 (7.0)
5425 (4.9)	5450 (1.3)
5815 (2.5)	5809 (1.1)
6825 (1.0)	6811 (1.0)

H_2-TBP	
OBS	DIB
3845 (0.5)	---------
4140 (3.7)	4180 (1.2)
4290 (5.8)	4180 (1.2)
4480 (0.2)	4428 (0.7)
5590 (0.1)	5544 (1.5)
5960 (0.8)	6004 (0.8)
6030 (1.1)	6010 (1.8)
6105 (0.6)	6113 (1.0)
6580 (1.0)	6597 (1.0)

H_2-Me_4CHLORIN	
OBS	DIB
4185 (29)	4180 (0.7)
5220 (1.7)	5359 (0.4)
5515 (1.3)	5508 (0.9)
5565 (1.3)	5544 (0.8)
6035 (1.0)	6010 (1.0)
6500 (3.0)	6494 (0.4)

CHLORIN	
OBS	DIB
3960 (2.5)	---------
4815 (0.1)	4779 (1.6)
4910 (0.2)	4882 (1.8)
5770 (0.04)	5772 (0.9)
5840 (0.05)	5843 (1.0)
5970 (0.03)	6004 (0.7)
6060 (0.04)	6089 (1.1)
6220 (0.1)	6234 (1.2)
6335 (1.0)	6318 (1.0)

Relative absorbances are in parentheses. TPP is ms-tetraphenylporphyrin; TBP is tetrabenzporphyrin; 4-CH3 and 4-OCH3 are the tetra(4-methylphenyl) or tetra(4-methoxyphenyl)porphyrins; T(2,6-Cl)Br4 is tetra(2,6-dichlorophenyl-2,4,6,8-tetrabromo)porphyrin; H2-TPPBr3 / 4 is TPP with 3 or 4 bromo groups on the beta-pyrrols; Me4-chlorin is 5,10,15,20 tetramethylchlorin and chlorin is the chlorin of porphin.

obviously desirable, and facilities are available at Howard for such work.

References

1. G. Herbig, Annu. Rev. Astrophys., 1995, 33, 19.

2. F. Johnson, Bull. Am. Astron.Soc., 1993, 25, 821.

3. J. Miles & P. Sarre, J. Chem. Soc. Faraday Trans., 1993, 89, 2269.

4. S. Volker & R. Macfarlane, J. Chem. Phys., 1980, 73, 4476.

5. B. Donn & R. Khanna, Astrophys. Space Sci., 1980, 68, 19.

6. F. Johnson, Proceedings of the symposium on The Diffuse Interstellar Bands, 1994, Boulder, Colorado, NASA Conference Publication 10144.

7. We thank Drs. Bert Donn, Fred Johnson , Prabhakar Misra and Arthur Thorpe for helpful discussions.

Modifications to the QCM/SAW Impactor For Data Acquisition

Julius R. Grant[1] and Arthur N. Thorpe
The Center for the Study of Terrestrial and Extraterrestrial Atmospheres
Howard University
Washington, D.C. 20059

Abstract

Consistent retrieval of in-flight data is an important function of studying aerosol particles in the stratosphere. Onset Computers' data-logger engine the Tattletale Model 7 controls the QCM/SAW impactor that CSTEA uses for its measurements in the stratosphere. This Model 7 is based on the MC68332 microprocessor. The QCM/SAW sensor was originally designed to measure particulates and gases in the plume of volcanoes in the stratosphere. CSTEA acquired the sensor on loan from NASA and reprogrammed it to measure plumes of aircraft exhausts in the stratosphere. CSTEA's field experience with the QCM/SAW has motivated modifications to the QCM/SAW to increase sensitivity and reliability.

Keywords: stratosphere, microprocessor

1. Introduction

The Quartz Crystal Microbalance Surface Acoustic Wave (QCM/SAW) instrument is an aerosol measuring device that takes data by mass loading Quartz crystal oscillators.[1] In a typical operational experiment for gathering aerosol particles in the stratosphere with the QCM/SAW impactor there are several processes that occur. The experiment is basically a three step process that includes the "on the ground prior to flight" step, the "in flight" step, and the "after flight" step. In a typical experiment before and during the flight, the instrument receives power to start various devices, then initiates the measuring process of the upper atmosphere, then stops measuring the upper atmosphere, and finally shuts down the power to the instrument. The instrument was originally designed to measure plumes from volcanoes.[2] Therefore, there was a need to modify the instrument for measuring exhaust wakes in the stratosphere. It is very important to note the old configuration and problems associated with them in order to understand the reasons for the modifications.

2. Original Design

The present QCM/SAW impactor was originally manufactured to measure the plume of volcanoes and is shown in the diagram in Figure 1.[3,4] The aerosol portion of the stack was configured, prior to reception by Howard University, with six QCM and 2 SAW stages which were the

[1] Currently pursuing a M.S. degree in Physics from Howard University. Primary research involves the study of magnetization of tertiary glasses.

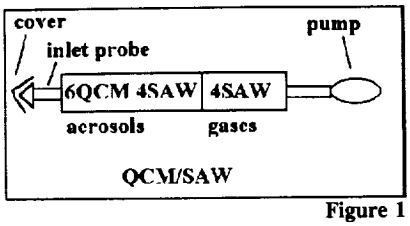


Figure 1

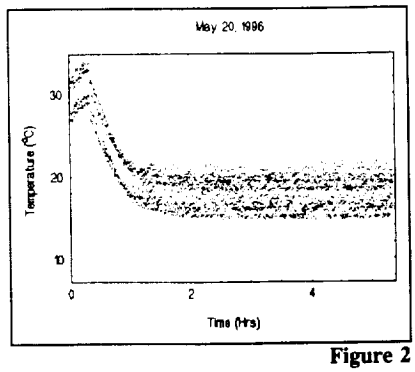


Figure 2

first eight stages of the instrument. There were an additional four SAW stages for stratospheric trace gas analysis. The SAW crystals were floating on springs and were broken when the stages were connected. Operationally upon receiving aircraft power, the Tattletale (TT7) would turn on the aerosol pump and constantly check to see if the pilot had pressed the pilot switch for the instrument to start taking data in the stratosphere. Once this switch was pressed the inlet cover would be jettisoned. Following this, the sampling valve would turn allowing atmosphere to flow through to the crystals. While the experiment is running, temperature data from several thermocouples is sent to the TT7. This digital signal was converted by the TT7 to a temperature reading which was stored. We found that the TT7 would lock up as a result of converting analog data to digital data at a rate exceeding twenty times per minute. Two heaters were used to control the temperature.

Figure 2 shows the temperature received by the he TT7 as a function of time for a long flight. It is easy to see that the data is unreliable and not reproducible. In the old configuration the TT7 controlled the failure lamp circuit and the jettisoning of the inlet cover. The failure lamp signal to the cockpit was connected from the TT7 to an optocoupler, then an amplifier, a driver circuit, then a relay that sends 28 V back to the aircraft. If the aircraft doesn't receive this 28 V then the failure lamp lights up which signals to the pilot that the experiment has failed. Flight preparation procedures did not allow the instrument to successfully operate and collect necessary data. There were problems that were encountered during flights, one problem was the inlet cover not being released. Another was that the flow of gas was not great enough to collect adequate data. In order for the QCM/SAW cascade impactor to be effective these problems had to be solved.

3. New Design

There was a great need to change the instrument from one that was designed to study volcano plumes to an instrument that would be useful for studying exhaust wakes in the stratosphere. Modifications to the control program, preparation procedures for the new program, temperature sensing, and airflow were necessary to enhance the QCM/SAW's reliability for operation.

The control program was replaced to allow better control of programming operations for the TT7. In the old program, the microprocessor was observing many

circuits. The new program only concentrates on logging data and detecting when it should be in sleep mode. The elimination of the watch dog restarter and the new program allows the TT7 to sleep while waiting for external power before collecting data. While sleeping, the instrument only uses 0.1 mA for sleeping and 0.2 mA while checking for the presence of power. In order to check the modifications in flight only six QCM aerosol stages are used while additional laboratory studies are being done on SAW stages. Although changes were made in the program changes were also made in the physical connections as well to totally isolate the microprocessor. (See Figure 3)

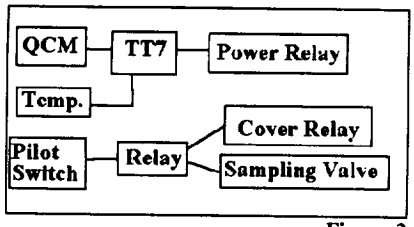


Figure 3

Modifications were made to insure that the pilot switch had no interaction with the TT7. The pilot switch goes directly to a relay which it triggers and the relay releases the inlet cover and controls the sampling valve. The microprocessor was given a new location on a new board dedicated just for it. The purpose for doing this because to supply only the minimal wires necessary. These wires are the power lines for the TT7 which are two(positive and ground),the communication lines which are four, and six lines for data collection, and one line for detection. The I/O connector for the laptop was replaced from the top circuit board to a new board that includes the microprocessor. The interface cable for the laptop had been attached to a location on the original board and attached to a new board. A modification was made to the signal wire that detected the presence of external power. This connection is needed for the microprocessor to remain in sleep mode until data collection is necessary. Once the Data Acquisition Box (DAB) receives external power a timer triggers a relay which grounds a special input of the TT7. The TT7 comes out of the sleep mode and begins to take data. The thermocouples and associated circuitry were replaced by a temperature sensitive voltage to frequency chip. This temperature sensor sends frequencies through the location where the signal for SAW #3 was detected. The conversion ratio of the chip is about 1mV/K. The use of two resistors and one capacitor are needed for its operation.

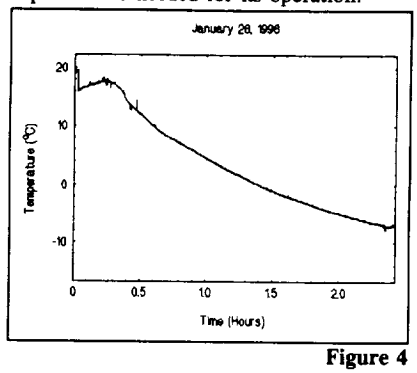


Figure 4

Figure 4 shows the operation of this sensor on an ER-2 flight and none of the previous problems associated with the thermocouples were observed. Sensitivity was changed from ± 1 degree Celsius to $\pm$ 0.1 degree Celsius. A second pump was added to the system for greater flow through the QCM/SAW. A new 28 V powerline along with a circuit breaker was added to power

the new pump and comes through the standard Power Circular Connector (POCC). The power supply was placed on the main frame a few inches away from the pump and two power wires were connected inside the DAB. The +28 V and ground lines were connected to the Pump Circular Connector (PUCC) and then directly through the POCC to the DAB.

4. Results of the Modifications

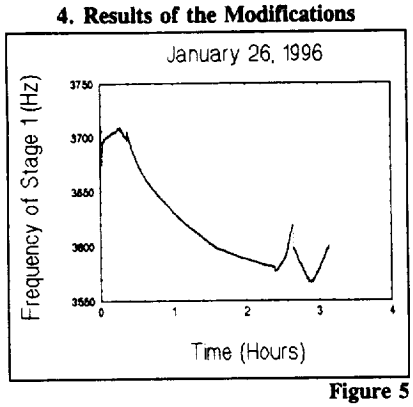


Figure 5

New modifications allowed us to change flight preparation procedures for NASA's ER-2 aircraft. Previously the procedure required us to download data on the aircraft ramp while providing additional external power excluding the power from the aircraft. The use of external power is no longer needed and data can be down loaded from inside the hangar. The modified QCM/SAW was flown on January 25th and 26th on NASA's ER-2. Each mission lasted approximately 3 hours. Data from the frequency counter on each flight shows no evidence of frequency shifting. A typical set of data collected from QCM stage 1 is shown in Figure 5. All data was collected with no abnormalities.

5. References

1. A. C. Hillier and M. D. Ward, "Scanning Electrochemical Mass Sensitivity Mapping of the Quartz Crystal Microbalance in Liquid Media", Analytical Chemistry, Vol. 64, no. 24, November 1, 1992.

2. R. L. Chuan and William Chiang, "Cascade Impactor Collects Airborne Samples for Particle Identification, Solid State Technology, July 1985.

3. Susanne V. Hering, "Calibration of the QCM Impactor for Stratospheric Sampling",Aerosol Science and Technology 7:257-274 1987.

4. W. D. Bowers and R. L. Chuan, "Surface acoustic-wave piezoelectric crystal aerosol mass microbalance",Rev. Sci. Instrum. 60 (7), July 1989.

New Polymer Coatings for Chemically Selective Mass Sensors

S. C. Sims, C. Wright, J. Cobb, T. McCalla, R. Revelle, V.R. Morris, and S. K. Pollack

Department of Chemistry
The Center for the Study of Terrestrial and Extraterrestrial Atmospheres
Howard University
Washington, D. C. 20059

Abstract

There is a current need to develop sensitive and chemically specific sensors for the detection of nitric acid for in-situ measurements in the atmosphere. Polymer coatings have been synthesized and tested for their sensitivity and selectivity to nitric acid. A primary requirement for these polymers is detectability down to the parts per trillion range. The results of studies using these polymers as coatings for quartz crystal microbalances (QCM) and surface acoustic wave (SAW) devices will be presented.

Introduction

Two of the most important scientific issues concerning stratospheric aviation are understanding what chemical and physical process in the atmosphere could be perturbed by aircraft emission and the long term climatic effects. The potential for negative impacts by nitrogen oxides (NO_x) in the exhaust of stratospheric aircraft was first pointed out by Harold Johnson over two decades ago.[1-2] The combination of anticipated ozone layer destruction and economic unfeasibility halted efforts to establish a commercial fleet of stratospheric aircraft in the early seventies. Improvements in aircraft engines and fuel have made the prospects of a commercial fleet of high speed stratospheric aircraft a practical venture, but the atmospheric impact remains nebulous. Current models predict that the photochemistry of the upper atmosphere is highly dependent on altitude and that increased NO_x at some altitudes may even

lead to a net ozone production.[1] The present state of knowledge is not sufficient for a conclusive plan of action to be developed.

The Stratospheric Wakes Analysis Project (SWAP) is trying to develop a scientific basis for the assessment of the impact of subsonic and supersonic aviation on the stratosphere. SWAP plans to measure aerosol size distributions and trace gas concentrations in the far wake region of the exhaust plume of the stratospheric aircrafts. A Quartz Crystal Microbalance/Surface Acoustic Wave (QCM/SAW) instrument is used to detect both aerosols and gases. The highlight of SWAP thus far has been a tandem flight involving two ER-2 airplanes, one equipped with the QCM/SAW used to track the wake of the lead ER-2 airplane. The data collected ranged from a few seconds to several minutes. These data represent the earliest sampling of a stratospheric wake. The success of this mission confirmed the feasibility of future studies of this sort for the characterization of wakes of high speed stratospheric aircraft.

The QCM/SAW is a cascade impactor that collects size fractionated distribution of aerosols on a series of 10 MHz quartz crystals, and employs SAW devices coated with a chemical sensors for the gas detection. We are calibrating several chemical coatings for SAW sensor applications in our laboratory for the detection of ozone, sulfur dioxide, nitric acid, and hydrochloric acid. The coatings for the nitric acid are the topic of this paper. Two chemical sensors have been used previously; nitron and poly-4-vinylpyridine-Co-styrene (P4VCS). The nitron coating was unstable and decomposed quickly. P4VCS yields a strong response to nitric acid, but it was not chemically specific. It tends to undergo reversible ion exchange reaction with hydrochloric acid and nitric acid in addition to and irreversible nitration with nitric acid.

The development of a nitric acid sensor is important because nitric acid is one of the major components total NO_y. The total reactive nitrogen (NO_y) is composed of all oxidized nitrogen compounds, those with and oxidation number two or greater[3] :

$$NO_y = NO + NO_2 + 2N_2O_5 + HONO + HNO_3 + PAN + \text{organic nitrates} + ClONO_2 + BrONO_2 + NOCl + \text{other halogen containing compounds} + \text{particulate nitrates}$$

There are several current instruments that measure other components of NO_y well.[4-5] The method used to measure nitric acid is the infrared diode laser. The diode laser is a high intensity narrow bandwidth infrared source. Complications arise from the demands for performing in-situ airborne absorption experiments of trace species in the 10-0.01 ppt concentration ranges.[6] Another reliable method for measuring stratospheric nitric acid would be useful.

Polymer chemical sensors are receiving a great deal of attention. They are cost efficient, tailored to have specific chemical properties, and can measure several different items simultaneously. The polymer coatings have been developed, whose reactivity is not based on ion exchange. The nitric acid sensitive polymer [Poly(styrene-co-3,4-dimethoxy styrene)] has been tested, and showed no interferences with carbon dioxide, sulfur dioxide, ozone, oxygen, hydrochloric acid, water, and nitrogen.

Experimental

A typical laboratory experiment consists of coating a QCM crystal with a dilute (0.1%) solution of polymer coating on the crystal. The initial tests were performed using the California Measurements QCM test stage and a gas delivery system. A single stage has two quartz crystals which oscillate in the shear mode. In the stage, the crystals are placed one above the other directly below the inlet nozzle (Figure 1). The QCM stack is composed of ten stages that

measure size-segregated fractions of aerosols and gases. The Metronics Dynacalibrator (Model 450) is a gas delivery system which employs permeation tubes for precise concentrations of selected calibration gases. Gaseous nitric acid was supplied using fuming nitric acid, the impurity in the nitrogen dioxide tank, and the standard Metronics permeation devices. The nitric acid permeation device is a mixture of nitric acid and water. The flow rate is controlled by a micrometering valve, monitored by a volume flow meter, and exhausted with a small mechanical rotary pump. The crystal is placed on the California Measurements QCM Cascade Impactor (Model PC-1) , coated down to a certain frequency using a Badger airbrush (Model 350), and stabilized. After the crystal is coated, a stream of nitric acid gas is pumped across the coated face of the crystals. During the exposure time, the characteristic mass loading due to the interaction between the gas and the coating causes a characteristic change in the beat frequency of the dual oscillator system. The time averaged response of the beat frequency corresponds to the concentration of the trace gas specific to the chemical sensor on the crystal. A typical response curve and its exponential fit is shown in Figure 2.

Summary and Future Directions

In the preliminary testing of poly(styrene-co-3,4-methoxy styrene) P2, we have determined the following:

1. P2 is stable and highly soluble in chloroform.
2. It has no reaction with carbon dioxide, sulfur dioxide, ozone, oxygen, hydrochloric acid, and water.
3. The reaction of P2 with nitric acid is irreversible.
4. P2 is sensitive down to 10 ppb for the 0.1% solution of P2 in chloroform.

The interferences of other nitrogen oxides with P2 are still under investigation.

Acknowledgments

I would like to acknowledge my advisor, Dr.Vernon R. Morris, and the entire SWAP team for support of this work and insightful encouragement and guidance. I would also like to acknowledge NASA #NAGW-2950.

References

1. Johnston, H. S.; Kinnison, D. E.; Wuebbles. D.J. J. Geophys. Res. 1989, 94, 16351.
2. Johnston, H. S.; Prather, M. J.; Watson, R. T. NASA Ref. Pub. 1250, 1991.
3. Crosley, David R. SRI International. 1994.
4. Fahey, D. W.; Kelly, K. K.; Ferry, L. R.; Poole, L. R.; Wilson, J. C.; Murphy, D. M.; Loewenstein, M.; Chan, K. R. J. Geophys. Res., 1989, 94, 11299-11315.
5. Gao, R. S.; Keim, E. R.; Woodbridge, E. L.; Ciciora, M. H.; Proffitt, M. H.; Fahey, D. W. J. Geophys. Res. 1994, 99, 20673-20681.
6. May, R. D.; Webster, C. R. J. Geophys. Res. 1989, 94, 16343-16350.

Biography

Shondalette C. Sims is a graduate student in the Department of Chemistry at Howard University. She earned her B.A. degree in chemistry from Talladega College in 1994. Her research experience includes the measurements of projection efficiency and projection color at the Xerox Corporation, and working with high energy density materials at Clark Atlanta University. Her graduate research involves the calibration of polymer coatings for nitric acid detection which are chemically specific and selective.

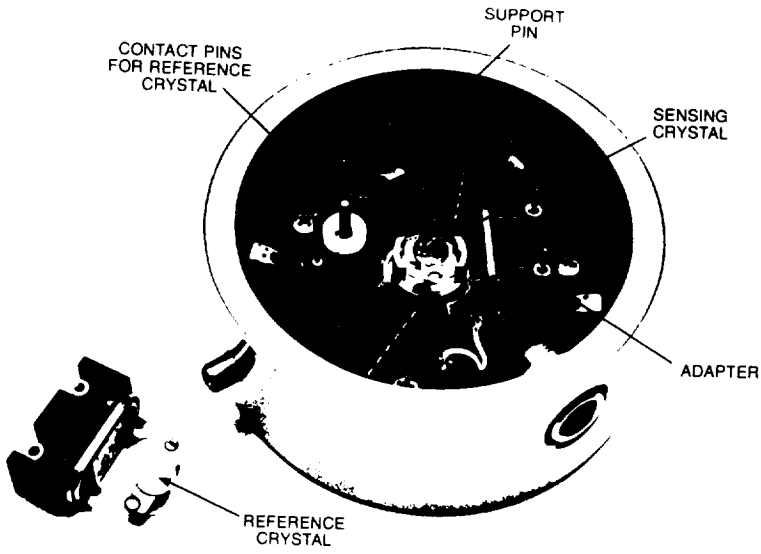


Figure 1-- Sensing Stage Housing

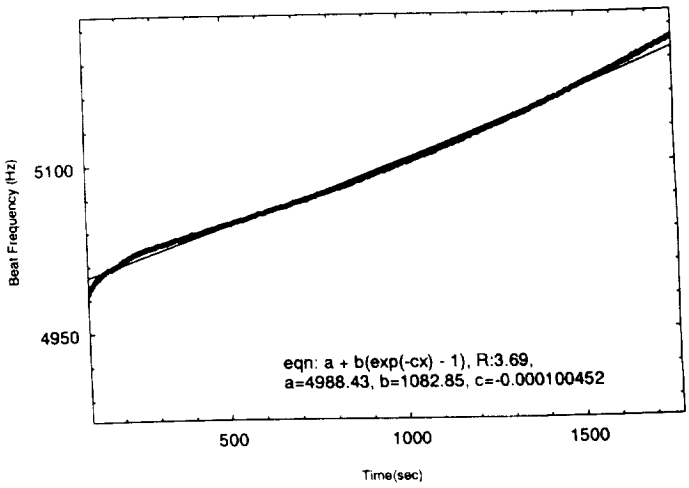


Figure 2--Response Curve for nitric acid on P2

290-95 P4

Measurement of Nitric Oxide in the Thermosphere and Mesosphere

A. Patrick[1], K. Kumar

The Center for the Study of Terrestrial and Extraterrestrial Atmostpheres
Howard University
Washington, D.C. 20059

Abstract. The Atmospheric Trace Molecule Spectrometer (ATMOS) is an instrument used to obtain solar occultation spectra, in the IR, of the earth's atmosphere from approximately 20 km altitude to 150 km. The instument has flown as a part of the suite of instuments on the ATLAS missions of the space shuttle. During the ATLAS 2 mission in late April 1993 ATMOS obtained a number of spectra in the 5.3 μm spectral region which includes the rotational lines of the fundamental (1-0) band of the NO molecule. This investigation is concerned with the analyses of these spectra to obtain the mixing ratios of nitric oxide in the mesosphere and lower thermosphere. The procedures for the measurement of slant column densities and their subsequent inversion to derive the VMR's will be described. The resulting NO concentrations will be compared with the predictions of 1-D models of atmospheric chemistry.

Keywords: nitric oxide, thermosphere, mesosphere, ATMOS, ATLAS 2

The Atmospheric Trace Molecule Spectroscopy (ATMOS) instrument is a shuttle based spectrometer which views sunlight through the earth's atmosphere and records the absorption at sunset and sunrise. It records the absorption of the sunlight in the spectral range 2 μm to 16 μm. The absorption spectra are taken between 20 km to 150 km tangent altitude above the earths surface with an altitude resolution of about 2 km. The occultation spectra taken during the ATLAS 2 (April 93) mission have been analyzed, to (i) obtain the density of nitric oxide (NO) molecules in the lower thermosphere and mesosphere and (ii) compare with correlative measurements of NO obtained the Halogen Occultation Experiment (HALOE) on board the Upper Atmospheric Research Satellite (UARS). The spectra in the 5.3 micron region (1900 cm^{-1}) obtained by ATMOS with spectral resolution of about

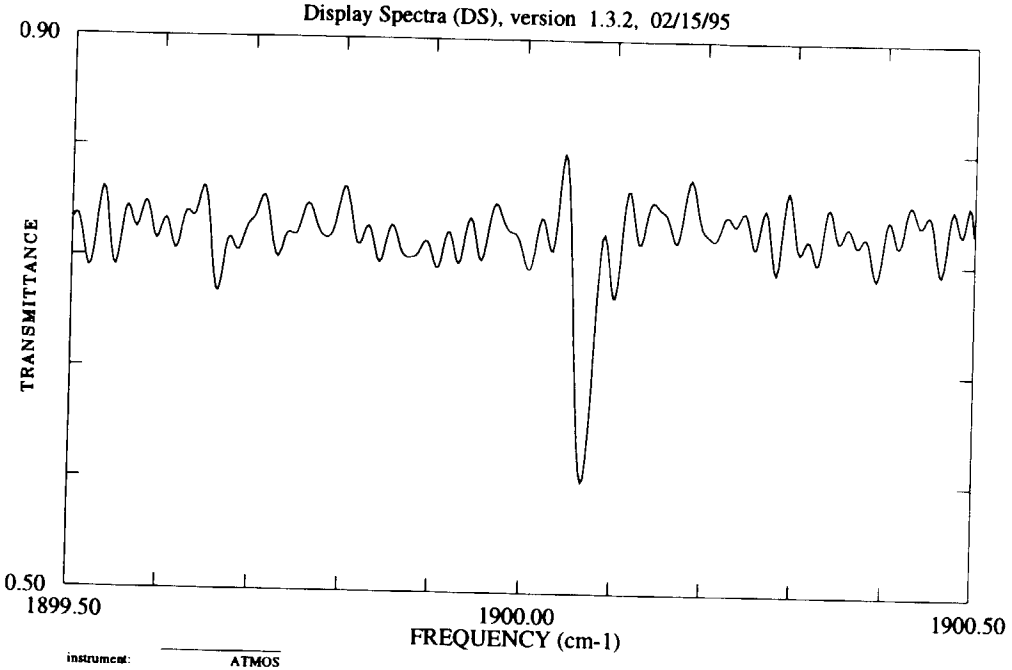
1. A student at Howard University pursuing a MS degree in Physics. Research interest involves computational simulation of neural networks. Primary research interest involve the study of nitric oxide in the earth's atmosphere.

0.01cm^{-1} were analyzed, as briefly described below.

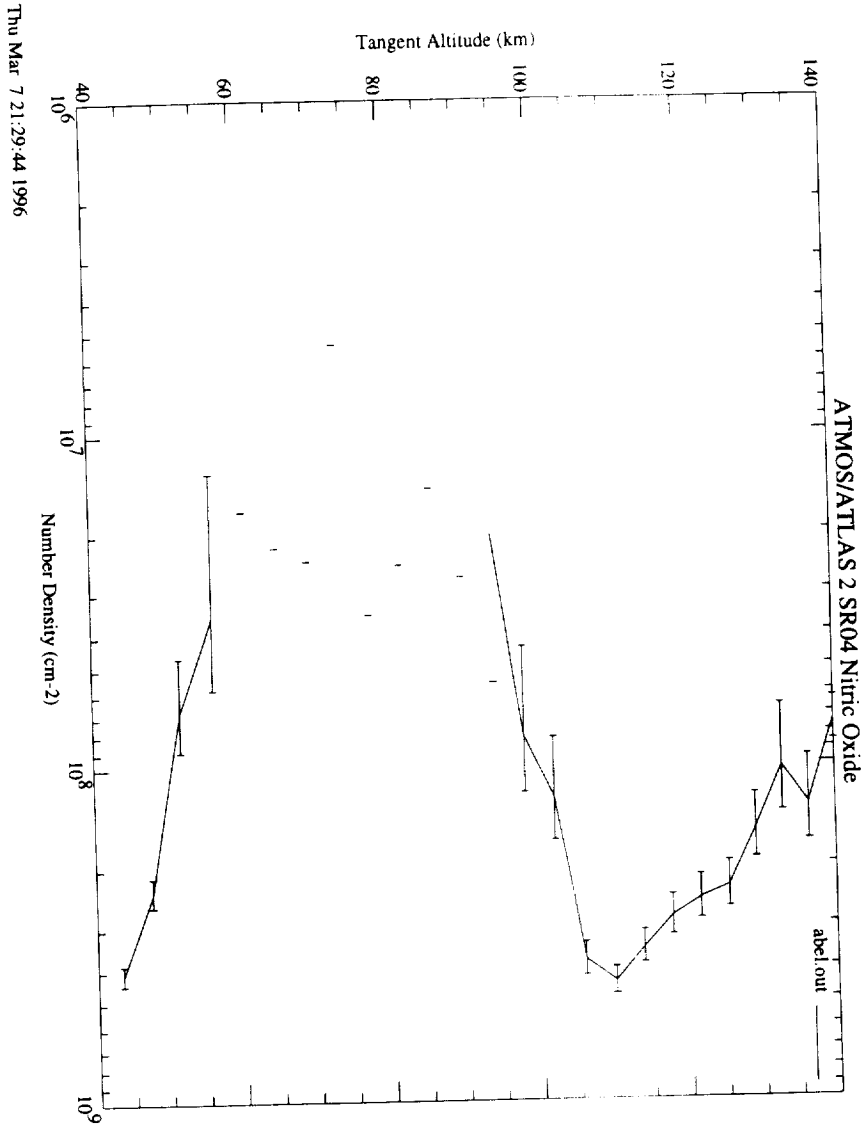
During 12 occultations the spectrum in the 5.3 micron region (~1900 cm^{-1}) was taken at tangent altitudes and various latitudes from 20km to 150km. The spectra was visually examined by the use of display spectra (DS) software routine which is resident at JPL. In particular, the strength of a line of NO at 1900.08cm^{-1} was examined as shown in figure 1. The highest altitude at which the above line was visible in all of the occultations above the noise in the spectrum was found at 140km.

By using occultation display spectra (ODS) software routine which is resident at JPL the slant column densities were obtained as a function of tangent altitude. ODS models the absorption at each of NO's spectral lines. The column density is adjusted to match the transition profile of the absorbed spectral line. There were 11 spectral lines, belonging to the (1-0) transition of NO, selected for further analysis. In each occultation 11 column densities and their associated errors were obtained at each altitude. It was found that the column densities at a given altitude agreed with measurement errors indicating that there was no systematic error. The average of the 11 values at each altitude was taken and the associated error in the mean was calculated. Thus the slantpath column densities of NO at the altitudes measured during an occultation were derived.

These slant path column densities were used to obtain the number density of NO at each altitude for each occultation. This was done by the use of retrieval software using abel inversion algorithms. The number densities was obtained as a function of tangent altitude using a constrained first difference method. NO's number density was finally plotted and showed a peak at about 110 km as shown in the figure 2. This research has shown that past studies based on the UV gamma bands of NO and what current models of NO chemistry predict are lower by a factor of 3. It is evident that the chemical models need to be revised significantly.



instrument:	ATMOS
run number:	r930410012720
apodization:	0
interpolation:	3
trans. max.:	0.9000
trans. min.:	0.5000
freq. shift:	-2.1888E-05
pt. spacing:	0.0025
obs. alt.:	306.0811
zenith ang.:	0.0000
tangent ht.:	105.3950
Phys. model:	


Thu Mar 7 21:29:44 1996
Tangent Altitude (km)
40
60
80
100
120
140
10^6
10^7
10^8
10^9
Number Density (cm-2)
ATMOS/ATLAS 2 SR04 Nitric Oxide
abel.out

Construction of a visible diode laser source for free radical photochemistry and spectroscopy experiments

Bronjelyn Newman and Joshua B. Halpern

Center for the Study of Terrestrial and Extraterrestrial Atmospheres
&
Department of Chemistry
Howard University
Washington, DC 20059

ABSTRACT: Tunable diode lasers are reliable sources of narrow-band light and comparatively cheap. Optical feedback simplifies frequency tuning of the laser diodes. We are building an inexpensive diode laser system incorporating optical feedback from a diffraction grating. The external optical cavity can be used with lasers that emit between 2 and 100 mW, and will also work if they are pulsed, although this will significantly degrade the bandwidth.

The diode laser output power and bandwidth are comparable to CW dye lasers used in kinetics and dynamics experiments. However, their cost and maintenance will be much less as will alignment time. We intend to use the diode lasers to investigate CN and C_2 kinetics as well as to study dissociation dynamics of atmospherically important molecules.

INTRODUCTION

Tunable diode lasers have been extensively used in atomic physics [1], and are beginning to be used for molecular spectroscopy [2], especially LIDAR applications [3]. They are reliable sources of narrow-band light and much cheaper than tunable dye lasers or solid-state lasers. Laser diode tuning depends on temperature, current, cavity size and fabrication among other things [4]. Their bandwidth is of the order of tens of Mhz. Individual diodes can tune over 10 nm, with many mode hops at repeatable, but unpredicable places. Optical feedback greatly improves laser frequency control.

We are building a simple, inexpensive diode laser system using diffraction gratings for optical feedback [4,5]. This system can be used with 2 and 100 mW CW and pulsed laser diodes, although with significantly degraded bandwidth in the latter case.

The laser system has four basic components, a diode laser, a collimating lens, a diffraction grating and a mirror. These components are mounted on a stable, vibration isolated baseplate.

We are building both Littrow and Littman type external cavities and intend to compare them at a later date. At present we are concentrating on completing and testing the Littman type external cavity system.

MacAdam, Steinbach and Wieman have provided a detailed description of how to build a small, external Littrow cavity diode laser [5].

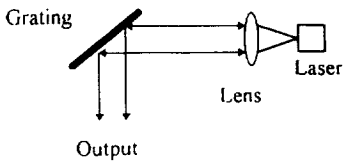


Figure 1. Littrow external cavity diode laser system.

As shown in Figure 1, for a Littrow cavity, the diffraction grating is mounted so that the light diffracted at a selected wavelength is fedback to the laser. The grating is one end of the external cavity, and the front end of the diode the other. For dye lasers [6], which are scanned over long regions, the Littrow mount grating, is rotated by a sine drive. In this case, where the diode is tuned over a small region, the grating can be coarsely tuned by the horizontal adjustment on the mirror mount. The output is the zeroth order grating reflection. To finely tune the laser frequency, a piezoelectric disk is used to rotate the grating through a small angle in response to an applied voltage.

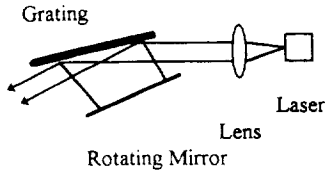


Figure 2. Littman external cavity is selected by moving this mirror, which is mounted on a rotation stage.

As shown in Figure 2, the Littman design places the grating at a very steep angle (>85°) to the laser, so that all of the lines on the ~ 25 mm long grating are covered by the laser beam [7]. The zeroth order specular reflection is again used as the output, but instead of tuning by rotating the grating, dispersed light from one of the grating orders is reflected back by a mirror placed opposite the grating. Single mode tuning requires careful adjustment of the laser to grating distance to match the distance between the tuning mirror and the point about which it is being rotated [8].

Advantages of the Littman (grazing incidence) design include higher dispersion and a spatially fixed output, which does not wander as the grating is tuned, as in the Littrow mounting. The principle disadvantages of the grazing incidence design for dye laser cavities, high output superfluorescence and a more difficult alignment for long scans, are not relevant to diode laser external cavity systems.

CONSTRUCTION

The diode lasers are purchased in TO-3 electronic packages. They are held in a Optima CDL3600 diode mount, with a 5 mm f1 lens that can be adjusted to collimate the laser beam. This allows us to very rapidly change the diodes.

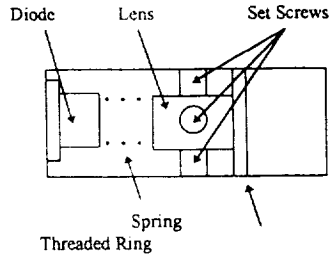


Figure 3. Optima collimated diode laser mount. The set screws adjust the lens position perpendicular to the laser. The lens position is adjusted along the laser beam by the threaded ring.

The assembly is temperature controlled to stabilize laser frequency, and to prevent diode thermal failure. The normal operating range is 15 - 25 K. Lower temperatures extend the tuning range and diode lifetime

The laser diode and lens assembly is a 10 mm diameter 20 mm long tube, held in a hole drilled at the end of a 37 mm, by 65 mm brass plate, roughly 6 mm thick. This plate and a 9 mm thick, 37 x 40 mm brass plate sandwich a 30 x 30 mm Peltier cooler (Melcor Corp. SH-1.0-127-05L). Heat sink compound improves thermal contact between the elements. The 9 mm plate is fastened to the front of a New Focus 9809 mirror mount.

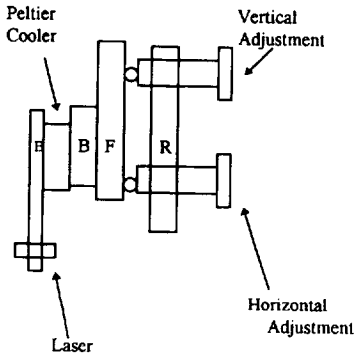


Figure 4. Laser diode mount. B indicates the brass plates and R and R are the front and rear plates of the mirror mount.

A 12.5 x 25 mm grating (Edmunds Scientific, 600 and 1200 l/mm)) blazed at 750 nm is held on another mirror mount. It is mounted at a high angle to the laser beam. The entire grating is covered by the laser light, maximizing dispersion. The laser output is the zeroth-order beam with the blaze pointing toward the output beam. Most of the light is coupled into the output. Very little is needed for the optical feedback.

Gratings can be easily damaged and should be handled as little as necessary and then with great care. Nothing should ever touch the grating surface, and it should not be cleaned, even when dirty.

The first order beam from the grating is reflected back by a high reflectivity dielectric mirror on an Aerotech ARS301 rotation stage. Coarse tuning is done by rotating the stage. A New Focus Model 3301 picomotor drive which replaces one of the mirror mount adjustment screw is used for finer tuning.

Finally, a piezoelectric disk is inserted between the picomotor and the mirror mount which can rotate the grating about a vertical axis and alter the cavity length electronically. These are thin 25 mm diameter brass disks with a small piezoelectric slice in the center, glued to the brass with adhesive. Voltage applied to the element induces a stress which forces the backing to "dish" on the opposite side. Two elements back to back, double the displacement of a single one.

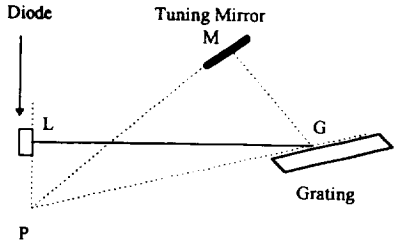


Figure 5. Littman cavity alignment for single mode scanning.

Figure 5 shows how the various component have to be adjusted for single mode scanning [8]. The "pivot" point P, for rotating the tuning mirror is the intersection between a line tangent to the laser's front facet and a line tangent to the grating. To maintain single mode scanning, the distance between the front facet of the diode laser and the

center of the grating , l_R, has to equal $l_p \sin \theta_{in}$, where l_p is the distance between the pivot point and the grating center and θ_{in} is the angle LPG in figure 5. Currently, "fine tuning" of the cavity via the piezoelectric elements has a shorter pivot, but the two arcs should be close enough over the less than 10 nm tuning range of a diode.

To reduce jitter from cavity vibration the baseplate can be set on a soft rubber cushion. To avoid air currents interfering with the temperature control the entire laser system can be placed in an insulated metal box. To operate the laser at reduced temperatures, the box can be purged with dry gas to avoid condensation.

We have built an aluminum enclosure to hold the laser. It has a sufficient thermal mass and conductivity to aid in temperature stabilization. The lid can be removed easily so that the cavity can be adjusted or the laser diode changed. Holes have been drilled in the box to allow for the laser beam to exit. Various electrical connections are provided for the Peltier cooler(s), the diode driver electronics and the picomotor/piezoelectric disk tuning elements. Initially we will operate the laser outside the box for testing.

Precise temperature control of the baseplate and the diode laser is essential for the long term wavelength stability. The same type of servo-driver in a feedback loop is used to control all Peltier elements. The sensing element for the servo are small thermistors in bridge circuits. The amplified and filtered error signal drives the thermoelectric-coolers.

A stable low-noise current source is also needed to operate the laser. We have both a commercial, older SDL 800 control unit from Spectra Diode Labs and control units built from circuits published in Reference 4.

CHARACTERIZATION

Several instruments will be used to characterize the laser After initial alignment of the cavity, the output wavelength of the laser should be within about 2 nm of the wavelength specified by the manufacturer, and near the center of the tuning range.

For accurate measurement of the frequency we have a Burleigh, near IR wavemeter and to characterize the mode structure of the laser, a Newport Model 6124 SuperCavity Fabry Perot

1. J.C. Camparo, Contemp. Phys. 26 (1985) 443.
2. L. Lundsberg-Nielsen, F. Hegelund and F.M. Nichaisen, J. Mol. Spectrosc., 162 (1993) 230.
3. J. Rall, D Cornwell, private communication, Goddard Space Flight Center.
4. C.E. Wieman and L. Holberg, Rev. Sci. Inst., 62 (1991) 1.
5. K.B. MacAdam, A. Steinbach and C. Wieman, Rev. Sci. Inst., 60 (1992) 1098.
6. T.W. Haensch, Appl. Opt., 11 (1972) 895.
7. M. Littman and H. Metcalf, Appl. Op., 17 (1978) 2224
8. K. Liu and M Littman, Opt. Lett., 6 (1981) 117

Bronjtelyn Newman is a resident of Ocean Springs , Mississippi who attended Ocean Springs High School. She is also a graduate of Alcorn State University with a major in chemistry. Currently Ms. Newman is attending Howard University seeking a Ph.D in Physical Chemistry. After graduation she hopes to work in industry.

Financial support for this work was provided by the Center for the Study of Terrestrial and Extraterrestrial Atmospheres under NASA Grant NAGW-2950.

FTIR Spectroscopy of HNO_3 and NO_2 Relevant to Stratospheric Wake Analysis

Rafiu A. Abina and Prabhakar Misra
Center for the Study of Terrestrial and Extraterrestrial Atmospheres
Laser Spectroscopy Laboratory
Department of Physics and Astronomy and Department of Mechanical Engineering
Howard University
Washington, DC 20059

Hideo Okabe
Department of Chemistry
Howard University
Washington, DC 20059

P.M. Chu
and
Robert L. Sams
Gas Metrology Research Group
Analytical Chemistry Division
Chemical Science and Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899

Abstract

The Fourier Transform Infrared (FTIR) technique has been employed to measure absolute concentrations of nitric acid (HNO_3) and nitrogen dioxide (NO_2) with 1 cm^{-1} resolution and an absorption pathlength of 4 m under quasi-static and flow conditions at atmospheric pressure and room temperature. Water features seen under quasi-static conditions diminished in intensity under flowing conditions. Nitric acid was observed in the 1660-1760 cm^{-1} range, while nitrogen dioxide was detected both in the 1536-1660 and 1213-1400 cm^{-1} ranges. Concentrations of nitrogen dioxide and nitric acid were determined to be 11.9 and 4.35 parts per million (ppm), respectively, with an uncertainty of 0.2 ppm. Experiments are underway with a 10 m cell to measure the absorption of nitric acid, water, sulfur dioxide, hydrochloric acid and ammonia on various materials such as glass, teflon, stainless steel and aluminum used for implementation of the flow system. Such materials will be used for the measurements of stratospheric trace gases by the Quartz Crystal Microbalance (QCM) and Surface Acoustic Wave (SAW) devices.

Keywords: FTIR; Trace Gases; Stratosphere; Wake; QCM-SAW

Introduction

Emissions from supersonic aircraft include water, carbon dioxide, nitrogen oxides and aerosol particles that participate in atmospheric photochemistry and radiative phenomena.[1] Nitrogen oxides (in the

[1] Rafiu A. Abina is pursuing a double major in Mechanical Engineering and Pure and Applied Mathematics and his research interests cover aerospace engineering and spectroscopy. He is a Junior and expects to graduate in May 1998.

form of NO and NO_2) and nitric acid (HNO_3) are present in significant nonequilibrium levels at the exit plane of the engine nozzle.[2]

The Fourier Transform Infrared (FT-IR) spectrometer fingerprints the identifiable chemical species and thereby provides a signature of the molecule. The spectra are recorded by either of two methods: static sample or the constant flow method. Intensities of prominent FTIR spectral features are measured and compared to determine the concentration of each sample (e.g. HNO_3 and NO_2) based on the spectral line intensity. Concentration estimates using this technique have been used for various samples of nitric acid. However, since nitric acid and water adsorb on the walls of the glass cell, their concentrations must be determined by a flow system.

Experimental

An FTIR spectrometer was used for recording the spectra in conjunction with a 4 m pathlength multipass absorption cell employing a flow system that used 100 cc/min of atmospheric air. Initially, zero air was used in the flow system to remove impurities that may be absorbed on the cell walls, such as water (H_2O), ammonia (NH_3), hydrochloric acid (HCl) and sulfur dioxide (SO_2). A given sample of NO_2 contains about 20 % of HNO_3 as impurity, because of the reaction of NO_2 with H_2O on the walls of the lecture bottle used for storing the sample. The FTIR spectra were recorded with a resolution of 1 cm^{-1} and a total of 50-100 scans were needed with the 4 m cell, whereas 1000 scans were required for the 20 cm pathlength cell to achieve comparable resolution.

The flow technique for recording FTIR spectra is being implemented using a 10-m cell, which would increase the sensitivity of the measurements. Such flow measurements will allow one to determine small concentrations (at parts per million level) of reactive species such as HNO_3, H_2O, NH_3 and SO_2.

Results

The HNO_3 feature was observed in the spectral range 1660-1760 cm^{-1}, while the NO_2 transitions were seen both in the 1536-1660 cm^{-1} and the 1213-1400 cm^{-1} ranges. The H_2O features were seen over the entire region 1213-1845 cm^{-1} in the quasi-static mode as shown in Fig. 1. These H_2O spectral lines were substantially diminished in intensity when a flow system was used as illustrated in Fig. 2. The HNO_3 transition around 1700 cm^{-1} showed a characteristic Q-branch appearance, and NO_2 around 1330 cm^{-1} also exhibited a prominent Q-branch. The concentrations determined by the FTIR spectrometer using the 4 m cell were 11.9 ± 0.2 ppm for NO_2 and 4.35 ± 0.2 ppm for HNO_3. Oscillator strength considerations were used for determination of these FTIR concentrations of HNO_3 and NO_2. These results were confirmed by the chemiluminescent detection (CD) technique.[3] For example, the CD value for NO_2 was 11.5 ± 0.2 ppm. Nylon can be used to remove HNO_3 preferentially from the mixture of NO_2 and HNO_3 using the CD technique, so that precise concentrations of both species can be obtained.

References

1. D.W. Fahey et al., Science **270,** 70 (1995).

2. G. Ritzhaupt and J.P. Devlin, J. Phys. Chem. **95,** 90 (1991).

3. L.J. Nunnermacker, R.R. Dickerson, A. Fried and R. Sams, Environ. Technol. **23,** 106 (1989).

Acknowledgment

Financial support for this work through the Center for the Study of Terrestrial and Extraterrestrial Atmospheres (NASA NAGW-2950) is gratefully acknowledged.

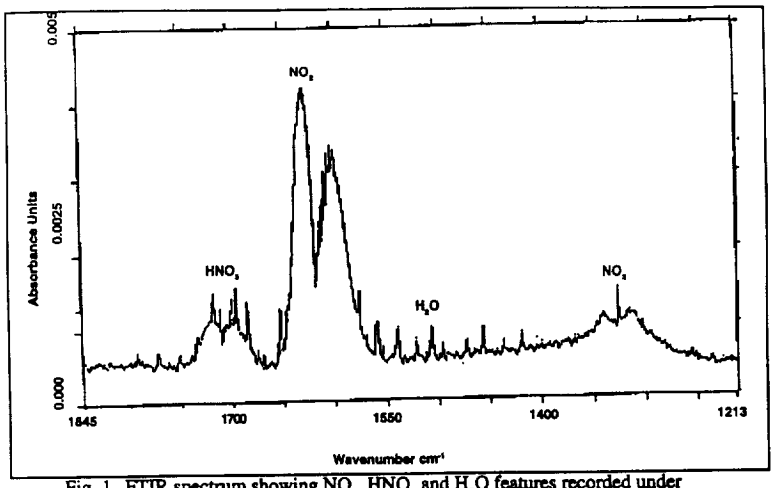


Fig. 1. FTIR spectrum showing NO_2, HNO_3 and H_2O features recorded under quasi-static conditions at 1 cm^{-1} resolution using a 20.7 cm cell and a gas pressure of 747.4 Torr.

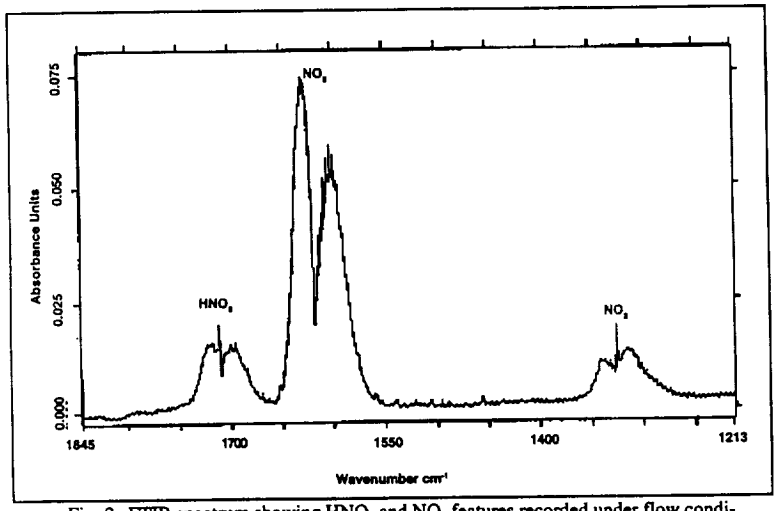


Fig. 2. FTIR spectrum showing HNO_3 and NO_2 features recorded under flow conditions at 1 cm^{-1} resolution using a gas pressure of 751.2 Torr. (Note: the water features go away.)

Modeling Standing Tidal Oscillations in Titan's Lower Atmosphere

Edward Colón[1] and James Lindesay

The Center for the Study of Terrestrial and Extraterrestrial Atmospheres
Howard University
Washington, D.C. 20059

Abstract

An investigation is made into the excitation of large-scale tidal oscillations within the atmosphere of Titan. Essentially, there exists both time-dependent and a time-independent components of tidal variability. A stationary equilibrium tide results from unperturbed tidal potential induced by Saturn[1]. Titan's highly eccentric (ϵ=0.0289)[2] orbit manifests itself in the form of a time-dependent first order perturbation in tidal potential. A numerical integration scheme was utilized to quantify the magnitude of air velocity divergence in Titan's atmosphere up to 575 km. This data is necessitated by Laplace's tidal equations which are employed in the calculation of mode-dependent barometric fluctuation and vertical air velocity for the time-dependent eccentricity induced tide. Systematic error analysis was performed with multiple procedures. The time-dependent, eccentricity-induced tidal perturbation was found to be of the order of 0.02 mbar for a zero zonal wave number. Overall implications will be examined.

Keywords: atmospheric tides, Titan, tidal potential, eccentricity, synchronous orbit

1. Introduction

Large-scale barometric fluctuations otherwise known as "atmospheric tides" are produced by two distinct mechanisms, namely, (a) differential gravitational forces exerted by local gravitating masses, and (b) the heating of various layers of an atmosphere by means of thermal absorption. These pressure oscillations are generally so small that they are totally obscured by localized weather conditions. In order to obtain a clear picture of the magnitude and periodicity of terrestrial tides, long term barometric observations are essential in isolating a true signal from the background noise.

Pioneering work in this field of study was performed by Laplace who wrote his influential text *Mécanique Céleste*. Much of the foundation of tidal theory is illuminated in this work. Laplace's theoretical calculations indicated a lunar atmospheric tide with a barometric pressure range in the tropics of 1/2 torr (0.67 millibars)[3]. Although his predictions were at least ten times too large in relation to observed data, they did spark Hough, Lamb, Bartels, Wilkes and others to pursue similar studies. Despite the great strides made in the dynamic analysis of atmospheric tides, there is still much to ascertain concerning the limitations of current models.

---

1 Currently pursuing a PhD degree in Physics from Howard University. Additionally, serving as an intern at NASA/GSFC Climate and Radiation Branch. Research involves simulating tropical intraseasonal oscillations using general circulation models. Primary research interests involve the study of atmospheric oscillations and related phenomena.

Thermal tides would be the dominant tidal component for all the terrestrial planets which possess appreciable atmospheres. However, as one examines the Jovian worlds with their accompanying satellite systems, gravitationally induced tides would play an increasingly significant role. Their vast distances from the sun dictate that they receive less direct solar radiation per square unit surface area than would be received by a terrestrial planet. Gravitationally induced tides may then tend to overwhelm thermal tides; the reverse of the situation occurring among the terrestrial planets. It may be asked what the useful purpose is in examining atmospheric tides on these distant worlds. Ioannou & Lindzen[2] suggest that the excitation of the outer planets by the gravitational tidal potential of their satellites may prove to be a useful probe of their planetary structure.

These planets are subjected to a forcing of known magnitude and frequency, and observation of the atmospheric response provides information about the mean thermodynamic structure. The converse of this argument may also be applied to the atmospheric structure of the aforementioned satellites. The paucity of information regarding dynamic responses can thus be negated with the use of tidal analysis.

2. Tidal Potential

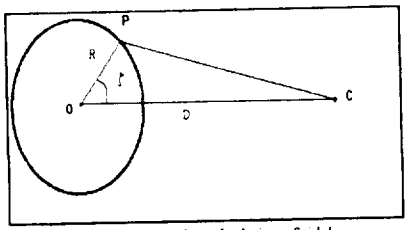


Figure 1. Geometry for the calculation of tidal potential; O is our planet's center, C is the location of our perturbing mass M, P is the location of our particle, and ζ is the angle between the position vector of our particle and the position vector of our perturbing mass.

In order to facilitate further investigation, we should examine an idealized planet-satellite system in which a significant perturbing mass will generate tides in the atmosphere of our planet. The geometry of our problem dictates that a particle in the atmosphere of our idealized planet will experience a potential

$$\Phi = \frac{GM}{\sqrt{D^2+R^2-2DR\cos\zeta}} \qquad (1)$$

where R is usually very small compared to D. For example, $R/D = 1.66 \times 10^{-2}$ for the Earth-Moon system and $R/D = 2.11 \times 10^{-3}$ for the Titan-Saturn system.

These facts allow us to expand to denominator of (1) using the binomial theorem as follows

$$\Phi = \frac{GM}{D}\left[1-2\left(\frac{R}{D}\right)\cos\zeta+\left(\frac{R}{D}\right)^2\right]^{-1/2} \qquad (2)$$

recognizing that the $(R/D)^j$ factors are merely Legendre polynomials, $P_j(\cos\zeta)$, and satisfy the recurrence relation:

$$P_{j+1}(x) = \frac{2j+1}{j+1}xP_j(x) - \frac{j}{j+1}P_{j-1}(x). \qquad (3)$$

One can now write the expansion in a more compact form. The third term is the tidal term in our analysis. It takes the form:

$$\Phi_2 = \frac{GMR^2}{D^3}\frac{1}{2}(3\cos^2\zeta-1) = \frac{GM}{D^2}\frac{1}{2}(2x^2-z^2) \qquad (4)$$

in cartesian coordinates. The force contributed by Φ_2 is:

$$F_2 = -\nabla\Phi_2 = \frac{GM}{D^2}[-2x\hat{i}_x+\hat{z}_z] = \frac{GMR}{D^2}[2(\cos\zeta)\hat{i}_x-(\sin\zeta)\hat{i}_z]. \qquad (5)$$

The tidal potential in terms of increasing powers of R/D can therefore be stated as:

$$\Phi_t = \frac{GM_t}{D}\left(\frac{R}{D}\right)^2 \sum_{j=2}^{\infty} \left(\frac{R}{D}\right)^j P_j(\cos\zeta) \qquad (6)$$

from which tidal force $\mathbf{F}_t$ can be evaluated. At this point we can treat the Titan-Saturn case specifically. If one defines the angle α as the angle subtended at Saturn by vectors **OC** and **OP** in figure 1, then

$$\sin\alpha = \frac{R\sin\zeta}{D - R\cos\zeta} \qquad (7)$$

Thus the acceleration A_T acting normal to **OP** is given by

$$A_T = GM_S R\left[\frac{2}{D^3} + \frac{1}{(D - R\cos\zeta)^3}\right]\sin\zeta\cos\zeta \qquad (8)$$

To calculate the pressure variation over the surface at a given altitude, the pressure distribution must balance the tidal force. Since for a given ζ the tidal force is constant, the pressure is constant also. Consequently,

$$\delta p = RA_T\rho\delta\zeta \qquad (9)$$

with p and p+δp being the pressures on either side of a small element of atmosphere and ρ being the density of the element. Before proceeding it is necessary to determine whether the ideal gas law can be utilized to determine ρ.

2. Titan's Lapse rate

Owing to the fact that N_2 is the principle constituent in the atmosphere and that its condensation temperature is at 77.37 K[4], the ideal gas law may not be applicable in describing the thermodynamic properties of Titan's lower atmosphere.

Using data from Lellouch, it is possible to determine the amount of deviation from the ideal gas law up to 600 km. Using rudimentary qbasic code, the following graphs were produced.

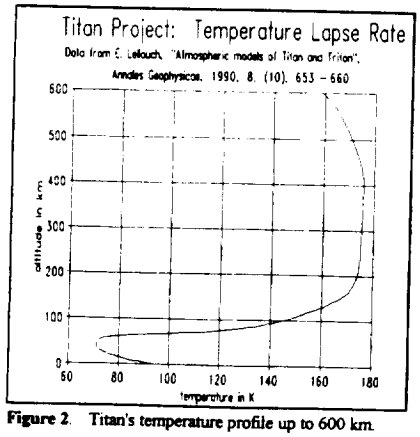


Figure 2. Titan's temperature profile up to 600 km.

Figure 2 illustrates the adiabatic decrease of mean temperature up to the tropopause(≈ 42 km) followed by a steady increase up to the stratopause. According to Hunten at alia[5], the likely heat source driving this escalation of temperature can be attributed to the absorption of solar radiation by brown high altitude haze. Of interest in this project is the fact that Coustenis[6] and Lindal et alia[7] both quantified the uncertainty of the temperature lapse rate at the top of the stratosphere to roughly one sigma (10-15K). Factors contributing to this uncertainty include noise in Voyager 1's radio recordings, deviation from the ideal gas law due to low temperatures - especially in the troposphere, deviations of molecular mass with respect to altitude and other ephemerides.

Figure 3 delineates the amount of sample variance generated by the ideal gas and Van der Waal's profiles assuming that the actual relative pressure profile can be designated as the "true" profile[7]. The Van der Waal's profile provided a slightly better agreement than the ideal gas profile but not significantly so. The Van der Waal's and ideal gas variances tend to converge with the actual profile above the tropopause which correlates with the steady rise in temperature highlighted in figure 2.

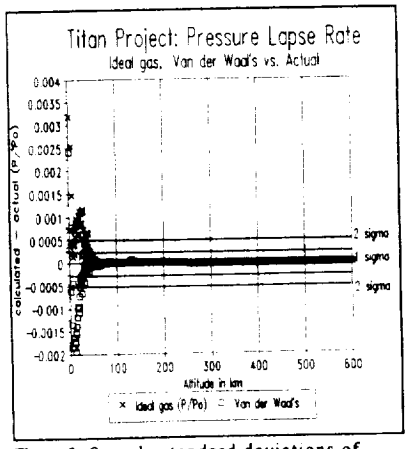


Figure 3. Sample standard deviations of ideal gas and Van der Waal's corrected relative pressures from actual pressure.

For the sake of simplicity, the ideal gas law will be utilized in further computations although the question of whether it is the best scheme is yet to be resolved.

3. Calculation of Equilibrium Tide

Returning to equation (9), one finds that

$$\rho = \frac{Pm}{R_o T} \tag{10}$$

where R_o is the universal gas constant, T is the air temperature (moisture in the form of condensible methane is ignored so that virtual temperature does not come into play), and m is the molecular mass of Titan's air = 29. Combining (8), (9) and (10) one obtains

$$\frac{dP}{d\zeta} = \frac{Pm}{R_o T} GM_T R^2 \left\{ \frac{2}{D^3} + \frac{1}{(D - R\cos\zeta)^3} \right\} \sin\zeta\cos\zeta \tag{11}$$

Rewriting (11) using a binomial expansion and rearranging terms yields

$$\frac{dP}{P} = \frac{3GM_T R^2 m}{R_o T D^3} \left\{ \cos\zeta + (R/D)\cos\zeta + (R/D)^2\cos^2\zeta + \cdots \right\} \sin\zeta\cos\zeta d\zeta. \tag{12}$$

Since the factor (R/D)is lass than 0.01, one can ignore terms after the first. Integrating gives the pressure variation with ζ for a given altitude so that

$$\ln(P_1/P_2) = \frac{3GM_T m(R_T + z)^2}{2D^3 R_o T} \cos^2\zeta \Big|_{\zeta_2}^{\zeta_1} \tag{13}$$

where R_T is Titan's radius, z is an altitude above Titan's surface, P_1, P_2, ζ_1 and ζ_2 are upper and lower bounds of pressure and position angle. ζ is a function of latitude θ and longitude ϕso that $\cos\zeta = \cos\theta\cos\phi$. Using (21), Lorenz [1] devised tidal correction factors which modulate pressure values according to altitude and position.

Table 1 Tidal Correction Factors

Zeta	0	30	45	60	90
z(km)					
0	1.0069	1.0051	1.0032	1.0014	0.9996
5	1.0071	1.0054	1.0035	1.0015	0.9996
10	1.0079	1.0058	1.0037	1.0016	0.9996
20	1.0086	1.0064	1.0041	1.0018	0.9995
30	1.0092	1.0068	1.0043	1.0019	0.9995
40	1.0095	1.007	1.0045	1.002	0.9995
50	1.0095	1.007	1.0045	1.002	0.9995
60	1.0089	1.0065	1.0042	1.0019	0.9995
80	1.0056	1.0041	1.0026	1.0012	0.9997
100	1.0049	1.0036	1.0023	1.001	0.9997
150	1.0045	1.0033	1.0021	1.0009	0.9998
200	1.0044	1.0033	1.0021	1.0009	0.9998
250	1.0045	1.0033	1.0021	1.0009	0.9998
300	1.0045	1.0033	1.0021	1.0009	0.9998
400	1.0048	1.0035	1.0022	1.001	0.9997
600	1.0055	1.0041	1.0026	1.0012	0.9997

It is seen from table 1 that surface pressure varies from 0.9996 to 1.0069 of the nominal which is defined as the normalized pressure for ζ=77° (the mean of the ζ

values for the Voyager 1 measurements on which the models are based)

Table 2. Pressure calculations plus tidal effects at Titan's equator

Altitude (km)	Tidal Factor	Pressure Nominal mbar	Pressure Nom+Tide mbar	% increase in Nominal Pressure
0	1.0069	1474.12380414	1485.27	0.756
5	1.0073	1166.05283448	1173.914	0.674
10	1.0078	902.745587586	910.125	0.817
20	1.0086	50.4757274483	50.9264	0.893
30	1.0092	27.3020292414	27.51533	0.781
40	1.0095	145.896365517	147.3553	1.000
50	1.0095	77.9632448966	78.80382	1.078
60	1.0089	43.0321172414	43.48509	1.053
80	1.0055	18.2604682345	18.37047	0.602
100	1.0049	9.79936777241	9.841425	0.429
150	1.0045	2.56582654345	2.575436	0.375
200	1.0044	0.77720132069	0.782216	0.645
250	1.0045	0.25417446124	0.255208	0.407
300	1.0045	0.091259051	0.09179	0.579
400	1.0048	0.013692728	0.01375	0.392
500	1.0052	0.00231954	0.0023	0.456
600	1.0055	0.00044333	0	0.596

The values of nominal pressure plus tidal fluctuations should be interpreted with caution since we are not taking zonal wind, and other factors into account especially above 200 km. It is readily apparent from the data presented above that the stationary equilibrium tide is of significant magnitude reaching up to 1% of nominal pressure. Essentially two areas of high pressure would be permanently parked at points facing and opposing Saturn along the equator. There may very well be dynamic effects unforeseen by current models attributable to the these permanent areas of high pressure. If anything, areas of high pressure would tend to trap haze in a manner analogues to subsidence inversions.

2. Calculation of Time-Dependent Tide

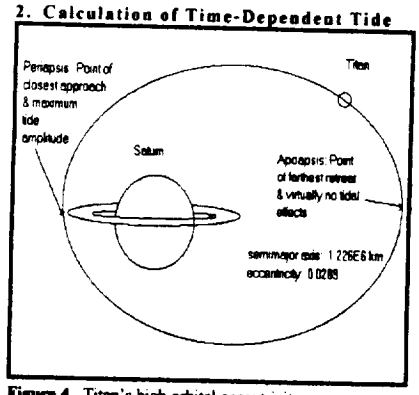


Figure 4. Titan's high orbital eccentricity promotes a time-dependent perturbation of the standing tidal oscillation in its atmosphere.

In order to calculate the time dependent tide, one must introduce classical tidal hydrodynamics. Oscillations induced by external forcings would generally abide by Laplace's tidal equation and the vertical structure equation within a given fluid. They are listed as follows:

Laplace's tidal equation:

$$\frac{d}{d\mu}\left(\frac{1-\mu^2}{f^2-\mu^2}\frac{d\Theta_n}{d\mu}\right) - \frac{1}{f^2-\mu^2}\left(\frac{mf^2+\mu^2}{f f^2-\mu^2} + \frac{m^2}{1-\mu^2}\right)\Theta_n + \frac{4R^2\Omega^2}{gh_n}\Theta_n = 0 \qquad (14)$$

$\mu = \cos\theta$

$f = \sigma/2\Omega$=driving frequency/2*(angular velocity of rotation)

m = integer indicating the periodicity of ϕ referred to as zonal wave number

n = positive integer value

h_n = equivalent depth

Θ_n=Hough function

and Vertical structure equation:

$$H\frac{d^2\chi_n}{dz^2} + (\frac{dH}{dz}-1)\frac{d\chi_n}{dz} + (\kappa + \frac{dH}{dz})\frac{\chi_n}{h_n} = 0 \quad (15)$$

where

$$\chi_n(x) = y_n e^{(x/2)}, \qquad \kappa = \frac{(\gamma-1)}{\gamma},$$

$$H = \frac{kT(z)}{mg_s(\frac{R}{R+z})^2}$$

Laplace's tidal equation possesses trivial solutions in only three cases: for a semidiurnal tide (f=1) with the zonal wave number set equal to zero(m=0), for a nonrotating planet (Ω=0), and for an infinite equivalent depth ($h_n=\infty$). Of interest here is the first case which leaves us with:

$$\frac{d^2\Theta_n}{d\mu^2} + \frac{4R^2\Omega^2}{gh_n}\Theta_n = 0 \quad (16)$$

To obtain our eigenvalue h_n, one assumes that Θ_n is single valued and continuous everywhere and that $\mu = \pm 1$ at the poles so that

$$\Theta^{0}_{n,f=1} = A\sin(\frac{2R\Omega\mu}{\sqrt{gh_n}}) + B\cos(\frac{2R\Omega\mu}{\sqrt{gh_n}}) \quad (17)$$

which contains both symmetric and antisymmetric wave types according to whether A = 0 or B = 0 respectively about the equator. In either case,

$$\frac{4R^2\Omega^2}{\sqrt{gh_n}} = \frac{n\pi}{2} \quad (18)$$

which gives:

$$h_n = \frac{16R^2\Omega^2}{n^2\pi^2 g} \quad (19)$$

The parameter h_n can be thought of as the height of an equivalent, constant density, incompressible, atmosphere. Sagan and Dermott[8] provide the time dependent part of the tide-raising potential at given point (R,θ,ϕ) on the surface of Titan as:

$$\Phi = -\frac{GM}{D}(\frac{R}{D})^2 3\epsilon\{[\frac{1}{2}(3\sin^2\theta\cos^2\phi - 1)\cos\sigma t + \sin^2\theta\sin 2\phi\sin\sigma t] \quad (20)$$

where
R = Titan's radius
D = distance between Saturn and Titan
M = Saturn's mass
θ = Titan's colatitude
ϕ = Titan's longitude
ϵ = eccentricity of Titan's of orbit.

Expanding the tidal potential in terms of the Hough function for n, one obtains:

$$\Phi_n^{\sigma,0} = \begin{cases} -k\frac{12}{n^2\pi^2}(-1)^{n/2} & even\ n \\ 0 & odd\ n \end{cases} \quad (21)$$

where k=radial component of tide. With (21) one can construct the model since one now has the driving term. Returning to equation (15), one can convert it into standard form:

$$\frac{d^2y_n}{dx^2} - \frac{1}{4}[1 - \frac{4}{h_n}(\kappa H + \frac{dH}{dx})]y_n = 0 \quad (22)$$

where

$$x = \int_0^z \frac{dz}{H},$$

$$y_n(x) = e^{x/2}\chi_n.$$

Fluctuations in mean atmospheric pressure, can now be expressed as follows:

$$\delta P_n^{\sigma,m} = \sum_n (\frac{P_0(0)}{H}[-\frac{\Phi_n}{g}e^{-x} + \frac{\gamma h_n}{i\sigma}e^{x/2}(\frac{dy_n}{dx} - \frac{1}{2}y_n)]\Theta_n e^{i(\sigma t + m\phi)}), \quad (23)$$

In order to construct a workable model, one should be able to evaluate (15) in terms of boundary conditions. However, two constraints should be made: (1) x & w_n = 0

at the surface of Titan, and (2) dH/dx=0 at and above some arbitrary height in the atmosphere. Utilizing the first constraint and equation (15) yields the boundary condition at the surface:

$$\frac{dy_n}{dx} + \left(\frac{H}{h_n} - \frac{1}{2}\right)y_n = \frac{i\sigma}{\gamma g h_n}\Phi_n \qquad (24)$$

For the upper boundary condition, constraint (2) must be employed which dictates that our model atmosphere must have an upper level where the scale height is essentially constant. That being the case, one can define the upper boundary condition as:

$$\frac{d^2y_n}{dx^2} - \frac{1}{4}\left[1 - \frac{4\kappa h}{h_n}\right]y_n = 0 \qquad (25)$$

For $h_n < 4\kappa H$ the solution to (25) can be written in the form:

$$y = Ae^{+\lambda x} + Be^{-\lambda x} \qquad (26)$$

where

$$\lambda = \sqrt{\left[\frac{\kappa H}{h_n} - \frac{1}{4}\right]}. \qquad (27)$$

The $Ae^{+\lambda x}$ term of (26) is associated with an upward propagation of energy and the $Be^{-\lambda x}$ with a downward propagation. Since there is no energy source at $x=\infty$ according to equation (26), then B=0. This being the case, one is left with what is referred to as the *radiation condition* mentioned by Chapman and Lindzen[3]. It is worth noting that (26) may not be applicable to all atmospheric conditions especially if λ is particularly small which can lead to the reflection of tidal energy back to the surface due to the opacity of upper atmospheric layers.

Utilizing the Lellouch[4] model for temperature lapse rate, a scale height profile was devised for Titan. Following subsequent computations of air velocity divergence(y(x)), and changes in air velocity divergence with respect to altitude parameter x, an integration refined by Richtmyer[9] was employed in order to solve (23). At this point the Runge Kutta method[10] was employed as a means of self checking to ensure by independent means that the truncation errors in the initial calculations did not lead to computational instabilities. The follwoing plots highlight the various parameters used in obtaining the sum of pressure oscillations for all even modes between 2 and 12.

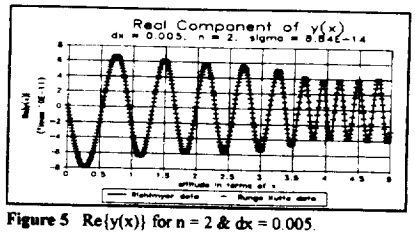


Figure 5 Re{y(x)} for n = 2 & dx = 0.005.

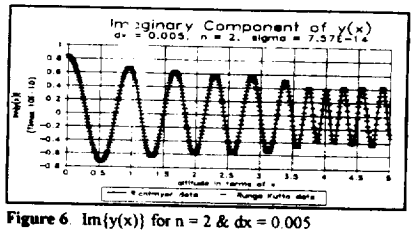


Figure 6. Im{y(x)} for n = 2 & dx = 0.005

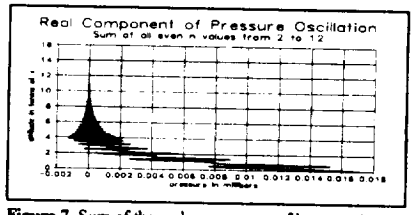


Figure 7. Sum of the real components of barometric fluctuation over a period of 15.95 days.

4. Conclusions

It appears that time-dependent mode of oscillation possesses a magnitude of approximately 0.02 mbar which is 500 times smaller than the magnitude of the corresponding time-independent equilibrium tide making its detection exceedingly difficult. However, even the minute perturbations induced by Titan's eccentric orbit may promoted transfers in angular momentum responsible for the satellites synchronous orbit[11].

The technical implications of standing tidal oscillations existing within the lower atmosphere of Titan may be more important than at first realized. According to Lorenz[1], the upcoming Cassini mission will launch the Huygens atmospheric probe at an entry point correlating to 22°N and 200°E which happens to fall in the locale of the antisaturnian tidal bulge. Since the atmosphere is disstended at the descent point, the probe will experience a lengthened descent time on the order of 30 seconds approximating that all of its descent occurs at the probe's terminal velocity. Therefore, this extended descent time must be taken into account during deployment.

Generally, it would be very difficult to detect these tidal oscillations above the boundary layer. According to Del Genio, Zhou, and Eichler[12], Titan's atmosphere may be dominated by a superrotating zonal winds in much the same manner as Venus's atmosphere is. In essence, much of the incoming solar flux is intercepted by upper tropospheric clouds. Subsequently, a statically stable equilibrium condition is created in the lower and middle troposphere. The radiative heating profile limits the depth of boundary layer convection and detaches the upper level flow from the surface. At Titan's rotation period, the general circulation model employed by these researchers produced equatorial winds in excess of 50/100 m/s with/without stratospheric drag.

5. References

1. R D Lorenz 1991, "Gravitational Tides in the Atmosphere of Titan", Proceedings Symposium on Titan, Toulousse, France, 9 - 12 September 1991.

2. P J Ioannou and R S Lindzen 1993, "Graviational Tides in the Outer Planets. Interior Calculations and Estimations of Tidal Dissipation Factor", The Astrophysical Journal **406**, pp. 266-278.

3. S Chapman and R S Lindzen 1970, Atmospheric Tides, Reidel.

4. E Lellouch 1990, "Atmospheric Models of Titan and Triton", Annales Geophysicae **8**, pp. 653-660.

5. D M Hunten et alia 1984, "Titan" in Saturn ed T Gehrels and M S Matthews.

6. A Coustenis 1991, "Titan: Recent Developments", Vistas in Astronomy **34**, pp. 11-34.

7. G F Lindal, G E Wood, H B Hotz, D N Sweetnam, V R Eshleman, and G L Tyler 1983, "The Atmosphere of Titan: An Analysis of the Voyager 1 Radio Occultation Measurements", Icarus **53**, pp. 348-372.

8. C Sagan and S F Dermott 1982, "The tide in the Seas of Titan", Nature **300**, pp. 731-733.

9. R Richtmyer 1957, Difference Methods for Initial Value Problems, Interscience.

10. G Arfken 1985, Mathematical Methods for Physicists, Academic Press.

11. S J Peale, "Tidal Dissipation, Orbital Evolution, and the Nature of Saturn's Inner Satellites", Icarus **43**, pp. 65-79.

12. A D Del Genio, W Zhou, and T P Eichler 1992, "Equatorial Superrotation in a Slowly Rotating GCM: Implications for Titan and Venus", Icarus **101**, pp. 1-17.

TITANIUM LAYER INFLUENCE ON THE STRENGTH OF A HYBRID TITANIUM COMPOSITE LAMINATE

David R. Veazie
Assistant Professor

Ronald O. Grover, Jr.
Student Assistant

Genine I. Bryant
Student Assistant

High Performance Polymers and Ceramics (HiPPAC) Center
Clark Atlanta University
Atlanta, GA 30314

ABSTRACT: An experimental study was under-taken to investigate the mechanical response of four hybrid titanium composite laminate (HTCL) systems, each prepared using a graphite fiber reinforced thermoplastic polyimide as the adhesive in a unidirectional prepreg. Two of the four HTCL systems were fabricated with the titanium Ti-15-3 alloy, while the other two systems were fabricated with the titanium Timetal β -21S alloy. Each HTCL system consisted of either three plies or four plies of the titanium alloy. Systems with only three plies of titanium had plies measuring 10 mils thick, whereas systems consisting of four plies of titanium had plies measuring 5 mils thick. The improvement in mechanical properties achieved by the influence of the titanium layer was assessed by comparing the uniaxial tensile results of static strength at room temperature. Results included stress-strain curves, ultimate strength, strain-to-failure, initial modulus of the HTCL's, and the description of the observed modes of failure.

KEYWORDS: hybrid technology, thermoplastic polyimide, laminated titanium, adhesive bonding graphite fiber,

Genine I. Bryant is a second year dual degree student at Clark Atlanta University. She is currently pursuing a B.S. in Mathematics, and plans to continue her studies at the North Carolina A&T State University in Mechanical Engineering. She plans to receive B.S. degrees from both institutions in May of 1999. As a HiPPAC Center Scholar, she has participated in research relating to the mechanics of polymer matrix composites. She is from Cleveland, Ohio.

Introduction

The next generation military and commercial aircraft requires applications for high strength, low weight structural components subjected to elevated temperatures. The hybrid titanium composite laminate (HTCL) is a material system that merits attention because of its capability to operate in structures at higher temperatures. In this system, thin plies of titanium are adhesively laminated together using a high temperature resin with high modulus fibers included in the bondline. Material systems such as the ARALL (Arimid Reinforced Aluminum Laminates) and GLARE laminates have essentially the same concept of the HTCL's and are now flying on several commercial and military aircraft. The history of laminated metals has shown definite mechanical advantages that can translate to weight savings in commercial and military aircraft applications [1, 2].

This experimental study will investigate the laminated/hybrid technology applied to high temperature titanium alloys, and a graphite fiber reinforced high temperature thermoplastic polyimide adhesive, in hopes of demonstrating that these systems will be useful in the next generation high speed aircraft. Four unidirectional hybrid titanium composite systems were fabricated with different titanium alloys and titanium layer thicknesses. The results were compared to each other to assess laminate strength, load-deformation behavior, and failure modes.

Materials and Processing

Four HTCL systems were fabricated as 17.8 cm by 17.8 cm panels in a hydraulic press with load cell and 30.5 cm by 30.5 cm induction heated platens. Two of the four HTCL systems were fabricated with the titanium Ti-15-3 alloy, while the other two systems were fabricated with the titanium Timetal β-21S alloy for the metal portion of the laminated panels. Prior to laminating, the titanium alloys were given a surface treatment available from Semco and designated Pasa-Jell 107.

A high temperature polyimide resin obtained from Imitec, designated LARC™-IAX {a version of LARC™-IA 4 mole % offset [8 mole % phthalic anhydride (PA) end cap] in N-methyl-2-pyrrolidone (NMP) in which the backbone was modified slightly with 10 mole % of a more rigid diamine, p-phenylenediamine, to improve solvent resistance}, was used as the adhesive [3]. The fiber, IM7, was an intermediate modulus carbon fiber manufactured by Hercules. Each panel was prepared using LARC™-IAX, 30 wt. % solids solution in NMP with a viscosity of 30,400 cp and 29 vol. % IM7 fibers solution coated as a unidirectional prepreg (adhesive tape) approximately 0.18 mm thick. The prepreg was made by the Composites and Polymers Materials Branch (CPB) of the NASA Langley Research Center on a multipurpose prepregging machine. The primer, which was also prepared by the CPB, was applied to the titanium sheets as a 7.5 wt. % solution of LARC™-IAX in NMP and air dried for one hour after which they were place in a convection oven and heated for one hour each at 150°C and 230°C. In Table 1, the constituent properties of the HTCL's are shown. The yield strengths were determined based on a 0.1% plastic strain.

The laminate assembly was arranged in an open-ended matched-die mold by alternating layers of titanium sheets with layers of adhesive tape. High temperature tape was used to hold the assembly and to prevent the laminate from sticking to the mold. A thermocouple was attached to the end of the laminate to provide feedback to the platen temperature controller. Panel 1 was prepared with three plies of the adhesive tape and four plies of the titanium Ti-15-3 alloy measuring 5 mils thick. Panel 2 was prepared with three plies of adhesive tape and four plies of the titanium Timetal β-21S alloy measuring 5 mils thick. Panel 3 was prepared with two plies of tape and three plies of the titanium Ti-15-3 alloy measuring 11 mils thick. Panel 4 was prepared with two plies of the adhesive tape and three plies of the titanium Timetal β-21S alloy measuring 10 mils thick.

Each laminate assembly was arranged in the bottom part of the die mold, placed in a forced-air oven without the top part of the die mold, then heated for 30 minutes each at 150°C, 175°C and 230°C. This staged process, approximately one hour before applying pressure, was used to remove some of the volatiles prior to the bonding procedure. The laminates were further processed under a pressure of 6.89 MPa at 360°C for one hour.

Dog-bone test specimens similar to those described in ASTM Specification D3552-77 (Standard Test Method for Tensile Properties of Fiber-Reinforced Metal Matrix Composites) measuring 15.2 cm. in length were machined from the laminated panels. A schematic diagram displaying specimen dimensions and typical thickness variations is shown in Figure 1. Three replicates of each specimen were used for each test.

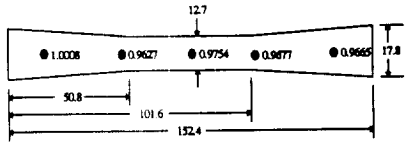


Figure 1. Schematic diagram of HTCL specimen dimensions and typical thickness variations. All dimensions in mm.

Experimental Procedure

Monotonic tensile tests were conducted in accordance with ASTM Specification D3552-77 on the HTCL specimens. From the limited supply of material manufactured, twelve HTCL specimens were fabricated, three from each panel. Each specimen was tested statically to failure at room temperature to examine the tensile strength and load-deformation behavior.

Prior to testing, all specimens were dried for at least 24 hours at 110°C in a convection oven [4]. Following the drying, the specimens were stored inside a desiccator until the start of the tests. The tensile response was evaluated using dog-bone straight-sided specimens in laboratory air on a 100 kN servo-hydraulic test frame equipped with digital controller and data acquisition. A cross-head

displacement rate of 2 mm/min was used to effect a constant strain rate. A plot showing the typical constant strain rate applied in these tests is shown in Figure 2. Hydraulic grips incorporating flat-faced wedges with a non-aggressive surface finish were used to allow for firm gripping of the composite without grip-induced failures. Uniaxial strain was measured on the surface of the specimen using clip-on extensometers with a 25 mm gage section. Stress was calculated based upon the applied load and the specimen cross-section before testing. This stress, along with the corresponding average measurement of two back-to-back extensometers aligned longitudinally on the specimen flats, was used to compute the mechanical properties of the HTCL's.

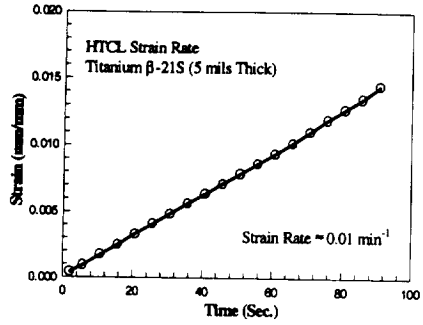


Figure 2. Typical constant strain rate.

Results and Discussion

The results of room temperature tension tests on the HTCL's are shown in Figures 3-5. For comparison purposes, the average response of the 5 mils thick Ti-15-3 and β-21S specimens is plotted in Figure 3, whereas the average response of 11 mils thick Ti-15-3 and the 10 mils thick β-21S specimens is plotted in Figure 4. The yield strengths were determined based on a 0.1% plastic strain. In Figure 5, the average response of all four HTCL's is compared using the equivalent strength-to-weight-ratios of each system.

In a comparison with the monolithic titanium sheets, all HTCL systems showed a significant improvement in the elastic modulus, as well as improved yield strength. The laminates that contained the titanium β-21S alloys approximately doubled the elastic modulus of the monolithic β-21S. Comparable mechanical properties were observed in laminates with equivalent titanium ply thicknesses, however the laminates that contained the titanium Ti-15-3 alloys experienced lower ultimate strength and failed at a lower strain. HTCL systems that contained the four 5 mils thick titanium plies showed a significant strength-to-weight-ratio improvement in elastic modulus and ultimate strength than those that contained three of the thicker titanium ply systems; however, these systems failed at a lower strain.

The fracture modes of the four HTCL's varied greatly. The 10 mils thick β-21S alloy specimens experienced very little delamination damage, mostly localized. All HTCL's, except the 10 mils thick β-21S alloy specimens, suffered severe delamination damage that extended throughout the specimen length. The 5 mils thick titanium plies fractured throughout the gage length of the specimens, while the thicker titanium plies experienced more localized fracture. In the Ti-15-3 alloy specimens, the IM7 fiber breakage was extensive throughout the specimen length; whereas fiber breakage was more localized in the β-21S alloy specimens. Ti-15-3 alloy specimens experienced considerably more fiber matrix debond damage than β-21S alloy specimens.

Conclusions

This investigation evaluated the tensile response of four HTCL systems at room temperature. Results of these test were compared to assess the improvement in mechanical properties achieved by the influence of the titanium layer. Four laminates were fabricated from two different titanium alloys with varying thicknesses, then tested statically to failure in tension. Results included stress-strain curves, ultimate strength, strain-to-failure, initial modulus of the HTCL's, and the description of the observed failure modes.

The experiments showed that the titanium layer in these HTCL systems influences the mechanical properties. Higher yield strength in the titanium alloys, in this case the β-21S alloy, results in HTCL's with greater ultimate strength. However, the stiffer Ti-15-3 titanium alloy did not result in a HTCL with a higher elastic modulus. If systems with higher strength-to-weight-ratios are of primary concern, as in applications for future high speed aircraft, HTCL systems that are fabricated with many thinner titanium plies show improvement over systems with fewer, thicker

titanium plies. Overall, the HTCL systems presented here provide stronger and stiffer alternatives to their corresponding monolithic metals.

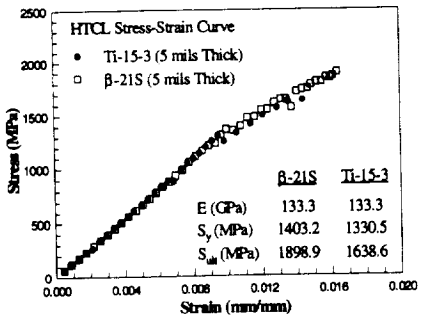


Figure 3. The average stress-strain response of the 5 mils thick Ti-15-3 and β-21S specimens.

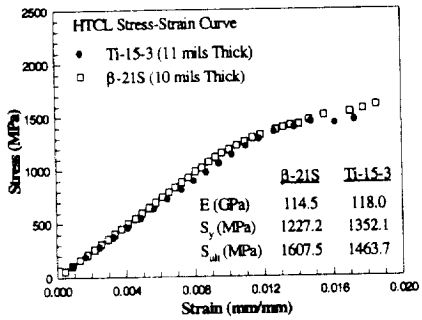


Figure 4. The average stress-strain response of the 11 mils thick Ti-15-3 and 10 mils thick β-21S specimens.

Acknowledgments

The authors gratefully acknowledge Dr. Terry L. St. Clair and Sharon Lowther of the Composites and Polymers Branch of the NASA Langley Research Center for providing the materials used in this study.

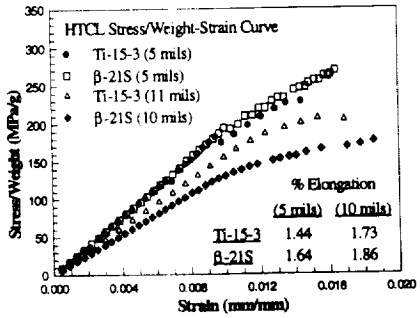


Figure 5. The average stress-strain response of all four HTCL systems using the equivalent strength-to-weight-ratios of each system.

Table 1. Constitutive material properties of the Hybrid Titanium Composite Laminates.

	β-21S[a]	Ti-15-3[a]	LARC™	IM7[b]
S_y MPa	1149.7	1016.4	71.7	N/A
S_{ult} MPa	N/A	N/A	121.9	5310.0
E GPa	64.0	74.3	3.34	275.8
% Elong	N/A	N/A	6.0	1.8

[a] data from Li and Johnson, CERC, Georgia Tech
[b] data from [1]

References

[1] Miller, J. L., Progar, D. J., Johnson, W. S., and St. Clair, T. L., "Preliminary Evaluation of Hybrid Titanium Composite Laminates," NASA TM 109095, April 1994.

[2] Schijve, J., "Development of Fibre-Metal Laminates, ARALL and GLARE, New Fatigue Resistant Materials," *FATIGUE '93*, May 3-7, 1993, pp. 3-20.

[3] Progar, D. J. and St. Clair, T. L., "A New Flexible Backbone Polyimide Adhesive," *Journal of Adhesion Science and Technology*, Vol. 4, No. 7, 1990, pp. 527-549.

[4] Veazie, D. R. and Gates, T. S., "Compressive Creep of IM7/K3B Composite and the Effects of Physical Aging on Viscoelastic Behavior," *SEM Spring Conference on Experimental Mechanics*, June 12-14, 1995, pp. 491-498.

PRELIMINARY EVALUATION OF LOW-DENSITY, LOW-CONDUCTIVITY PAN ABLATIVE COMPOSITES

David R. Veazie
Assistant Professor

Keva Vaughn
Student Assistant

High Performance Polymers and Ceramics (HiPPAC) Center
Clark Atlanta University
Atlanta, GA 30314

ABSTRACT: In this study, the mechanical response of four low-density, low conductivity polyacrylonitrile (PAN) based carbon fiber reinforced phenolic resin composites was investigated, each prepared with ceramic microballoon fillers. Two of the four PAN composite systems were fabricated with 20% microballoon filler content, while the other two systems were fabricated with 15% microballoon content. Composites with 20% microballoon content were cured at pressures of 1.65 MPa and 0.35 MPa, whereas systems with 15% microballoon content were cured at 1.65 MPa and 0.69 MPa. The improvement in mechanical properties achieved by the influences of the microballoon content and the cure pressure was assessed by comparing the uniaxial tensile results of static strength of the composite at room temperature. Results included stress-strain curves, ultimate strength, strain-to-failure, initial modulus of the PAN composite, and the description of the observed modes of failure.

KEYWORDS: phenolic resin, ablative composite, polyacrylonitrile fiber, ceramic microballoons, ply lift.

Keva Vaughn is a third year dual degree student at Clark Atlanta University. She is currently pursuing a B.S. in Mathematics, and plans to continue her studies at the Georgia Institute of Technology in Civil Engineering. She plans to receive B.S. degrees from both institutions in May of 1998. As a HiPPAC Center researcher, she has participated in research relating to the mechanics of polymer matrix composites. She is from New Orleans, Louisiana.

Introduction

Concerns about the stability of the domestic supply of carbonizable aerospace grade rayon for ablative composites have led to interest in an alternate carbon fiber precursor. Polyacrylonitrile (PAN) based carbon fibers have been shown through various studies to be an acceptable alternative. Another approach to developing a replacement for rayon is to reduce the thermal conductivity and density of the composite by the inclusion of filler ceramic microballoons to the PAN prepreg. These microballoons are an evolution from previous rayon low density materials which used carbon microballoons, but provide cost and thermal conductivity advantages.

The ply lift anomaly which occurs in ablative composites causes erratic erosion and spallation, leaving large multiple voids between plies of the material. This can lead to loss of the thermally protective char cap and result in increased heat-affected depth of the material. The fiber/matrix bond strength in low-density PAN based carbon fiber reinforced phenolic composites is believed to influence this event. In particular, higher cross-ply tension and interlaminar shear strength properties are necessary to accommodate the high cross ply thermal expansion (cool-down dilation) experienced in ablative applications.

An experimental study was undertaken to investigate the mechanical properties believed to influence ply lifting in ablative, PAN composite liner materials, in hopes of demonstrating that these systems will be useful as viable substitutes

for rayon-based ablative composites. Four PAN fiber ablative composite systems were fabricated with different ceramic microballoon levels and cure pressures. The results were compared to each other to assess laminate strength, load-deformation behavior, and failure modes.

Materials and Processing

Four PAN ablative composite systems were autoclave cured as 35.56 cm. long by 44.64 cm. wide by 10.16 cm. thick panels. Two of the four PAN composite systems were fabricated with 20% ceramic microballoon filler content, while the other two systems were fabricated with 15% microballoon content. Optical microscopy and vibrational spectroscopy were used to study the size, size distribution and integrity of the virgin microballoons prior to layup.

The prepreg broadgoods, obtained from ICI Fiberite and designated MX-134 LDR, included the ceramic microballoon fillers in SC-1008 Phenolic Resin, along with the T300 PAN fibers. Normal certification testing of the prepreg including a minimum of volatile content, resin content, filler content and flow was performed by ICI Fiberite. In addition, flow and Differential Scanning Calorimetry (DSC) testing of the prepreg was performed by Thiokol. The composite panels were fabricated at Thiokol and the Productivity Enhancement Center of the NASA Marshall Space Flight Center (MSFC).

From the as-received rolls, 35.56 cm. long by 44.64 cm. wide plies were cut and piled into 8 stacks of 32 plies per stack. Each stack was double sealed in foil bags and placed in cold box storage until layup. After warming to ambient (approximately two hours), these 8 "sub-stacks" were preheated in a convection oven at 60°C for 30 min. To simulate tape wrap debulk, the stacks were then arranged in a hydraulic press with load cell and induction heated platens to form a 256 ply composite. Teflon tape and high temperature, non-perforated release film was used to hold the assembly and to prevent the laminate from sticking to the tool surfaces, respectively. Thermocouples attached to the center of each platen provided feedback to the platen temperature controller. A maximum pressure of 2.76 MPa was applied to the stacks at 60°C for 30 min.

Following the press, the panels were bagged and placed under a minimum of 24.0 in. Hg vacuum for 1 hour. Each panel was monitored to assure that the vacuum drop did not exceed 1.0 in. Hg within 15 min. and held to a minimum of 21.0 in. Hg for a period of 15 min. Each bagged panel was then placed in a programmable temperature and pressure cycle autoclave. Autoclave pressure was held constant throughout the cure cycle. Thermocouples placed directly on top and under the center of the first debulk stack was used to monitor the composite temperature. The profile of the 10 segment temperature cycle used for the autoclave cure is shown in Table 1.

Panel 1 was prepared with 20% ceramic microballoon filler content and cured at a pressure of 1.65 MPa. Panel 2 was prepared with 20% microballoon content and cured at 0.35 MPa. Panel 3 was prepared with 15% ceramic microballoon filler content and cured at a pressure of 1.65 MPa. Panel 4 was prepared with 15% microballoon content and cured at 0.69 MPa. The microballoon filler contents selected were based on the widest range that are of interest and are predicted to result in composite bulk densities from 1.1g/cc to 1.4 g/cc. The skewed distribution to the higher process pressures is prompted by the desire to be able to co-process these materials with standard density materials to further reduce cost.

Table 1. Autoclave temperature cure cycle.

Segment	Temp (°C)	Time (Min.)
1. Start	RT	0
2. 1st Ramp	RT → 79	105
3. Apply Press	79	15
4. 1st Hold	79	90
5. 2nd Ramp	79 → 107	60
6. 2nd Hold	107	60
7. 3rd Ramp	107 → 160	90
8. 3rd Hold	160	300
9. Cool Down	160 → 71	330
10. Cool Down	71 → RT	120

Rectangular test specimens similar to those described in ASTM Specification D3039M-95 (Standard Test Method for Tensile Properties of Polymer Matrix Composite Materials) measuring 25.4 cm. in length, 2.54 cm. wide, and approximately 0.3175 cm. thick were machined from the laminated panels. A schematic diagram displaying specimen dimensions and typical thickness variations is shown in Figure 1. Four replicates of each specimen were used for each test.

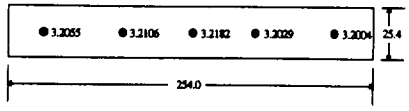


Figure 1. Schematic diagram of PAN specimen dimensions and typical thickness variations. All dimensions in mm.

Experimental Procedure

Monotonic tensile tests were conducted in accordance with ASTM Specification D3039M-95 on the PAN composite specimens. From the limited supply of material manufactured, 16 PAN composite specimens were fabricated, four from each panel. Each specimen was tested statically to failure at room temperature to examine the tensile strength and load-deformation behavior.

The tensile response was evaluated using rectangular specimens in laboratory air on a 100 kN servo-hydraulic test frame equipped with digital controller and data acquisition. A loading rate of 150 N/sec was used to effect a constant strain rate. A plot showing the typical constant strain rate applied in these tests is shown in Figure 2. Hydraulic grips incorporating flat-faced wedges with a non-aggressive surface finish were used to allow for firm gripping of the composite without grip-induced failures. Fine grit abrasive paper, held in place by the pressure of the grip, was used as friction tabs between the wedges and the specimen. Uniaxial strain was measured on the surface of the specimen using clip-on extensometers with a 25 mm gage section. Stress was calculated based upon the applied load and the specimen cross-section before testing. This stress, along with the corresponding average measurement of two back-to-back extensometers aligned longitudinally on the specimen edges, was used to compute the mechanical properties of the PAN composites.

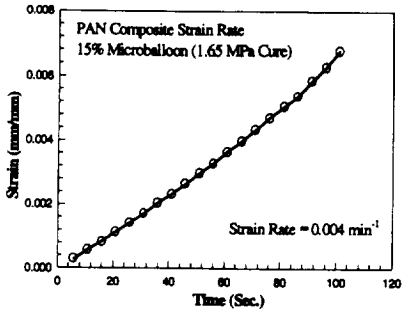


Figure 2. Typical constant strain rate.

Results and Discussion

The results of room temperature tension tests on the PAN composites are shown in Figures 3-5. For comparison purposes, the average response of the 20% ceramic microballoon filler content specimens is plotted in Figure 3, whereas the average response of the 15% microballoon content specimens is plotted in Figure 4. In Figure 5, the average response of all PAN composite specimens is compared.

In a comparison with the specimens from panels that were cured at lower pressures, the specimens from panels cured at 1.65 MPa showed an improvement in the initial modulus. Comparable initial moduli were observed in laminates with equivalent cure pressures, however the laminates that contained 20% ceramic microballoon filler content experienced lower ultimate strength. All PAN specimens showed little plastic deformation and failed at similar strains.

The fracture modes of the four PAN composites varied greatly. The specimens from Panel 1, prepared with 20% ceramic microballoon filler content and cured at a pressure of

1.65 MPa, experienced very little delamination damage, mostly localized. All PAN composites, except the specimens from Panel 1, suffered severe delamination damage that extended throughout the specimen length. The PAN fibers in specimens prepared from panels with 15% ceramic micro-balloon filler content experienced more distortion at fracture. In the specimens from panels that were cured at a pressure of 1.65 MPa, the PAN fiber breakage was more localized; whereas fiber breakage was extensive throughout the specimen length in the specimens that were cured at lower pressures. Specimens from panels cured at lower pressures also experienced considerably more fiber matrix debond damage and crumbling than specimens cured at 1.65 MPa.

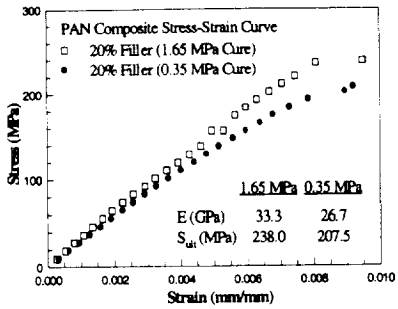


Figure 3. The average stress-strain response of the 20% ceramic microballoon filler content specimens.

Conclusions

This investigation evaluated the tensile response of four PAN composite systems at room temperature. Results of these test were compared to assess the improvement in mechanical properties achieved by the influences of the ceramic microballoon content and the cure pressure. Four laminates were fabricated with two microballoon contents and varying cure pressures, then tested statically to failure in tension. Results included stress-strain curves, ultimate strength, strain-to-failure, initial modulus of the PAN composites, and the description of the observed failure modes.

The experiments showed that the ceramic microballoon content and the cure pressure in these PAN composite systems influences the mechanical properties. If systems with higher moduli are of primary concern, PAN composites that are fabricated with higher cure pressures show improvement over systems cured at lower pressures. If higher tensile strength is necessary to accommodate the high cross ply thermal expansion causing ply lift in ablative applications, PAN composites fabricated with a lower ceramic microballoon filler content yield stronger composites.

Acknowledgments

The authors gratefully acknowledge the Productivity Enhancement Center of the NASA MSFC and the Thiokol Corporation for providing the materials used in this study.

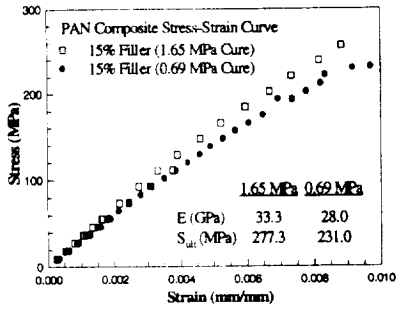


Figure 4. The average stress-strain response of the 15% ceramic microballoon filler content specimens.

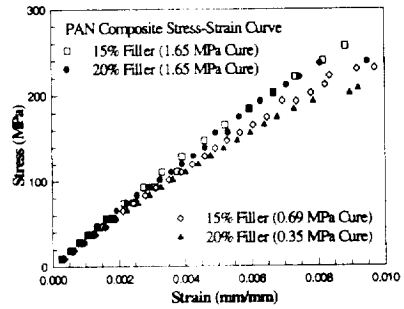


Figure 5. The average stress-strain response of all PAN composite specimens.

SYNTHESIS AND CHARACTERIZATION OF HIGH PERFORMANCE POLYIMIDES CONTAINING THE BICYCLO[2.2.2]OCT-7-ENE RING SYSTEM

M. Alvarado,[1] I. I. Harruna, K. B. Bota, Department of Chemistry and High Performance Polymers and Ceramics Center, Clark Atlanta University, Atlanta, GA 30314 USA

ABSTRACT: Due to the difficulty in processing polyimides with high temperature stability and good solvent resistance, we have synthesized high performance polyimides with bicyclo[2.2.2]-oct-7-ene ring system which can easily be fabricated into films and fibers and subsequently converted to the more stable aromatic polyimides. In order to improve processability, we prepared two polyimides by reacting 1,4-phenylenediamine and 1,3-phenylediamine with bicyclo[2.2.2]-7-octene-2,3,5,6-tetracarboxylic dianhydride. The polyimides were characterized by FTIR, FTNMR, solubility and thermal analysis. Thermogravimetric analysis (TGA) showed that the 1,4-phenylenediamine and 1,3-phenylenediamine containing polyimides were stable up to 460 and 379 °C, respectively under nitrogen atmosphere. No melting transitions were observed for both polyimides. The 1,4-phenylenediamine containing polyimide is partially soluble in dimethyl sulfoxide, methane sulfonic acid and soluble in sulfuric acid at room temperature. The 1,3-phenylenediamine containing polyimide is partially soluble in dimethyl sulfoxide, tetramethyl urea, N,N-dimethyl acetamide and soluble in methane sulfonic acid and sulfuric acid.

Introduction

Due to the increased performance characteristics demanded of polymers in various fields, the use of rodlike aromatic polymers such as polycondensates, polybenzoxazoles,[1] polybenzthiazoles[2] and polyimides[3] are growing steadily. Although, the first synthesis of an aromatic polyimide was reported in 1908,[4] it was not until the late 1950's[5] that high molecular weight polymers were prepared. Interest has developed largely as a result of the increased need for advanced materials in the electronic and aerospace industries. Thermally stable polyimides are used as structural adhesives, coatings, selectively permeable membranes and composite matrices.

Polyimides exhibit excellent physical properties such as, low density, remarkable thermal stability, excellent solvent ressistance and excellent mechanical and electrical performance. The use of polyimides is limited due to their poor solubility in common solvents, the high processing temperature and the evolution of volatiles during imidization. The low solubility of these polymers in organic and inorganic solvents and their high melting transitions are due essentially to the rigidity of the aromatic ring system which result in large van der Waals forces between individuals polymer molecular chains.

Different approaches have been used to improve solubility and reduce melting transitions of polyimides. These efforts include structural modifications such as the introduction of bulky substituents, kinks and crank-shaft, randomly spaced semi-flexible chain fragments, and the use of non-coplanar biphenyl moieties.

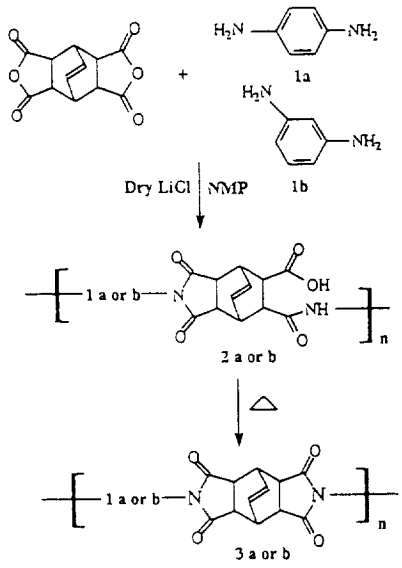
It has been demonstrated that the benzene ring and the bicyclo[2.2.2]octane ring systems are similar in linearity, geometry[6] and polarizability,[7] For example, Harruna et al.,[8] demonstrated that 2,5-disubstituted-1,4-bicyclo[2.2.2]octyl containing polyamides could be converted to the more stable polyamide by pyrolysis in the presence of a Lewis acid. This paper describes the synthesis and characterization of two polyimides containing the bicyclo[2.2.2]-oct-7-ene ring system. Further pyrolysis of the polyimides result in the formation of the more stable desired aromatic polyimide.

Results and Discussion

Polymer Synthesis. The schematic representation of the synthesis of the precursor polyamic acids and polyimides are shown on Scheme I. The reactions were performed under inert and dry

[1]Graduate student at Clark Atlanta University, working toward a master of science degree in organic chemistry (specialization in polymers). Research interests include the synthesis and characterization of high performance polymers and the study of structure-property relationship of substituents on polymer backbones, particularly effects on thermal stability, liquid crystallinity and solubility. Other interests include the synthesis of novel monomers for the preparation of liquid crystalline polymers. Expected date of graduation is July 1996.

conditions in anhydrous NMP. Dried LiCl was added to enhance the polymer solubility. The polyimides (3a, 3b) were precipitated in water and extracted with acetone for 48 hrs. The polyimides were dried under vacuum at 100 °C for 24 hrs. Polyamic acids (2a, 2b) were isolated and treated in the same manner as mentioned above. Both the precursor polyamic acid and precursor polyimides were characterized by elemental analysis, FTIR, differential scanning calorimeter (DSC), viscosity and solubility.



Scheme I
Thermal Imidization of PDA and MDA containing polyamic acids (2a, 2b)

Elemental Analysis. The elemental analyses data: for PDA precursor containing polyamic acid, Anal.: Calcd. for $C_{18}H_{14}O_5N_2$: C, 63.9; H, 4.17; N, 8.28. Found: C, 63.10; H, 4.28; N, 8.02; for PDA containing polyimide, Anal.: Calcd. for $C_{18}H_{12}O_4N_2$: C, 67.5; H, 3.78; N, 8.75. Found: C, 66.02; H, 3.87; N, 8.62. The elemental analyses data for MDA precursor containing polyamic acid, Anal.: Calcd. for $C_{18}H_{14}O_5N_2$: C, 63.9; H, 4.17; N, 8.28. Found: C, 63.13; H, 4.69; N, 8.07; for MDA containing polyimide, Anal.: Calcd. for $C_{18}H_{12}O_4N_2$: C, 67.5; H, 3.78; N, 8.75. Found C, 66.7; H, 3.87; N, 8.69.

Solution properties of the PDA (2b) and MDA (3b) containing polyimides. Table I summarizes the solubility and inherent viscosity properties of precursor PDA and MDA containing polyimides.

Table I. Solution behavior of PDA and MDA containing polyimides (3a, 3b)

		Solubility			
polymer	Aromatic	η_{inh} dL/g	NMP	DMAc	DMF
3a	1a	0.3	-	-	-
3b	1b	0.4	-	+-	-
		DMSO	H2SO4	MSA	TMU
3a	1a	+-	+	+-	-
3b	1b	+-	+	+	+-

Solubilities are given at room temperature: (+) soluble; (-) insoluble; (+-) partially soluble.

Thermal analyses. Thermal properties of the polymers were obtained by TGA and DSC. All polyamic acids and precursor polyimides were found to be more stable under nitrogen than in air. Thermogravimetric analysis indicated that the PDA polyamic acid (2a) intermediate was stable up to 414 °C under nitrogen atmosphere. The PDA containing polyimide was stable up to 460 °C under nitrogen atmosphere. No melting transitions were observed by differential scanning calorimetry analysis. Differential scanning calorimetry analyses showed that polyamic acid (2b) and polyimide (3b) were stable up to 368 and 379 °C under nitrogen atmosphere, respectively (Figure 1).

FTIR Spectroscopy. Infrared spectra were obtained in KBr pellets. The IR spectra of the PDA and MDA containing polyamic acid show characteristic amide and hydroxy absorptions at approximately 3552 to 2729 cm^{-1} for the N-H and O-H stretch, respectively. As the imidization proceeds, these absorptions decrease in intensity and the characteristic imide absorptions (imide carbonyl stretch at 1782 cm^{-1} and 1374 cm^{-1} for the N-C stretch) appear. An absorption at 1709 cm^{-1} for the carboxylic carbonyl indicate that imidization was not completed.

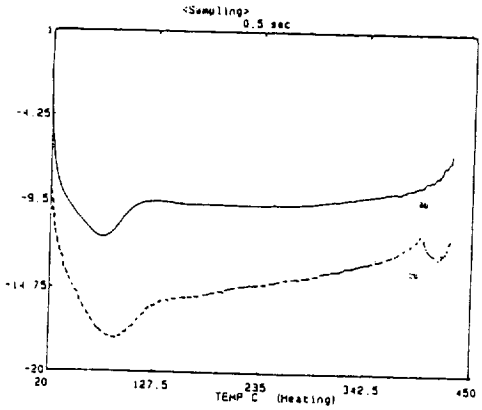


Figure 1. Differential scanning calorimetry thermogram for *m*-phenylenediamine containing polyamic acid and polyimide 2b, 3b.

Experimental Section

Materials. Bicyclo[2.2.2]-7-octene-2,3,5,6-tetracarboxylic dianhydride was obtained from Chiskev Company and was purified by recrystallization from acetonitrile. *m*- and *p*-Phenylenediamine were obtained from Aldrich Chemical Company. *p*-Phenylenediamine was purified by sublimation. *m*-Phenylenediamine was used as received.

Characterization Methods. Infrared spectra were obtained on KBr pellets using the Nicolet Impact 400 infrared spectrometer in the FTIR mode. Thermal analysis of the polymers were performed using the Seiko TG/DTA 320 thermogravimetric analyzer. Open platinum pans were used as reference and sample holders with heating rates of 20 °C/min. Differential scanning calorimetry (DSC) thermograms were obtained on the Seiko DSC 220C differential scanning calorimeter at a heating rate of 10 °C/min. The DSC thermograms were recorded using sealed aluminum sample pans and sealed aluminum reference pans. Inherent viscosity measurements were obtained using a calibrated Cannon-Fenske viscometer at 30 °C with concentrations of 0.5 g/100 mL of precursor polyamic acid and precursor polyimide in concentrated sulfuric acid. Solubilities were determined at room temperature.

Polymerization. In a three-necked round bottom flask (250 mL) equipped with a mechanical stirrer, thermometer, condenser, drying tube and bubbler was added 1.09 g (0.01007 mol) of *p*-phenylenediamine, 1.56 g of dry LiCl and 60 mL of NMP. The mixture was stirred until all the solids dissolved. The solution was cooled to 0 °C. A mixture of anhydrous NMP (25 mL) and bicyclo[2.2.2]-7-octene-2,3,5,6-tetracarboxylic dianhydride (2.5 g, 0.01007 mol) were injected into the stirred solution. The reaction mixture was stirred for 1 h at 0 °C, and then slowly heated to 30 °C overnight. By means of a pasteur pipet half of the reaction solution was removed. The solution collected at 30 °C was precipitated in 200 mL of distilled water and the solids collected by suction filtration. The solids were washed twice with 60 mL of ethanol and dried under vacuum for 1 h at 50 °C to give the crude polyamic acid (2.46 g). The polyamic acid was dried under vacuum at 100 °C for 12 hrs to give the *p*-polyamic acid (2.30 g).

The remaining reaction mixture was stirred and heated at 180 °C overnight. The reaction mixture was precipitated in distilled water. The precipitate was collected by suction filtration. The solids were washed twice with ethanol and dried under vacuum at 50 °C for 1 h. The crude *p*-phenylenediamine containing precursor polyimide (3.11 g) was obtained. The solids were extracted with acetone using a Soxhlet extraction apparatus for two days. The polymer was dried under vacuum at 100 °C for 12 hrs to give *p*-phenylenediamine precursor polyimide (2.95 g).

Conclusion. Bicyclo[2.2.2]-oct-7-ene containing polymers exhibited improved solubility over the all aromatic systems. Ongoing work is focused on studying the effects of various diamines in combination with the bicyclo[2.2.2]-oct-7-ene ring system in polyimides.

Acknowledgment. The authors are grateful to NASA for financial support through the High Performance and Ceramics Center at Clark Atlanta University.

References

(1) Wolfe, J. F; Arnold, F. E., *Macromolecules*, 1981, 14, 909. (2) Roberts, M. F; Jenekhe, S. A., *Polym. Commun*, 1990, 31, 215. (3) Harris, F. W.; Hsu, S. L. C., *High Perform. Polym.*, 1989,1,3. (4) Bogert, T. M.; Renshaw, R. R., *J. Am. Chem. Soc.*, 1908, 30, 1140. (5) E. U. du Pont de Nemours & Co., *French Pat.*, 1960, 1, 239, 491. (6) Dewar, M. J. S; Goldberg, R. S., *J. Am. Chem. Soc.*, 1970,92,1582. (7) Roberts, J. D.; Moreland, W. T., *J. Am. Chem. Soc.*, 1953,25,2167. (8) Harruna, I. I.; Bota, K. B.; McLamore, S. D., *Polymer*, 1993,34,15,3328.

290708 P4

THREE-DIMENSIONAL POLYBENZOBISOXAZOLES AND POLYBENZOBISTHIAZOLES

M. Bray,[1] I. I. Harruna, and K. B. Bota, Department of Chemistry and High Performance Polymers and Ceramics Center, Clark Atlanta University, Atlanta, GA 30314, USA

ABSTRACT: Due to the poor compressive strength properties of high performance liquid crystalline polymers such as polybenzobisoxazoles (PBOs) and polybenzobisthiazoles (PBTs), we have prepared homopolymers and copolymers with PBO and PBT pendant groups on a central star-like unit, 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene, in order to improve upon their compressive strength properties. The fluorene moiety was prepared by the reaction of 2,7-dinitro-9-fluorene with aniline and aniline hydrochloride, followed by reduction with palladium on carbon. The central star-like unit was characterized by FTIR, FTNMR, and elemental analysis. The PBO and PBT pendant groups were synthesised by the polycondensation of 4,6-diaminoresorcinol dihydrochloride with terephthaloyl chloride and 2,5-diamino-1,4-benzendithiol dihydrochloride with terephthaloyl chloride in poly(phosphoric acid), respectively. The resulting linear polymers containing the dicarboxylic end groups were attached to the central star-like unit by refluxing with 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene to give the star-like polymers. The star-like PBO and PBT were soluble in methanesulfonic acid. Further characterization of the polymers is ongoing.

Introduction

Considerable research efforts in recent years have been directed toward the synthesis, characterization, and fabrication of extended-chain or rodlike polymers. The increased activity in the development of high performance polymer fibers possessing high tensile strength and modulus and excellent thermal and environmental resistance has resulted in new ultra high-modulus, ultra high-strength, and thermally stable fibers. Two thermooxidatively stable, aromatic heterocyclic polymers which exhibit liquid crystalline behavior in concentrated solutions are poly[(benzo[1,2-*d*:5,4-*d'*]bisoxazole-2,6-diyl)-1,4-phenylene (PBO) and poly[(benzo[1,2-*d*:4,5-*d'*]bisthiazole-2,6-diyl)-1,4-phenylene (PBT).[1] The films and fibers of these polymers also possess high resistance to solvents.

The observation of ultra high-strength and ultra high-modulus properties for wet spun fibers and biaxially oriented films led to the expectations of utilizing these materials in such applications as transatmospheric vehicles with weight reductions of up to 50%, and in advanced aircraft and aerospace systems.[2]

However, rod-like polymers like PBO and PBT in their highly extended state in fibers and films were discovered to have poor compressive properties. Fibers which are spun from lyotropic solutions of extended-chain polymers exhibit high strength and modulus in the direction of the fiber axis, but the strength and modulus are poor in the direction perpendicular to the fiber axis. The modulus and strength perpendicular to the chain axis are lower by one or two orders of magnitude than that in the chain direction because of weak van der Waals forces.[3]

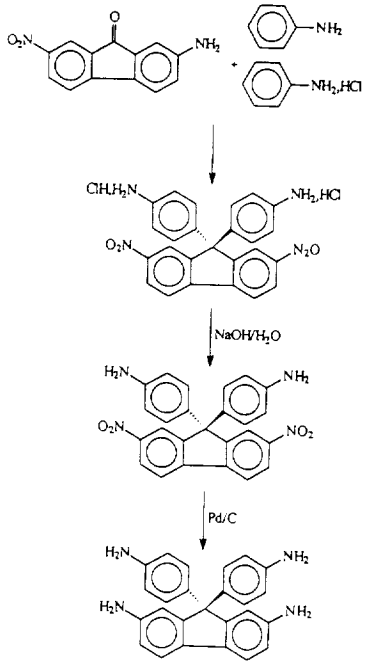
This paper describes an attempt to improve upon the transverse mechanical properties of PBO and PBT films and fibers by utilizing the favorable geometry of an organic nucleus, 2,7-diamino-9,9'-bis(4-aminophenyl)-fluorene. The formation of star-like polymers will lead to radially oriented chains in films and fibers leading to improved transverse mechanical properties.

---

[1]Graduate student at Clark Atlanta University working towards a master of science degree in organic chemistry with specialization in polymers. Research interests include the synthesis and characterization of high performance polymers and the study of the effects of molecular orientation on the compressive strengths of these polymers. Expected date of graduation is July 1996.

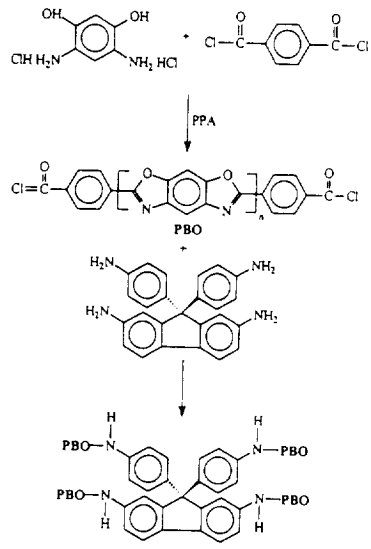
Results and Discussion

Monomer Synthesis. The schematic representation of 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene is shown in Scheme I. The reaction was performed in a nitrogen atmosphere. 2,7-Diamino-9,9'-bis(4-aminophenyl)fluorene was precipitated out of solution followed by the removal of ethanol and hydrazine monohydrate and followed by cooling of the solution. The fluorene unit was characterized by elemental analysis, FTIR, NMR and solubility.



Scheme I
Synthesis of 2,7-Diamino-9,9'-bis(4-aminophenyl)fluorene

Polymer Synthesis. The schematic representations of the polycondensation of PBO and PBT with 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene are shown in Scheme 2 and Scheme 3. The polymers were precipitated in water and extracted with water for 48 h and dried under vacuum at 100 °C for 24 h. The polymers were characterized by FTIR, solubility and elemental analysis. Further characterization is ongoing.



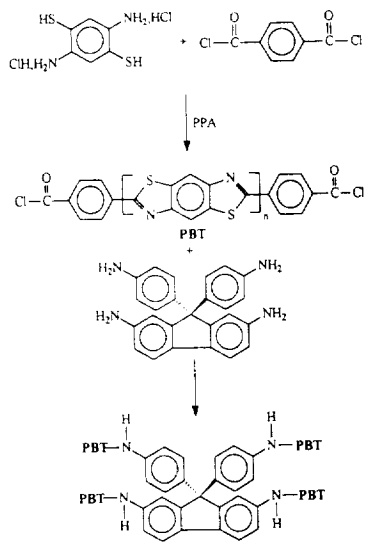
Scheme 2
Synthesis of PBO with 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene

Elemental Analysis. The elemental analyses data: for 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene, Anal.: Calcd. for $C_{25}H_{22}N_4$: C, 79.34; H, 5.86; N, 14.80. Found: C, 79.07; H, 5.96; N,14.92; for PBT containing the fluorene moiety, Anal.: Calcd. for $C_{85}H_{42}N_{12}S_8O_4$: C, 65.78; H, 2.73; N, 10.83; S, 16.53. Found: C, 59.99; H, 3.06; N, 9.14; S, 17.47; for PBO containing the fluorene moiety, Anal.: Calcd. for $C_{85}H_{42}N_{12}O_{12}$: C, 71.73; H, 2.97; N, 11.81. Found: C, 62.91; H, 3.08; N, 9.32.

Solubility Properties. PBO and PBT were partially soluble in methanesulfonic acid and trifluoroacetic acid

Experimental Section

Materials. Commercially available materials used include 2,7-dinitro-9-fluorenone, purchased from Aldrich and 4,6-



Scheme 3
PBT with 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene

diaminoresorcinol dihydrochloride, purchased from TCI. Materials were used as received. Terephthaloyl chloride was recrystallized from hexane.

Characterization Methods. Infrared spectra were obtained on KBr pellets using the Nicolet Impact 400 infrared spectrometer in the FTIR mode. Elemental analysis data was obtained from Atlantic Microlab, Inc., Norcross, GA. Solubilities were determined at room temperature.

Polymerization. In a thoroughly dried three-necked round bottom flask (250 mL) equipped with a mechanical stirrer, thermometer, condenser, nitrogen inlet, drying tube and bubbler was added 2.0 g (0.00829 mol) of 2,5-diamino-1,4-benzenedithiol dihydrochloride and 21.1 g of PPA. The mixture was stirred at room temperature for 24 h under nitrogen. It was then heated for 34 h at 70 °C. After heating to 110 °C, terephthaloyl chloride (1.813 g, 0.0107 mol) was added followed by 21.0 g of PPA. The temperature was kept between 110 °C and 165 °C for 5 h under vacuum, 165 °C for 12 h under nitrogen, 180 °C for 12 h under vacuum, and 195 °C for 12 h under nitrogen. To this mixture, 0.768 g (0.00203 mol) of 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene and 20.0 g of PPA were added and heated at 195 °C for 72 h. The reaction mixture was precipitated in water and then extracted with water for 48 h. The PBT containing the 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene was then dried at 100 °C under vacuum for 24 h to yield 3.3 g.

To prepare PBO with the 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene star-like unit, 2.0 g (0.0094 mol) of 4,6-diaminoresorcinol dihydrochloride and 21.1 g PPA were allowed to stir at room temperature for 24 h under nitrogen in a thoroughly dried three-necked round bottom flask (250 mL) equipped with a mechanical stirrer, thermometer, condenser, nitrogen inlet, drying tube and bubbler. It was then heated for 34 h at 70 °C. After heating to 110 °C, terephthaloyl chloride (2.10 g, 0.0124 mol) was added followed by 21.0 g of PPA. The temperature was kept between 110 °C and 165 °C for 5 h under vacuum, 165 °C for 12 h under nitrogen, 180 °C for 12 h under vacuum, and 195 °C for 12 h under nitrogen. To this mixture, 0.889 g (0.0023 mol) of 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene and 20.0 g of PPA were added and heated at 195 °C for 72 h. The reaction mixture was precipitated in water and then extracted with water for 48 h. The PBO containing the 2,7-diamino-9,9'-bis(4-aminophenyl)fluorene was then dried at 100 °C under vacuum to yield 2.9 g.

Ongoing Studies. The thermal analyses of the polymers will be performed using the Seiko TG/DTA 320 thermogravimetric analyzer and the Seiko DSC 220C differential scanning calorimeter. Inherent viscosity measurements will be obtained using a calibrated Cannon-Fenske viscometer.

Polycondensation of 2,5-diamino-1,4-benzenedithiol dihydrochloride and 4,6-diaminoresorcinol dihydrochloride with excess terephthaloyl chloride in PPA will result in the formation of the random copolymer of PBO and PBT. Further reaction with the fluorene moiety will give the star-like random copolymer of PBO and PBT.

Acknowledgment. The authors are grateful to NASA for their financial support through the High Performance and Ceramics Center at Clark Atlanta University.

References

(1) Evers, R. C.; Arnold, F. E.; Helminiak, T. E., *Macromolecules*, 1981, 14, 925.
(2) Evers, R. C.; Moore, G. J., *Polymer*, 1970, 25, 1863
(3) Evers, R. C.; Arnold, F. E.; Helminiak, T. E., *Macromolecules*, 1981, 14, 925.

290712 p6

Impact Resistance of the Sandwich Shells used in Aerospace Applications

Mr. Derke R. Hughes, Graduate student

William J. Craft, Professor of Mechanical Engineering

Ajit D. Kelkar, Associate Professor of Mechanical Engineering

Center for Aerospace Research
North Carolina A&T State University
Greensboro, NC 27411

Abstract

The purpose of the investigating low velocity impact on sandwich shell composites is to determine the amount of damage at different energies and the failure mode versus the energy of the impact. There is also a desire to develop an analytical model emulating damage, and to compare the projected damage with the actual damage determined by the experiments.

Sandwich shells of this study form a layered flat plate of dimension 0.1778 Meter by 0.0254 Meter. The thickness of the plate is a variable depending on the thickness of the individual layers. The sandwich plate is composed of two external sheets of stiff, dense, strong material, and one internal layer called the core. The external plates of this study are constructed of graphite epoxy laminate. The face sheets of this study are composed of a laminated plane weave graphite reinforced prepreg. For this study the sheets are 9 layers in thickness. The internal plates are fabricated from one or more layers of an isotropic plastic foam called Rohacell. The purpose of the sandwich shell construction is to reduce the amount of overall weight of a structure while maintaining high stiffness and strength, particularly due to bending loads. Both stiffness and strength are exceedingly important qualities in an aerospace structural application.

Low velocity impact is an important test on any composite including a sandwich to quantify the actual integrity of the structure after impact degradation. Composites are generally quite susceptible to such damage and show a dramatic loss in stiffness and strength after impact.

In addition, low energy impact damage is difficult to detect in the traditional laminated composites, and there are no test results in the open literature on new sandwich shell composites. Not only is the reaction of sandwich shells to low velocity impact unknown, our ability to detect threshold damage is also unknown. There has been substantial work on impact damage of composites in the department, and, to date, we have a total of eight full-size panels awaiting Wright Patterson Flight Dynamics Laboratory Dynatub drop weight facility.

While the number of samples is low, specimens are difficult to manufacture, and the test program is expected to answer several fundamental questions: **What is the threshold energy which produces incipient damage?** and **What is the maximum damage a panel can sustain which produces backface spalling?**

In addition, the comparison between sandwich plates and laminated composite plates will shed new light on the practicability of composites in aerospace structural design.

Biography: Mr. Derke Hughes is a Master of Science student in the Department of Mechanical Engineering at North Carolina A&T State University. His thesis work is in composite materials within the Mechanics and Materials Group. His undergraduate degree is in mechanical engineering from the University of Texas at Austin. He has an interest in testing and analyzing composite materials with applicability to high performance aircraft. Dr. Craft is the faculty advisor. Dr. Kelkar serves on the committee of Mr. Hughes. Both faculty have an interest in sandwich composites and low velocity damage analysis.

Introduction

Theory of sandwich shell structures was developed after World War II in order to reduce the weight of a structure while maintaining the structure*s mechanical properties. In the 1950*s and 1960*s, the effort to create lighter structures intensified. This spawned the development of Sandwich Shell Technology so that the military aircraft, especially, would be more fuel efficient due to its lightweight design.

Currently, composites materials are increasingly used in the defense and space industry as a further weight reduction as compared to metals used in sandwich construction. With an increase desire for fuel efficiency, advances in composites laminates will be accelerated by further investigations into their mechanical properties. Some properties shared by many fiber-reinforced composite laminates include a low ultimate strains and very low strength in thickness [1]. Sandwich shells composites have low shear strength and the internal sheets may also fail in compression before the failure of the face sheets. Meanwhile, sandwich shell structures exhibits excellent stiffness in bending, the primary loading source in many applications. This benefit makes sandwich beam and plate structures of interest for a number of applications.

While this is true, a low velocity impact resistance restricts the use of such structures where public safety is an issue and in other cases where impact is a likely design consideration. Accordingly, low velocity testing can demonstrate the intrinsic limitations of composite laminates including sandwich shells. Low velocity impacts are classified as impacts with velocities ranging under 12.2 m/sec. Another point of interest is to determine the relative damage tolerance of a sandwich shell to that an equivalent weight laminate.

The objective of the report is to investigate the low velocity impact damage at different energies and failure mode versus the energy of the impact. Also, an analytical model predicting the actual damage will compare experimental to projected results.

Sandwich Shell Construction

The simplest symmetric sandwich shell structure is composed of two external sheets with an internal core or sheet. The external sheets are composite laminates, and the internal sheet is constructed of a less dense material. The external sheets in the study are several laminae made from single plain weave carbon graphite fibers with an epoxy resin (Diaminodiphenly Sulfone). The laminates are made of nine laminae which have produced a final sheet of from 0.00165 to 0.00173 Meters. The stacking sequence is [0/90], a common one, which parallels that of the Cristeu [2] study in order to develop an effective matter to utilize and cure prepreg laminae. The study performed by Gause and Huang [3] yields results similar to Cristeu [2] study. In Gause and Huang study, the stacking sequence of [+\-45/0/+\-] had comparatively similar results with regards to low velocity impact effects. So, the stacking sequence of [0/90] was employed on the test specimens* external sheets.

Of importance is the selection of an adequate internal material or core material. For this study, we selected a new foam sheet called Rohacell. Rohacell is a polymethacrylide manufactured with various densities and thickness. This study selected the density and thickness which could be most easily constructed in our laboratory environment. In practice, the selection of the core would dictated as part of complex design process in which stiffness, weight, manufacturability, or cost might be parameters to be optimized under a particular loading or series of loading conditions. Other factors such as temperature and environment could also represent important design considerations.

Rohacell foam can be acquired with a variety of densities and shear and compressive strengths. One important consideration in using it for this work is that so many thicknesses, densities, and strengths exist. This makes it possible for us to experiment with combinations of sandwiches to see how far these structures can be driven toward an optimized state. Tables 1 and 2 indicate some of the important stiffness and strength parameters for the core material, Rohacell, a new Polymethacrylimide (PMI) rigid cellular

Keywords: Composite Materials, Sandwich Shells, Low Velocity Impact

foam.

Density (Kg/M^3)	32	52	75
Youngs Modulus (N/mm^2)	36	70	92
Shear Modulus (N/mm^2)	13	19	29

Table 1. Rohacell Stiffness Properties

Density Kg/M^3	32	52	75
Tensile Strength N/mm^2	1.0	1.9	2.8
Compressive Strength N/mm^2	0.4	0.9	1.5
Shear Strength N/mm^2	0.4	0.8	1.3
Flexural Strength N/mm^2	0.8	1.6	2.5

Table 2. Rohacell Strength Properties

In this study, Rohacell is bonded to the external sheet by an adhesive film known as AF-163. The adhesive film has excellent operating temperature range and bonding strength.

The procedure for manufacturing a sandwich shell with the previously mentioned material requires several steps. The stacking sequence of [0/90] determines how the prepreg will be cut and laid up. The outer and inner face sheets are laminated in a hydraulic press. The Hydraulic Press is initially heated to 121° C so the time required for the press to reach a pressure and temperature 689.5 kPa and 177 ° C throughout all layers. The coefficient of expansion for the resin and fibers are substantially different, and they can cause parasitic internal stresses unless processing is done carefully and consistently. After the laminate has been maintained at the prescribed temperature and pressure, the laminate must be cooled at 2 degrees per minute. This rate allows enough time for an annealing action reducing the parasitic stresses due to thermal expansion mismatch of the resins and fibers. Once the face sheet is formed, the internal sheet is bonded to each of the external sheet in an oven. The AF-163 adhesive also has a different thermal expansion. Again, a prescribed heating and cooling rate as part of the curing process help to control parasitic stresses. In this process, the oven is set to 149 °C, and remains at that temperature for 45 minutes. The constructed sandwich shell cools initial in the oven for ten minutes. Then, ambient air cools the specimen to roughly 38 °C. Once the sandwich shell laminated composites structures are assembled, they are checked for warping and dimension. They are edge cut to their final shape required for low velocity impact tests, measured for dimensional tolerances, and weighed. The dimensions are shown in **Figure 1a.** (plane view) and **Figure 1b.** (edge view) on the last page.

Discussion of Tests for Mechanical Properties and Impact Test

The gross failure of a sandwich shell will depend on delamination. For total failure, delamination will occur in at least one of the tested specimens. In impact, the rounded indentor is expected to penetrate one face sheet, most of the core, and produce back-face spalling in the opposite face sheet. This delamination of the back face indicates total failure.

The determination of incipient damage is much more subtle to determine, and it depends a great deal on the quality of the instrumentation and test equipment. The test for minimal damage concerns primarily the front face sheet. C-Scan analysis of the sandwich is the best way to determine incipient damage.

Once the incipient and maximum damage values have been determined, additional impact tests with impact energies set to ranges between these values will indicate the progression of the failure from the sandwich shells. The Dynatup machine and computer software packages help by calculating the

Keywords: Composite Materials, Sandwich Shells, Low Velocity Impact

energy delivered to the sandwich by the drop weight system. The Dynatup machine performs in accordance with the American Standard for Test and Materials (ASTM) test, D-2430.

Discussion on Related Research

In the Caprino and Teti [4] study, the stress intensity factor, defined below, was used as an analytical approach for developing some correlation between untested and tested structures:

$$F_c / F_o = (C_o / C)$$

The residual stress is indicated by F_o in the untested structure. The variable, F_c is the corresponding stress due to the estimated stress caused by the tested structure.

A second relationship defines the size of an impact-produced damage region, the equivalent hole, C_{eq}, to the impact energy and to the equivalent length where A and m represent a power law

$$C_{eq} = A\ U^m$$

The constants are dependent on the material and the shape of the impact. Also, the energy reaches a state at which the strength reduction does not change properties U_o. So, the equivalent equation becomes the following:

$$C_o = A\ U_o^n$$

This substituted into the previous equation producing the correlation between the initial strength of the materials before impact. In this research, impact energy will be related to the specimen damage and reduction in residual strength.

Analysis of Impact Damage

The sandwich panels already produced are not optimized to resist impact damage. In fact, sandwich panels are generally already optimized for stiffness or strength. Because impact is so complex, there are no standards nor analytical methods for the design of sandwich panels to resist impact. In order to consider such tools, we must begin with a very thorough understanding of the material and laminated properties of sandwich panels as well as a knowledge of what happens during impact damage.

To begin with, the material properties must be determined by performing tension and compression tests on the external sheets and core sheet. In addition, failure loads must also be determined for the face and core sheets in bending, tension, compression, and in shear.

Finally, the applicability of existing failure criteria should be reviewed to determine the onset of damage.

What is required next is substantially more complex. What is required now is an understanding of post-incipient damage and the mechanisms which are brought into play. As an example, the next possible failure mode could be the wrinkling phenomenon. This theory investigated in the Kuhhorn and Schoop [5] study provides some insight into the wrinkling phenomenon implicit in sandwich shell structural failure. This local failure mode of sandwich shells has been addressed via some simplifying assumptions making the analysis tractable. The assumptions for geometrically nonlinear sandwich theory include:

- Sandwich plates must be three - layered and symmetrical with respect to the middle surfaces.
- The core material is much more ductile compared with the external plates. Both are modeled as separate homogeneous materials.
- The thickness of core and skins should be constant.
- The plane membrane stresses of the core are small and may be neglected in comparison to those of the skins. Therefore, the core transmits only transverse shear and normal stresses.

The external plate behavior is characterized by Kirchoff/ Love Theory which assumes plate flexibility in stretching and bending but not in shear.

Keywords: Composite Materials, Sandwich Shells, Low Velocity Impact

These assumptions for the independent bending of the skins and a thickness stretching of the core which is linearly distributed allow the formation of the geometrically nonlinear sandwich theory equations.

Conclusions

This paper is both a statement of how we are extending the state of the art by developing both an experimental and analytical capability in sandwich shell composites. The fabrication of panels has been described as have been the test procedures. A description of how the composite panels are to be tested has also been included. Unfortunately, we have no results yet as testing will begin in April.

Part of our accomplishments to date include our development of a fabrication of sandwich panel components and the creation of panels which meet the requirements of generally accepted low velocity impact testing research.

This research will assist us and others in developing better and more durable aerospace transportation vehicles.

In closing, I wish to thank the NASA Center for Research Excellence for funding my assistantship. I also wish to express my sincere thanks to Dr. V. S. Avva and Professor Sadler of the Center for Composite Materials Research for the use of their facilities as well as Dr. Craft and Dr. Kelkar for their assistance in developing this topic and manuscript.

References

1. Kelkar, A., "Analysis of Composite Laminates Subjected to Low Velocity Impact Loading," Wright-Patterson Air Force Base, AFB 45433-6553, June 1992, pp. 4-5.

2. Cristeu, N., Malvern, L. E., and Sierakowski, R. L., "Failure Mechanisms in Composites Plates Impacted by Blunt-Ended Penetrators," Foreign Object Impact Damage to Composites, ASTM STP 568, *American Society for Test and Materials*, 1975, pp. 159-172.

3. Gause, L. W., and Huang, S. L., "Compression Fatigue of Impact Damaged Graphite Epoxy Sandwich Beams," Naval Air Development Center, NADC-77305-60, April 1978.

4. Caprino, G., and Teti, R., "Impact and post-impact behavior of foam core sandwich structures," *Composite Structures*, 1994, pp. 47-55.

5. Kuhhorn, A., and Schoop, H., "A nonlinear theory for sandwich shells including the wrinkling phenomenon," *Archives of Applied Mechanics*, 1992, pp. 413-427.

Keywords: Composite Materials, Sandwich Shells, Low Velocity Impact

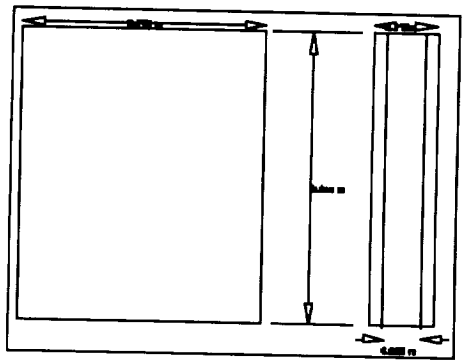


Figure 1a. Plan View of the Sandwich Specimen **Figure 1b.** Edge View of the Test Specimen

Keywords: Composite Materials, Sandwich Shells, Low Velocity Impact

Heat Transfer and Skin Friction Comparison of Dimpled Versus Protrusion Roughness

Mark E. Kithcart and David E. Klett
Mechanical Engineering Department
North Carolina A&T State University
Greensboro, NC 27411

ABSTRACT

This paper presents results of skin friction and heat transfer measurements made on flat plates with closely-spaced, three-dimensional surface roughness elements (hemispherical dimples, hemispherical protrusions, and rectangular protrusions) in turbulent boundary-layer flow at velocities ranging from 18 to 40 m/s (roughness Reynold's nos. from 1500 to 4000). The roughness element densities for each plate were chosen using Simpson's sand-grain roughness correlation to study the effects of spacing on drag and heat transfer in the vicinity of the peak of the Simpson-Dvorak correlation curve. The data is presented in terms of the efficiency factor, $\eta = (St_r/St_s)/(C_{fr}/C_{fs})$, to provide a means of comparing the effects on skin friction and heat transfer collectively. Dimpled surfaces have the advantage of significantly increasing heat transfer at a lower penalty in increased drag compared to the protrusion roughness studied.

KEY WORDS:Surface Roughness, Dimples,Efficiency, Turbulent Boundary Layer, Forced Convection, Drag

INTRODUCTION

The effects of surface roughness on skin friction and heat transfer in wall-bounded flows has long been an area of major interest in fluid mechanics research. Applications of the acquired knowledge to this point range from the study of atmospheric boundary layers to the enhancement of heat transfer. Nikuradse [1950] and Schlichting [1937] conducted the earliest in-depth studies on rough-wall flow behavior. In an effort to develop a useful surface roughness correlation which took into account his data on uniform, discrete roughness along with Nikuradse's results for sand-grain roughness, Schlichting devised the equivalent sand-grain roughness parameter, k_s. This concept has been widely used as an analytical tool in understanding and predicting both rough-wall friction and heat transfer; see for example, Dipprey and Sabersky [1963], Dvorak [1969], Webb et al. [1971], and Dirling [1973]. An extensive summary and bibliography of recent rough-wall heat transfer research can be found in Webb [1994]. Presently, with the availability of high-speed computers, the necessity of relying on the sand-grain correlation for prediction purposes has been mitigated, especially in cases where uniform roughness elements are found (see Olsen and Smith [1985], for example); More detailed modeling of the flow physics is now possible.

For situations where the boundary-layer approximation of the Navier-Stokes equations is used for modeling and prediction of non-sand-grain rough-wall flow, input is needed in the form of drag and heat transfer correlations for the individual roughness elements to effect closure (instead of the k_s correlation method). This has become known as the discrete element method, as described by Finson et al.[1980], Taylor et al. [1984], [1985], [1992], and Hosni et al. [1991]. One of the goals of the experimental program detailed herein was to provide a database of skin friction and heat transfer measurements from rough plates with uniform, closely-spaced three-dimensional roughness elements (hemispherical dimples, hemispherical protrusions, and rectangular protrusions), which could be used for correlation and validation of rough-wall boundary-layer codes. In the course of analyzing the data obtained in the experiment described below, insight into the nature of the friction/heat transfer relationship of the rough surfaces tested was obtained using the efficiency factor described by Webb [1994], $\eta = (St_r/St_s)/(C_{fr}/C_{fs})$, where the subscripts r and s refer to rough and smooth wall conditions, respectively. The results of this analysis form the topic of this paper.

EXPERIMENT

The experiments were performed in the open-circuit subsonic wind tunnel shown schematically in Figure 1. The test section is 3.05 m long x 0.51 m high x 0.71 m wide. The tunnel is powered by a 37 KW variable speed drive, producing a maximum test section velocity of 40 m/s. An adjustable ceiling in the test section is used to create a zero pressure gradient in the vicinity of the test plates. A boundary-layer plate is mounted 5 cm above the tunnel floor. A suction system is incorporated at the leading edge to remove the upstream (contraction section) boundary-layer, allowing the development of a new turbulent boundary layer over the plate. A 1.6 mm wire was used to trip the boundary layer 5 cm downstream of the suction slot.

An air bearing drag balance, designed at NASA Langley, was used for skin friction measurements in conjunction with a PC-based data acquisition system. As shown in Figure 1, the drag balance was supported on a concrete base below the tunnel test section. The 303-m wide x 608-mm long x 12.7-mm thick aluminum test plates were mounted on the drag balance and placed in the test section through rectangular openings cut in the test section floor and the boundary-layer plate. The smooth surface of the test plates was typically flush with the boundary-layer plate with the protrusion roughness elements extending into the boundary layer. In addition, the rectangular roughness element plates were also tested with the tops of the roughness elements flush with the boundary-layer plate. The respective plate surfaces were made flush with the boundary-layer plate to within ±0.05 mm using a metal straight edge. Gaps between the periphery of the test plates and the boundary-layer plate were less than 1.5 mm on average, allowing free movement of the drag balance during friction tests without disrupting the boundary layer at the test plate leading edge. Airflow within the gaps was eliminated by sealing the enclosure containing the balance (Figure 1), thus equalizing the pressure between the enclosure and the tunnel test section.

For heat transfer measurements, electric heating pads (powered by a variable ac supply) were sandwiched between the test plate and the bottom insulation, providing uniform heating of the lower plate surface. All test plates were instrumented with 12 copper-constantan thermocouples, spaced evenly over the test plate area and installed in holes drilled in the bottom of the plates to within 0.4 mm of the top plate surface. The thermocouple measurements were supplemented with infrared measurements of the test plate surface. The gaps between the test plates and the boundary-layer plate described above were filled with insulation during heat transfer experiments. For more details on the experimental setup, see Klett and Kithcart [1992].

The roughness densities for the test plates were determined using the Simpson [1973] roughness correlation, with the intent of quantifying the effects of spacing on skin friction and heat transfer, especially in the vicinity of the peak roughness effect which occurs at a value of 4.68 for the correlation parameter, A_s/A_p, where A_s is the plate area and A_p is the total projected frontal area of the protrusion roughness. Plates with hemispherical and rectangular protrusion roughness were designed to have A_s/A_p values ranging between 2.7 to 11.

For hemispherical dimple roughness A_s/A_p is not an appropriate roughness-density correlation since there is no projected frontal area, therefore a different correlation parameter was needed for the dimpled roughness data. An error in machining resulted in non-uniformity of depth of the dimples from plate to plate, adding another variable in addition to spacing, which affected both drag and heat transfer. The depth variation, and its effect, was discovered when two test plates with different spacings (plates D4 and D5 in Table 1) produced the unexpected result of nearly identical drag and heat transfer measurements. After several confirming retests with consistently reproducible results, a variation in dimple depth was determined to be the cause. A correlation parameter was sought that would account for variations in depth and diameter of dimples (diameter is a function of depth) as well as spacing density. A parameter that suitably correlates both the heat transfer and skin friction data was determined to be $A_{sp}/A_{fp}*d/D$, where A_{sp} is the smooth plate area not covered by dimples, A_{fp} is the total footprint area of the dimples ($A_{sp}+A_{fp} = A_s$), and d and D are the average dimple depth and diameter, respectively. The choice of this parameter was influenced by the results of Wieghardt [1943], who studied the effects of depth and diameter on the drag coefficient of a circular cavity (also see Schlichting [1979]).

When experiments on the dimple roughness had been completed, aluminum spheres were placed in the depressions and covered with a thin coat of heat-

resistant spray paint to hold them in place. The rectangular roughness patterns were created using a cross hatch milling process which resulted in elements of uniform size, with each plate having a different element spacing. The rough plate design specifications are shown schematically in Figures 2 and 3, and the roughness density parameters are given in Tables 1-3.

The average plate surface temperature was calculated from the readings of the 12 plate thermocouples. Radiative heat transfer from the top surface was calculated based on the average surface temperature and a surface emissivity of 0.95. Edge and bottom conduction losses were calculated from the readings of thermocouples placed around the periphery of the plate and on either side of the insulation. The convective heat transfer could then be determined from the power input to the heaters minus the conductive and radiative losses. Radiative losses amounted to approximately 2 percent of the total power input while conduction losses accounted for about 10 percent. The Stanton number for each plate was calculated from the convective heat transfer using

$$St = \frac{Q}{\rho C_p A_t U_\infty (T_w - T_\infty)} \qquad (1)$$

where Q is the convective heat transfer, T_w is the average surface temperature, T_∞ is the freestream temperature, ρ is the air density at the film temperature (average of T_w and T_∞), A_t is the plate surface area, C_p is the air specific heat and U_∞ is the freestream velocity.

Likewise, the measured friction force obtained from the drag balance was converted to the friction coefficient with

$$C_f = \frac{F}{\frac{1}{2}\rho A_t U_\infty^2} \qquad (2)$$

where F is the measured drag force and the other parameters are as described above.

An uncertainty analysis was carried out for St and C_f based on the estimated uncertainties in the parameters used to calculate the coefficients, using the methods described by Coleman and Steele [1989]. The error bars shown in Figs. 4, 5, 6, and 7 represent an uncertainty in the St and C_f values of no more than 3.5% at 95% confidence level. A smooth plate with a polished surface was used as the baseline for determining heat transfer and drag augmentation. The measured drag force on the smooth plate was compared with the value predicted by the rough-surface boundary-layer code of Taylor, Coleman, and Hodge [1984] with the roughness turned off, and the values agreed within four percent over the entire range of freestream velocities tested.

RESULTS AND DISCUSSION

Figures 4, 5, 6 and 7 present the measured skin friction and heat transfer data for each of the three classes of surfaces measured; Figure 4 for the dimpled surfaces, Figure 5 for the hemispherical protrusions, and Figures 6 and 7 for the rectangular protrusions. The values of C_f and St shown plotted in the Figures are the average values for each plate over the range of freestream velocities used in the experiments, 18 - 40 m/s, which resulted in roughness Reynold's numbers in the range $1500<e^+<4000$, i. e. fully rough flow, where $e^+ = e\,u^*/\nu$ where e is the roughness height (defined for the protrusion roughness only, u^* is the friction velocity, and ν is the kinematic viscosity. The dimpled plates display monotonically decreasing drag and heat transfer as the dimples become more widely spaced, i.e. for increasing values of $A_{sp}/A_{fp}*d/D$. At the largest spacing tested, C_f and St approach each other in value and, had plates with yet wider spacing been tested, the trend exhibited would lead one to speculate that St would eventually exceed C_f. These results give rise to the efficiency factor behavior exhibited in Figure 8 for the dimpled plates. The value continues to increase with spacing, exceeding 1.0 at a spacing of 6.88 mm.

This is in contrast to the behavior exhibited by the protrusion roughness cases which exhibit peaks in both drag and heat transfer for certain element spacings due to recirculation zones and overlapping wake regions behind roughness elements. The relative magnitudes of C_f and St for the hemispherical protrusion plates results in a nearly constant value of η of about 0.5 as seen in Figure 9. The rectangular roughness exhibits different behavior depending on the orientation of the elements relative to the surrounding boundary-layer plate. The orientation (i.e. tops of the elements flush with or

protruding above the boundary-layer plate) has a significant effect on the drag, but little effect on the heat transfer. In terms of efficiency, as seen in Figure 10, the protruding rectangular elements produce a nearly constant value of about 0.6, with little effect due to spacing. The flush rectangular elements, on the other hand, produce nearly equivalent heat transfer with lower drag penalty, and produce efficiency values that increase with closer spacing with a peak value of about 0.82.

In order to facilitate direct comparison, the efficiency factor results for all surfaces are plotted together versus roughness element spacing in Figure 11. The dimpled surfaces exhibit the highest values of η, which can be explained in terms of the individual roughness geometries and their effect on total friction and heat transfer. The hemispherical and rectangular protrusions would be expected to have a higher form drag, relative to the dimple roughness, because the protrusions extend upward from the surface into the flow field. All three geometries, however, produce nearly the same levels of turbulence intensity which results in similar levels of heat transfer augmentation for all three roughness types (the dimples create the highest turbulence levels at the surface, due to the vorticity generated in the depressions as discussed by Ezerskii and Shekhov [1990]). Additionally, the protrusions actually have a damping effect on heat transfer at the highest roughness densities, because of the recirculation zones in the spaces between the roughness elements. In this flow regime, described by Morris [1963] as quasi-smooth flow, the interaction of the boundary layer with the roughness elements is minimized, thus reducing turbulence intensity, form drag and heat transfer as compared to most of the surfaces tested with more widely spaced elements (Klett and Kithcart [1992]).

These effects are illustrated in Figure 12, showing sample turbulence intensity profiles for Plates D1, H1, and R1, representing the highest roughness density tested for each element type (see Tables 1, 2, and 3 for dimensions). The velocity measurements were obtained using a Dantec 55P14 right angle hot wire probe in conjunction with a Dantec model 55D01 anemometer, which allowed measurement of profiles down to the surface, between the protrusion elements. The profiles shown were taken at the rear of the plates (2.54 cm from the back edge), at a velocity of 27 m/s. As seen in Figure 12, significantly greater turbulence is generated by the dimple roughness near the wall, while there is less momentum loss in the boundary layer near the wall compared to the higher momentum loss near the wall for the two protrusion roughness types, as illustrated in Figure 13, which shows velocity profiles for the plates having the highest skin friction (plates H5 and R5), along with the dimpled surface at the same density (D5).

The above explanation is best exemplified by the results for the rectangular roughness, Figure 10. Of the two positions tested, the flush position generates the higher efficiency factor values, a result of the roughness elements not protruding into the flow field, thus reducing the form drag contribution of the elements to the total plate friction. The heat transfer is slightly lower in this position, but the reduction is not as great as the reduction in form drag. Also of note is the peak in η, created by the changes in the flow field as described above and by Morris [1963], as the roughness density increases (value of A_r/A_p decreases).

CONCLUSION

Skin friction and heat transfer data have been presented in the form of the efficiency factor, $\eta = (St_r/St_s)/(C_{fr}/C_{fs})$, for turbulent boundary-layer flow over flat plates covered with three-dimensional roughness elements. It has been shown that hemispherical dimples significantly augment heat transfer with a lower penalty in increased skin friction relative to protrusion roughness. Dimples are therefore considered to be better suited for heat transfer enhancement in forced convection applications where pressure losses are an important consideration.

ACKNOWLEDGEMENTS

This work was supported by the United States Air Force Systems Command, Flight Dynamics Directorate, Wright Laboratory, under Contract Number F33615-C-87-3022. The authors wish to acknowledge Dr. Edmund G. Brown-Edwards for his encouragement and assistance during the conduct of this research.

NOMENCLATURE

A_s	Test Plate Surface Area
A_p	Projected Frontal Area of Roughness Elements
A_{sp}	Smooth Plate Area not Covered by Dimples

A_{fp}	Foot Print Area of Dimples
C_f	Skin Friction Coefficient
C_p	Specific Heat of Air
d	Dimple depth
D	Dimple Diameter
e	Roughness height
e^+	Roughness Reynold's number
F	Drag Force
Q	Convective Heat Transfer
St	Stanton Number
T_w	Average Plate Surface temperature
T_∞	Freestream Air Temperature
U_∞	Freestream Velocity
u^*	Friction velocity = $(\tau_o/\rho)^{1/2}$
ρ	Density of Air
τ_o	Average wall shear stress
ν	Kinematic viscosity

REFERENCES

Coleman, H. W. and Steele, W. G., 1989. *Experimentation and Uncertainty Analysis for Engineers*, John Wiley & Sons, New York.

Dipprey, D. F. and Sabersky, R. H., 1963. "Heat Transfer in Smooth and Rough Tubes at Various Prandtl Numbers," *Int. Journal of Heat and Mass Transfer*, Vol. 6, pp. 329-353.

Dirling, R. B., 1973. "A Method for Computing Rough Wall Heat Transfer on Re-entry Nosetips," AIAA Paper 73-763.

Dvorak, F. A., 1969. "Calculation of Turbulent Boundary Layers on Rough Surfaces with Pressure Gradient," *AIAA Journal*, Vol. 7, pp. 1752-1759.

Ezerskii, A. B. and Shekhov, V. G., 1990. "Visualization of the Heat Flux in Flow over Isolated Spherical Depressions," Fluid Dynamics - A Translation of Izvestya Akademii Nauk SSSR, Mekhanika Zhidkosti i Gaza, No. 6, pp. 161-164, Plenum Publishing Co.

Finson, M. L. and Clark, A. S., 1980. "The Effect of Surface Roughness Character on Turbulent Re-entry Heating," AIAA Paper 80-1459.

Hosni, M. N., Coleman, H. W., Taylor, R. P., 1991. "Heat Transfer Measurements and Calculations in Transitionally Rough Flow," *Transactions of the ASME, Journal of Turbomachinery*, Vol 113, pp. 404-411.

Klett, D. E. and Kithcart, M. E., 1992. Uniform Roughness Studies, USAF Flight Dynamics Directorate, Wright Laboratory Report WL-TR-92-3041.

Morris, H. M., 1963, *Applied Hydraulics in Engineering*, Roland Press Co., New York.

Nikuradse, J., 1950. Laws for Flows in Rough Pipes, NACA TM 1292.

Olsen, G. C. and Smith R. E., 1985. Analysis of Aerothermal Loads on Spherical Dome Protuberances, *AIAA Journal*, Vol. 23, No. 5, pp. 650-656.

Schlichting, H., 1937. Experimental Investigation of the Problem of Surface Roughness, NACA Translation, TM 823.

Schlichting, H., 1979. *Boundary Layer Theory*, 7th Edition, p. 555, McGraw-Hill Inc., New York.

Simpson, R. L., 1973. A Generalized Correlation of Roughness Density Effects on the Turbulent Boundary Layer, *AIAA Journal*, Vol. 11, pp. 242-244.

Taylor, R. P., Coleman, H. W. and Hodge, B. K., 1984. A Discrete Element Prediction Approach for Turbulent Flow over Rough Surfaces, Mech. and Nuc. Eng. Dept., Mississippi State Univ., Rept TFD-84-1.

Taylor, R. P. Coleman, H. W., and Hodge, B. K., 1985. "Prediction of Turbulent Rough-Wall Skin Friction Using a Discrete Element Approach," *Journal of Fluids Engineering*, Vol. 107, pp.251-256.

Taylor, R. P., and Hodge, B. K., 1992. "Fully-Developed Heat Transfer and Friction Factor Predictions for Pipes with 3-Dimensional Roughness," *Fundamentals of Forced Convection Heat Transfer*, M. S. Ebadian, and P. H. Oosthuizen, Eds., ASME Symposium Vol. HTD-210, ASME, New York, pp. 75-84.

Webb, R. L., Eckert, E. R. G., and Goldstein, R. J.,1971. "Heat Transfer and Friction in Tubes with Repeated-Rib Roughness," *Int. Journal of Heat and Mass Transfer*, Vol. 14, pp. 601-617.

Webb, R. L., 1994. *Principles of Enhanced Heat Transfer*, John Wiley & Sons, Inc., NY.

Wieghardt, K., 1943. "Erhöhung des turbulenten Reibungswiderstandes durch Oberflächenstörungen," Techn. Berichte 10, No. 9.

TABLE 1 - Parameters For Dimpled Roughness Plates

Plate No.	Hole Spacing (mm)	Avg Hole Depth (mm)	A_{tp}/A_{tp}*d/D
D1	3.45	1.62	0.274
D2	3.71	1.44	0.351
D3	3.96	1.37	0.438
D4	4.22	1.73	0.706
D5	4.45	1.58	0.765
D6	4.88	1.41	0.915
D7	5.64	1.32	1.303
D8	6.88	1.44	2.329

TABLE 2 - Parameters For Hemispherical Roughness Plates

Plate No.	Element spacing (mm)	Avg Element Height (mm)	A_t/A_p
H1	3.45	1.56	3.15
H2	3.71	1.74	3.17
H3	3.96	1.81	3.43
H4	4.22	1.44	5.19
H5	4.45	1.60	5.04
H6	4.88	1.76	5.37
H7	5.64	1.86	6.81
H8	6.88	1.74	10.97

TABLE 3 - Parameters For Rectangular Roughness Plates

Plate No.	Element Spacing (mm)	Groove Width (mm)	A_t/A_p
R1	3.28	0.79	2.74
R2	3.68	1.19	3.45
R3	4.08	1.59	4.26
R4	4.47	1.98	5.05
R5	4.87	2.38	6.02
R6	5.66	3.18	8.07

(Height is 1.59 mm and base dimensions are 2.49 mm x 2.49 mm for all rectangular elements)

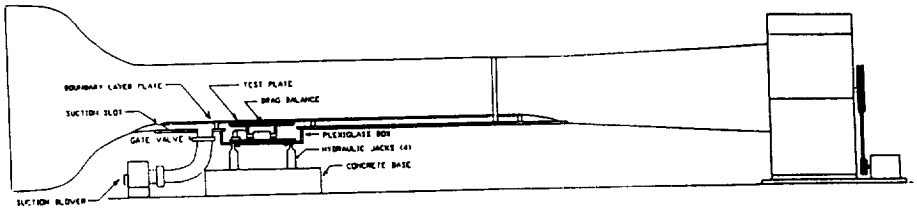
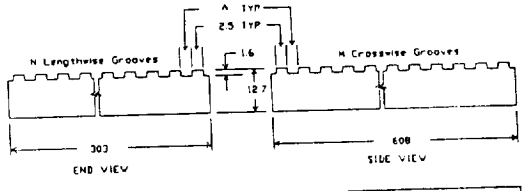


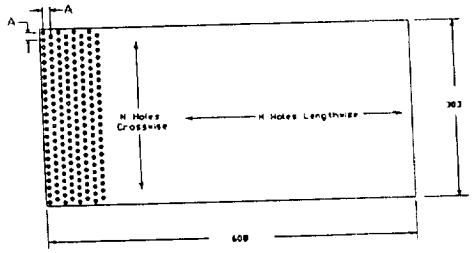
Figure 1. Schematic of Wind Tunnel and Drag Balance



NOTES:
Not to Scale
Dimensions in mm
12.7 mm Thick Aluminum Plate
1.59 mm Deep Grooves Milled Lengthwise and Crosswise
With Spacings as Indicated

Plate No.	A	N	M
R1	0.79	91	184
R2	1.19	81	164
R3	1.59	73	147
R4	1.98	67	135
R5	2.38	61	124
R6	3.18	53	106

Figure 2. Schematic of Rectangular Roughness Plates



NOTES:
Not to scale
Dimensions in mm
12.7 mm Aluminum Plate
Hemispherical Holes
Nom. 3.18 mm Dia. x 1.59 mm Deep
With Spacings as Indicated

Plate No.	A	N	M	No. of Holes
D1	3.45	87	175	15138
D2	3.71	81	163	13122
D3	3.96	76	153	11552
D4	4.22	71	144	10152
D5	4.45	68	136	9188
D6	4.88	62	124	7626
D7	5.64	53	107	5618
D8	6.88	44	87	3795

Figure 3. Schematic of Hemishpherical Roughness Plates

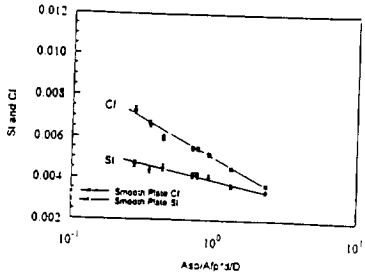


Figure 4. Cf and St for Dimpled Plates

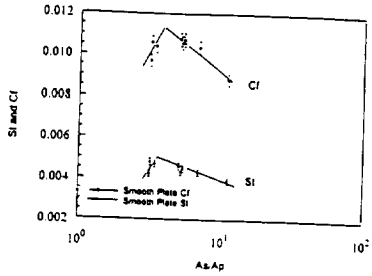


Figure 5. Cf and St for Hemisphencal Roughness Plates

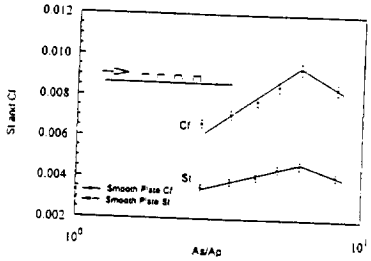


Figure 6. Cf and St for Rectangular Roughness Plates With Elements Protruding Above BL Plate

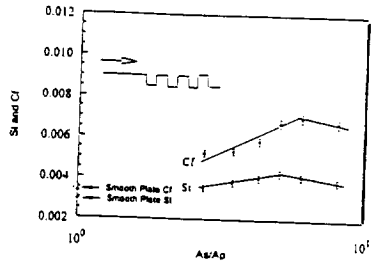


Figure 7. Cf and St for Rectangular Roughness Plates With Elements Flush With BL Plate

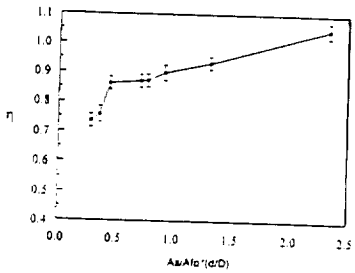


Figure 8. Efficiency Factor for Dimpled Plates

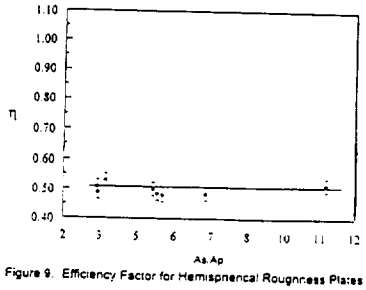


Figure 9. Efficiency Factor for Hemispnencal Roughness Plates

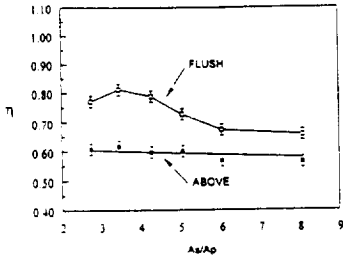


Figure 10 Efficiency Factor for Rectangular Roughness Plates

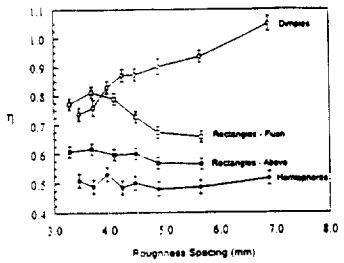


Figure 11 Efficiency Factor for all Surfaces Versus Spacing

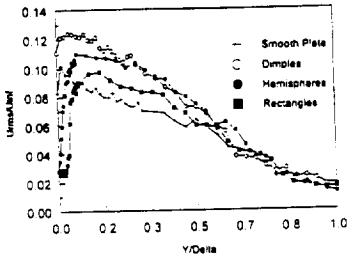


Figure 12. Sample Turbulence Profiles

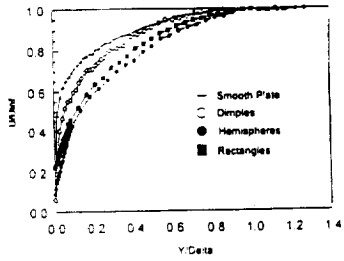


Figure 13 Sample Velocity Profiles

Fabrication, Testing and Optimization of Sandwich Composites

by

Mr. Michael D. Walton, Graduate student

William J. Craft, Professor of Mechanical Engineering

NASA Center of Research Excellence for Aerospace Research
North Carolina A&T State University
1601 East Market St.
Greensboro, NC 27411

Abstract

The optimum design of a composite sandwich structure involves the determination of both core and face thickness for minimum weight. Optimization analysis is based on a specified parameter such as strength or stiffness. The objective of this research is to develop an analytical tool to optimize and predict the performance of a composite sandwich given its component material properties. Results will be compared with experimental data obtained in accordance with ASTM Standards.

Introduction

A structural sandwich [1,2,3] consists of two high strength face sheets separated and bonded to a relatively thick low density core. This configuration increases the flexural rigidity of the structural member by increasing its moment of inertia. Substantial weight savings is achieved through the incorporation of the low density core. Examples of sandwich structures can be found in the building industry, surfboards, and aircraft structures where minimum weight considerations are paramount. Thus, an optimized sandwich structure is a minimum weight design that satisfies a specific strength or stiffness requirement.

Several failure mechanisms must be considered in a composite sandwich beam subjected to three point loading. First, wrinkling or buckling may occur in the face sheet where the beam is in compression. The core material must withstand the shear forces that develop in the beam. Compressive yield of the core can also occur in areas close to the applied load. During construction of the sandwich beam, the face sheets are bonded to the foam core. Delamination at the bond line is another possible failure mode. A properly designed sandwich beam is optimized when two or more modes of failure occur at approximately the same load.

Biography Mr. Michael D. Walton is a Master of Science student in the Department of Mechanical Engineering at North Carolina A&T State University. His thesis work is in sandwich composites within the within the Mechanics and Materials Group. His undergraduate degree is in mechanical engineering from North Carolina A&T State University. He has an interest in developing test methods for sandwich panels and beams, and for developing new materials systems for high performance aircraft. Dr. Craft is the faculty advisor, Dr. V. S. Avva and Dr. K. N. Shivakumar serve on the graduate advisory committee of Mr. Walton. These faculty share an interest in sandwich composites and in developing experimental methods for verifying the integrity of these structures.

The work of Allen [1] provides a solid foundation for the analysis of sandwich beams. Failure criteria for the foam core sandwich is explored by Gibson and Triantafillou [3]. Lingard [2] develops an approach to the optimization of sandwich composites. Optimization for stiffness, strength and weight of foam core sandwiches is developed by Gibson and Triantafillou [3,4,5].

The objective of this research is to develop an analytical procedure to optimize and predict the performance of composite sandwich beams with a stratified core. The work involves a review of the literature on sandwich structures in view of new material systems. Quality experimental data is dependant on consistent fabrication techniques and testing procedures. This data will be compared with revised optimization equations that model a sandwich beam with a stratified core.

For the current paper, sandwich beams have been constructed and tested. The face sheets are formed by compression molding carbon fiber reinforced plain weave fabric. The core material is a Polymethacrylimide (PMI) rigid cellular foam. The bonding adhesive is Epoxy 828 and Jeffamine or structural film adhesive AF 163-2.

Analysis

Typical sandwich beam construction is shown in the Figure on the last page. Face sheet and core thickness are given by t and c, respectively. The distance between the neutral axis of each face sheet is denoted by d, and h is the overall thickness of the sandwich beam. The beam has width b and a span length l, where l is the distance between the supports.

The flexural rigidity D of a homogeneous beam is the product of its Young's Modulus E and the second moment of area I [1].

$$D = EI \tag{1}$$

The sandwich beam is a composite structure. The flexural rigidity of the sandwich beam is the sum of the flexural rigidities of its components about the centroid of the beam [1].

$$D = \frac{1}{12}E_c bc^3 + \frac{1}{6}E_f bt^3 + \frac{1}{2}E_f btD^2 \tag{2}$$

An approximation for the flexural rigidity of a sandwich beam is appropriate under the following conditions. The second term represents the local stiffness of the faces and is generally small compared to the third term in equation 2. The first term in equation 2 represents the bending stiffness of the core. Normally, the contribution of this term to the overall stiffness of the sandwich is small. Therefore, the flexural rigidity of the sandwich beam may be approximated by equation 3 [1].

$$D = \frac{1}{2}E_f btd^2 \tag{3}$$

The shear force V and bending moment M at any point along the beam are obtained from the shear and bending moment diagrams. The bending moment is maximum at the midspan of the beam.

The normal stress at a distance z from the neutral axis is obtained from bending theory [1,3]. The maximum normal stress occurs at the midspan of the beam and at

the outer most fiber of each component.

$$\sigma_f = \pm M\frac{E_f h}{D\,2} \quad ; \quad \sigma_c = \pm M\frac{E_c c}{D\,2} \qquad (4)$$

The shear stress distribution in a homogeneous beam is given by:

$$\tau = \frac{V I}{Q b} \qquad (5)$$

The shear force is V, and Q is the first moment of area above the position where the shear stress is wanted. For a rectangular beam Q is the product of the area above this selected point, times the distance from the centroid of the beam to the centroid of this area.

Typically, the shear stress formula for a sandwich beam must be modified to account for the differences in the Young's Modulus of its components.

$$\tau = \frac{V}{Db}\Sigma(QE) \qquad (6)$$

If the core does not provide a significant contribution to the flexural rigidity of the sandwich, the shear stress may be assumed constant thru the thickness of the core [1].

Optimization for minimum weight sandwich beams is rooted in an analysis of strength or stiffness parameters. Previous work provides insight for the optimization analysis of composite sandwich beams. Lingard [2] develops theory for the optimization for sandwich composites based on strength and stiffness. Gibson and Triantafillou [4,5] develop failure mode maps or graphical representations of the failure modes that occur simultaneously in an optimized sandwich structure. The failure mode maps are used in the optimization of foam core panels for a given strength. Gibson [3] develops equations for the optimization of stiffness in foam core sandwich beams.

The current research goal is to develop an optimization procedure for composite sandwich beams with stratified or variable density core. The scope of this work is two fold.

The interface between failure modes must be investigated. Failure of the sandwich beam should occur under the influence of several failure mechanisms.

The stratified or variable density core technology will be an integral part of the minimum weight optimization . Rigid foam core is available in different densities. The strength of the foam core is a function of density. A stratified core is simply a core made of several layers. Each layer has a different density and is uniformly thick.

The variable density core is different from a layered core. In the variable density core, the density changes continuously thru the thickness.

Face sheet thickness is a third variable in the optimization equation.

Fabrication

Composite sandwich fabrication is a three step process. The composite face sheets must be molded. The core material must be machined to appropriate dimensions. Finally the face sheets must be bonded to the core material. All elements of the sandwich are constructed with dimensions

of the final specimen in mind.

The face sheets are compression molded carbon fiber reinforced plain weave fabric with 934 epoxy resin. Cold storage insures maximum shelf life.

Initially, the prepreg is removed from cold storage and allowed to reach room temperature. To minimize moisture content the material is stored in an air tight nylon bag. Condensation that collects on the storage bag is removed during the conditioning process. Upon reaching room temperature the prepreg is removed from the storage bag. Each ply is cut from the roll of material. Care is taken to preserve the fiber orientation of the prepreg. Each ply is marked to indicate its orientation to the roll.

Place the first ply of the laminate on a clean flat surface. Applying firm pressure with a soft roller, flatten the initial ply. Each subsequent ply is placed on the stack and the roller is used to flatten the panel and remove as much air as possible between the sheets. Each ply should be oriented properly on the stack.

Before the panel can be molded in the compression press, it must be properly sealed. The panel is removed from the mylar film and placed between sheets of release cloth. The release cloth is a Teflon coated fiberglass that will not bond to molded specimen. The release cloth and panel are sandwiched between two sheets of breather cloth. Breather cloth allows air to escape from the surface of the panel As a result, the surface quality of the molded panel is improved. The entire package is sealed in a nylon bag. The nylon bag collects the resin that flows out of the panel.

The panel is placed in the compression press. The mold pressure is set to (689.48 kPa).

The heating cycle starts at room temperature. Both platens are heated to (177 C) at a rate consistent with the material data sheet. The target temperature is held for two hours. The panel is cooled and remove from the press.

The sealed nylon bag must be inspected to assure that the resin flow is uniform.

Core materials should be the same size as the composite panels. The actual test specimens will be machined from the sandwich panel.

Bonding of the face sheets to the core is the final step in sandwich construction. Possible candidates for the adhesive are AF 163-2 or Epoxy 828 and Jeffamine with silicon flower. AF 163-2 is a structural film adhesive that is heat activated. The Epoxy 828 and Jeffamine with silicon flower is cured at room temperature for 24 hours.

The sandwich components are placed on a flat surface, and compressed in a vacuum (30 in Hg) until the cure cycle is complete.

The fabrication procedure is complete. It is important to make the sandwich panels large enough to allow for machining of the test specimens.

Testing

Mechanical testing is used to determine the properties of the composite sandwich and its components. The mechanical properties of the face sheet material and core are used in the optimization models. The actual sandwich specimens are tested to gauge the accuracy of the optimization

theory.

The tensile and compressive properties of the composite skins are determined from ASTM standards. The standard for the tensile properties of fiber-resin composites is D 3039-76. Compressive properties of crossply fiber resin composites are obtained from D 3410-87. The flexural properties of flat sandwich structures are obtained by ASTM standard C 393. The in plane shear properties of a flat sandwich can be obtained from ASTM standard C 273-61.

Preliminary testing was done on a few sandwich beams with homogeneous foam cores. The testing was done to get a feel for the overall performance of sandwich beams made from these materials.

The first failure mode to occur was compressive yield of the core under the midpoint load. The forces are applied as line loads across the width of the beam. The stress close to the load points is very high. Testing continued until the specimen failed along the bond line.

Conclusions

The literature search of related topics is nearly complete. The previous work provides a solid foundation for the advancement of this research.

Fabrication techniques for have been developed. This includes molding of the composite face sheets. Manipulation of the core material. Fixtures for sandwich construction have been built, and sandwich specimens have been prepared.

Preliminary testing has been done on a few of the sandwich specimens. The premature failure of the core as a result of the loading, must be investigated further.

Future work includes testing of the sandwich components to determine the mechanical properties. Development of the optimization theory is needed in order to predict the dimensions for each sandwich component. Specimens will be tested to evaluate the accuracy of the analytical models.

The author wishes to thank Dr. V. S. Avva and Professor Sadler of the Center for Composite Materials Research for their guidance, and use of the facilities. Invaluable assistance from Dr. Craft is also greatly appreciated.

References

1. Howard G. Allen, Analysis and Design of Structural Sandwich Panels, (London : Pergamon Press, 1969).

2. J. R. Lingard, Development of a Theory For Optimising Sandwich Composites, *Proceedings of the International Conference on Advanced Composite Materials*, 1993.

3. L. J. Gibson, Optimization of Stiffness in Sandwich Beams with Rigid Foam Cores, Materials Science and Engineering, 67 (1984), 125-135.

4. T. C. Triantafillou, L. J. Gibson, Failure Mode Maps for Foam Core Sandwich Beams, Materials Science and Engineering, 95 (1987), 37-53.

5. T. C. Triantafillou, L. J. Gibson, Minimum Weight Design of Foam Core Sandwich Panels For a Given Strength, Materials Science and Engineering, 95 (1987), 55-62.

6. Ferdinand P. Beer, E. Russell Johnston, Jr., Mechanics of Materials, (McGraw-Hill Book Company, 1981).

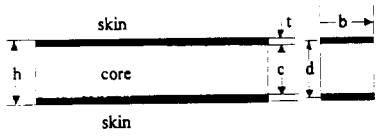
7. Standard Test Method for Compressive Properties of Unidirectional or Crossply Fiber-Resin Composites, ASTM D 3410 - 87.

8. Standard Test Method for Tensile

Properties of Fiber-Resin Composites, ASTM D 3039 - 76 (Reapproved 1989).

9. Standard Test Method for Flexural Properties of Flat Sandwich Constructions, ASTM C 393 - 62 (Reapproved 1988).

10. Standard Test Method for Shear Properties in Flatwise Plane of Flat Sandwich Constructions or Sandwich Cores, ASTM C 273 -61 (Reapproved 1988)



Sandwich Beam Construction

Figure - Sandwich Beam with Side and End Views of Construction

Keywords: Composite Materials, Sandwich Shells, Stratified core, variable density core

555-61

The Utilization of the MACSYMA Software Package in Multidisciplinary Optimization Problems

290719

P7

Wanita Dunston and Dr. Meldon Human
NASA Center of Research Excellence
North Carolina A&T State University
Greensboro, North Carolina

Keywords: MDO, optimization, aerospace design, modeling, simulation

Abstract

Multidisciplinary Optimization (MDO) is an important consideration in many design situations. MDO involves the simultaneous manipulation of design variables across interdisciplinary lines, such as structures and fluids. In an attempt to create balanced designs, the necessity of a means of computer based optimization becomes apparent. Limited time, manpower, and/or resources result in the need to eliminate the tedious "cut and try" methods, and an ultimate design to improve the quality for equal design time.

MDO requires several mathematical manipulations, such as model reduction and approximations, derivative calculations, etc. These mathematical manipulations may be performed using commercially available software packages. The utilization of these software packages for analysis is quite effective. Our objective in such analysis is to obtain a system/component which satisfies design and performance requirements for a set of given input factors and considerations. We seek a means of optimization which possesses reasonably efficient and productive run times. Our goal in this endeavor is to develop a tool which applies to MDO problems, and to finally test this tool on an atypical MDO situation. MACSYMA, a software package, will be used to accomplish this.

I. NEED FOR MDO

MDO has found great use in the aerospace industry. Figure 1 shows a flow diagram which shows a flow diagram which shows the relationship among various disciplines for such applications. In design, we often encounter variables which involve myriad fields of study, thus, the term: "multidisciplinary". In the pursuit of a balanced design, the engineer may be faced with more design topics and functions than he/she can expertly handle. Since it is necessary to efficiently and economically yield results, the need for a means of "freeing the engineer to think" becomes self-evident. Mathematical operations may serve as a vehicle to reach this goal.

II. MDO ANALYSIS

MDO can be defined as a methodology for design of complex engineering systems that are governed by mutually

[1] Biography, Wanita Dunston

interacting physical phenomena and made up of distinct interacting subsystems. The MDO methodology exploits the state of the art in each contributing engineering discipline and emphasizes the synergism of the disciplines and subsystems. In large systems, since the optimization and analysis of a given engineering system may differ significantly from those routinely encountered in application to a single part of the system, or in a single discipline, MDO is necessary. In these "large systems", the number of constraint functions, design variables, and unknowns may be quite large. We must also consider the existence of qualitative differences. This refers to the fact that in general, a single mathematical model which can handle the entire engineering system does not exist. What we do have, in fact, is an assemblage of several mathematical models, each with different functions/purposes, used to address a certain physical part of the system, a specific engineering discipline, or a particular physical phenomenon. The engineering analysis is therefore, iterative, as each model depends on the other for input through the usage of computer codes, which in turn, form a modular software system.

MDO involves several steps. For each discipline, we often need to simplify the governing fundamentals into a reduced, tractable model. This may be referred to as approximation theory. Linking the various disciplines involves relating changes in one set of variables to how other variables are altered. This is accomplished via sensitivity analysis, a formal procedure coupling partial derivatives with respect to design variables. A major aspect of MDO is optimization, and involves mathematical techniques to minimize (maximize) a defined objective subject to predescribed constraints. In Figure 2, we show a diagram from Reference 1 which shows a typical MDO flow.

[2] MACSYMA, Inc., Arlington, MA.

III. MATHEMATICAL OPERATIONS

There are a number of mathematical operations required for performing MDO functions. We shall highlight optimization. It is relevant to note that optimization may be performed through a series of mathematical operations, including (but not limited to) Newton's Method, Sequential Linear Programming, Design Variable Linking, and Sequential Quadratic Programming. For unconstrained function minimization, Newton's Method be used. This is a two-step repetitive process: after we find a usable search direction, we may then perform a predimensional search. Alternately, for constrained function minimization, sequential linear programming are two approaches available. Design variable linking involves the usage of approximate symbolic solutions of equations, and numerical optimization, as well.

IV. MACSYMA

MACSYMA[2] is an interactive, symbolic algebra program that assists the user in solving complex mathematical problems, offering a wide range of capabilities, including differentiating and integrating expressions, factoring polynomials, plotting expressions, solving equations, manipulating matrices, and computing Taylor series. MACSYMA is also a

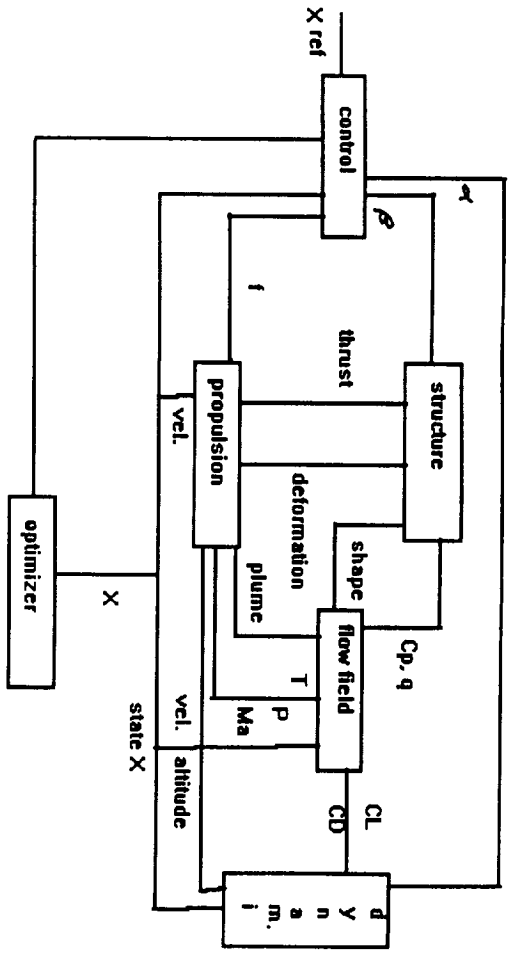


Figure 1 Interdisciplinary Flow Chart

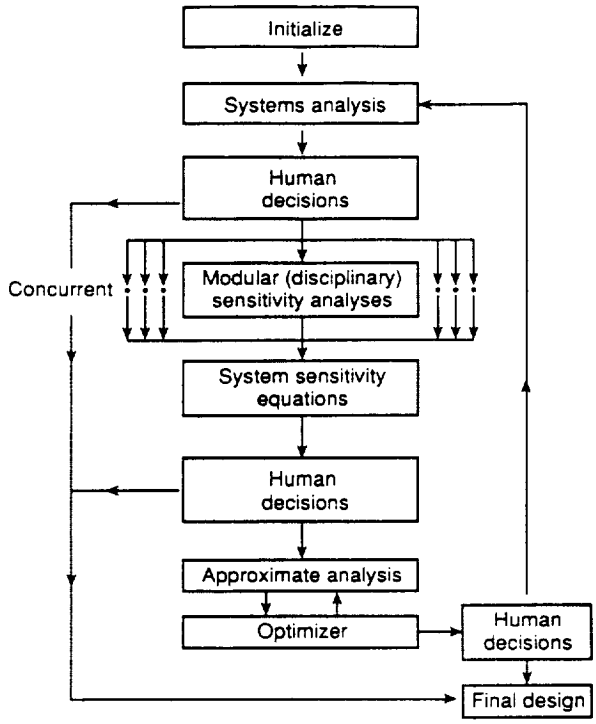


Figure 2. An optimization procedure for a non-hierarchic system (note an opportunity for concurrent processing at the level of disciplinary sensitivity analyses).

programming environment in which the user has the power to define virtually any mathematical procedures tailored to his/her personal needs. With the above description of MACYSMA in mind, we chose to use this package for our study.

V. MACSYMA-MDO OPERATION MATCHING

With a basic working knowledge of the MACSYMA software package's capabilities, and an understanding of basic multidisciplinary optimization on operations, a correlation may be developed. This correlation involves matching of MDO operations with MACSYMA's intrinsic mathematical operations. Figure 3 shows a matching between some required operations and corresponding MACSYMA functions. A list of cross references has been prepared for easy access and quick indexing. This list may serve as the predecessor to the final algorithm used to perform MDO test examples in the near future.

VI. THE OVERALL PROGRAM

The purpose of this study will be to construct a working MDO model which uses MACSYMA capabilities to perform MDO functions. The model will be applied to an aerospace design problem involving several cross disciplines.

REFERENCES

1. Sobieszczanski-Sobieski, J., "Multidisciplinary Design Optimization: An Emerging New Engineering Discipline". World Congress on Optimal Design, Brazil, 8/93.

2. Hillier, F.S., G.J. Lieberman, Operations Research, Holden-Day, San Francisco, CA, 1974.

3. Luenberger, D.G., Introduction to Linear and Nonlinear Programming, Addison-Wesley, Reading, MA, 1973.

Figure 3: Cross References

MDO Function	MACSYMA Operation
A nonlinear, constrained optimization task	~Solving Non-linear Equations ~Optimization of Symbolic Expressions for Numerical Evaluation ~Approximate Symbolic Solutions of Equations
Iteration	~Using Iteration
Vector Search Direction	~Plotting Vector Fields in 3-D
Calculation of S^q gradients	~Generating Finite Difference Approximations
Newton's Method	~Newton's Method
Polynomial Interpolation	~Simplifying Functions ~Exact Roots of Polynomials ~Interpolation of Curves
Golden Section Method	~Fibonacci Numbers
Linear Programming	~Equalities ~Inequalities ~Solving Linear Equations
Sequential Unconstrained Minimization Techniques	~Functions with a Variable Number of Arguments
Reduced Gradient Methods	~Generating Finite Difference Approximations
Augmented Lagrange Multiplier Method	~Other Number-Theoretic Functions
Hessian Matrix, B	~Generating Finite Difference Approximations
Direct Linking	~Basic Matrix Arithmetic Operations ~Determinant, Inverse, and Related Matrix Functions ~Solving Matrix Equations ~Eigenf Package for Iterative Eigen-Analysis of Matrices
Design Variable Linking	~Approximate Symbolic Solutions of Equations ~Numerical Optimization
Linear Approximation Example	~Numerical Solutions of Systems of PDE's
Nonlinear Approximation Example	~Numerical Optimization
Rayleigh Quotient Approximation	~Finding Eigenvalues and Characteristic Polynomials for Eigenvalues

Footnote: Biography for Wanita Dunston, Principal Author

Wanita Evette Dunston, a native of Oxford, North Carolina, is a Senior Mechanical Engineering major at NC A&T State University. Throughout her undergraduate career, Ms. Dunston has been involved with the American Society of Mechanical Engineers (ASME), and the National Society of Black Engineers (NSBE). In addition, she has over one year's work experience working with the Environmental Protection Agency, (US EPA), and has been published through a national American Institute of Aeronautics and Astronautics (AIAA) conference. Her research interests are in the area of thermal sciences, and she will graduate in December 1996.

356-05

290720 9

A FUZZY METRIC FOR COOPER-HARPER FLYING QUALITY MEASURES

Dara Strickland[1] and Celestine A. Ntuen
Human-Machine Systems Engineering Group Center
for Aerospace Research, NASA URC
College of Engineering
North Carolina A&T State University
Greensboro, NC 27411

ABSTRACT: A flight handling quality (FHQ) is a composite measure of performance used to assess how well a pilot responds to flight conditions. For many years the Cooper-Harper (CH) FHQ has been considered as the industrial standard. Because the CH-FHQ consists of opinions characterized by vagueness and impreciseness, fuzzy set models can be used to aggregate opinions expressed on the CH-FHQ scale. This paper presents a fuzzy composite measure for CH - FHQ. The metric is robust and insensitive to the scale of judgments.

1. INTRODUCTION

For many years, the Cooper-Harper (1996) flying handling quality measures (FHQ) have been considered a standard for assessing the pilot performance during and after flight. An FHQ consists of subjective measures of demands on pilot in selected tasks or operations. Once a rating database has been collected, it is possible to correlate the pilot perceptions with the task and aircraft performance. An example Cooper-Harper (CH) FHQ scale is shown in Table 1.

The CH - FHQ consists of opinions that are vague and imprecise on linguistic scale of preference. Many pilot performance models that use the CH - FHQ (see, e.g., Craig and Ashkenars, 1971; Shomber and Gertsen, 1967; Hess, 1977) have failed to consider the elements of ambiguity in their FHQ-based performance metric. In order to consider this ambiguity and variations in scale of judgments, we have propose the use of fuzzy theory as a tool for data aggregation. The fuzzy FHQ is, by properties of fuzzy models, robust and sensitive.

2. FUZZY THEORY

Zadeh (1973, 1975) introduced the theory of fuzzy set to address the issues associated with vagueness and impreciseness by hedging subjective opinions on a cognitive scale of preference. Eshrag and Mamdami (1979) developed a general approach to linguistic approximation to weight the "behavioral preference" of choice in multi attribute decision making problem. Others, for example, Baas and Kwakernaak (1977) developed methods for ranking subjective alternatives. In general, fuzzy metrics, developed on subjective scales, are known to follow certain laws of comparative judgment (Thurstone, 1927).

The fundamental definitions of a fuzzy set theory are given as follows (Zadeh, 1973):

Let X = {x} be a set of attributes, then a fuzzy set A $\in$ X is a set of ordered pairs

$$A = \{x, \mu_A(x)\}, x \in X \qquad (1)$$

where $\mu_A(x)$ is called the characteristic function or graded membership of x in A (Zadeh, 1975). The membership function $\mu_A(x)$ maps the fuzzy set A onto the interval [0, 1], that is μ_A: A $\rightarrow$ [0, 1]. Similarly, let Y = {y} be a set of criteria variables, then a fuzzy set B$\in$Y is a set of ordered pairs

$$B = \{y, \mu_B(y)\}, \quad y \in Y \qquad (2)$$

and μ_B: B$\rightarrow$ [0, 1]. Note that $\mu_A(x)$, $\mu_B(y)$ can be assumed to have a known distribution - mathematically or perceptually. The fuzzy distribution can be a real continuous phenomenon or may represent a discrete countable event. The interaction of A and B is defined by:

$$\mu_A(x) \wedge \mu_B(x) = \min \{\mu_A(x), \mu_B(x)\} \qquad (3)$$

The union of A and B is defined by:

$$\mu_A(x) \cup \mu_B(x) = \max \{\mu_A(x), \mu_B(x)\} \qquad (4)$$

[1] Dara Strickland is a graduate of NC A & T State University. She earned her BS in Industrial Engineering on May 1995. She is currently pursuing a Masters Degree in I.E. specializing in Human Machine Systems Engineering. She is expecting to graduate December, 1996.

The extended maximum operator combines the definitions in (3) and (4). This is defined by :

$$\mu_A(x) \cup \mu_B(x) = Sup\{\min[\mu_A(u), \mu_B(v)]\} \quad (5)$$
$$((u, v / x = \max(u, v)), \forall x \in \Re$$

3. APPLICATION OF FUZZY THEORY TO CH-FHQ

3.1 The Main Concept

The C-H scale in Table-1 classifies FHQ into three levels:

- Level 1: Flying qualities adequate for the mission flight phase.
- Level 2: Flying qualities adequate to accomplish the mission flight phase, but some increase in pilot workload or degradation in mission effectiveness exists.
- Level 3: Flying qualities such that the airplane can be controlled safely, but pilot workload is excessive, or mission effectiveness is inadequate, or both.

The interaction of these levels can be represented in a linguistic geometric space as shown in Figure 1; where A is for Level 1, B for Level 2, and C for Level 3 respectively. In Figure 1, the fuzzy boundary between A and B, and between B and C represent some "occluded" interaction of cognitive opinions; and can be modeled as event interactions. Specifically, let $\mu_A(x)$, $\mu_B(x)$, and $\mu_C(x)$ represent the fuzzy membership of levels A, B, and C, the aggregate cognitive fuzzy rating is given by

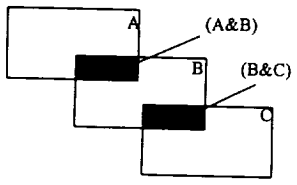


Figure 1: Set representation of FHQ levels with fuzzy boundaries.

$$\mu_S(x) = \mu_A(x) \cup \mu_C(x) - [\mu_A(x) \wedge \mu_B(x)] - [\mu_B(x) \wedge \mu_C(x)] \quad (6)$$

Where $\mu_s(x)$ is the fuzzy membership for describing the FHQ rating.

From equation (6), it is easy to show that

$$\mu_S(x) = \mu_A(x) \cup \mu_B(x) \cup \mu_C(x) - \mu_B(x)[\mu_A(x) \cup \mu_C(x)]$$
$$= \mu_A(x) \cup \mu_B(x) \cup \mu_C(x) - g(x) \quad (7)$$

where g(x) is a fuzzy function describing the degree of overlapping opinions. Therefore, g(x) can be obtained by using the general overlapping function (Eshrag & Mamdami, 1979).

$$\mu_{MN}(x) = \frac{\mu_M(x) \cup \mu_N(x) - \{\mu_M(x) + \mu_N(x)}{\mu_M(x) \wedge \mu_N(x)} \quad (8)$$

$$\mu_{MN}(x) \in (-1, \infty)$$

We want $\mu_{MN}(x) \in [0, 1]$, hence from equation (8), the scaled overlapping function is

$$g(x) = \begin{cases} 1/2(\mu_{MN}(x) + 1), & -1 \le \mu_{MN}(x) < 0 \\ 1 - \dfrac{1}{\mu_{MN}(x) + 1}, & \mu_{MN}(x) \ge 0 \end{cases} \quad (9)$$

3.2 Experiment

An experiment to validate the fuzzy metric was conducted with the Manual Control Laboratory (MCL). The MCL is a software for tracking tasks—both compensatory and pursuit. We experimented with the pursuit tracking at three levels of task difficulties defined by the plant transfer functions: k (easy), k/s (hard) and k/s^2 (very hard). Eight graduate students with an average age of 24.3 years participated. Each subject received the C-H FHQ questionnaire shown on Table 1. A control bandwidth of 0.33HZ, 0.67HZ and 1.0HZ were selected similar to a previous workload study by Moray and Waterton (1983).

4. RESULTS

The fuzzy rating by levels and by task difficulties is shown in Figure 2. This rating represents an aggregate across the control bandwidth.

As suggested, the purpose of this model was to estimate an aggregate fuzzy metric based on C-H FHQ rating scale. Further studies will be needed to:

a) develop a fuzzy prediction equation for each level of performance; b) develop a fuzzy prediction equation for an entire task situation; c) develop fuzzy models for predicting the pilot workload.

REFERENCES

1. Baas, M. S. , and Kwakernaak, H. (1972). Rating and ranking of multiple-aspects alternatives using fuzzy sets. Automatica, 13, 47-58.

2. Cooper, G. E., and Harper, R. P. (1969). The use of pilot rating in the evaluation of aircraft handling qualities. NASA, Tech. Note d-5(53), NASA, Washington, DC.

3. Craig, S. J. and I. L. Ashkenas (1971). Background data and recommended revisions for MIL - 8785 B (ASG), "Military Specifications — Flying qualities of piloted planes", Technical Report TR - 189-1, Systems Technology, Iuco, Hawthorne, CA.

4. Eshrag, F., and Mamdami, E. H. (1979). A general approach to linguistic approximation. International Journal of Man-Machine Studies, 11, 5001-519.

5. Hess, R. A. (1977). Predictions of pilot opinion ratings using an optimal pilot model. Human Factors, 19(5), 459-475.

6. Moray, N., and Waterton, K. (1983). A fuzzy model of rather heavy workload. Proc. 8th. Annual Conference on Manual Control, Dayton, Ohio, 120-126.

7. Shomber, H. A., and Gertsen, W. M. (1967). Longitudinal handling qualities criteria: and evaluation. Journal of Aircraft, 4(4), 371-376.

8. Thurstone, L. L. (1927). A law of comparative judgment. Psychological Review, 34, 273-286.

9. Zadeh, L. A. (1975). The concept of a linguistic variable and its applications to approximate reasoning-I. Information Science, 8, 199-249.

10. Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision process. IEEE Trans. on Systems, Man, and Cybernetics, SMC-3, 28-44.

TABLE-1 Pilot Opinion Rating on CH Scale

Aircraft Characteristics	Demands on Pilot in Selected Task or Required Operation	Pilot Rating	Flying Qualities Level
Excellent; highly desirable	Pilot compensation not a factor for desired performance	1	
Good; negligible deficiencies	Pilot compensation not a factor for desired performance	2	1
Fair; mildly unpleasant deficiencies	Minimal pilot compensation required for desired performance	3	
Minor but annoying deficiencies	Desired performance requires moderate pilot compensation	4	
Moderately objectionable deficiencies	Adequate performance requires considerable pilot compensation	5	2
Very objectionable but tolerable deficiencies	Adequate performance requires extensive pilot compensation	6	
Major deficiencies	Adequate performance not attainable with maximum tolerable pilot compensation Controllability not in question	7	
Major deficiencies	Considerable pilot compensation required for control	8	3
Major deficiencies	Intense pilot compensation required to retain control	9	
Major deficiencies	Control will be lost during some portion of required operation	10	

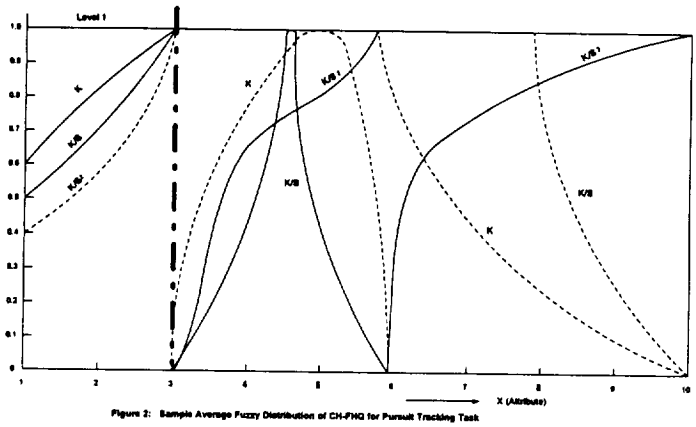


Figure 2: Sample Average Fuzzy Distribution of CH-FHQ for Pursuit Tracking Task

"EFFECTS OF TASK DIFFICULTY ON PILOT WORKLOAD

Alexandria R. Watson
Dr. Celestine Ntuen
Dr. Eui Park
NASA Core
North Carolina Agricultural and Technical State University
Greensboro, North Carolina

Abstract: In this study, we develop a model of workload as a function of task difficulty. Experiments on compensatory tracking tasks are used to assess the task difficulties. First, second, and third order plants that are controlled under three damping frequencies, no oscillation, critically damped, and overdamped, are investigated. The task difficulty index is defined as the ratio of root means square of path error to the root means square control velocity.

Keywords: workload, damping, task difficulty

Introduction

As the characteristics of systems change, the human operator is expected to adapt to these changes and perform the necessary tasks accordingly. Additionally, operator performance becomes a critical issue. According to Lysaght, Dick, Hill, et al. (1988), two major factors affecting performance are the tasks defined by the objectives of the mission (what is to be accomplished) and the environment and the stable and transitory traits of the human operator. Operator performance and system performance are determined by the interaction of these factors. The interaction also determines the operator workload.

Although there is not a universal definition of workload, it can be described as the information processing demands placed on the operator by a task (Sanders and McCormick, 1993). The concept of workload is based on the difference between the amount of resources demanded by the task situation and the amount of resources available by the operator to perform in the task situation (McCloy, Derrick, and Wickens, 1983). In other words, workload can be changed by (1) altering the demands of the task on the operator, (2) altering the amount of resources available within the operator, or (3) a combination of both (Sanders and McCormick, 1993).

Pilot Workload

Within the aviation community, reduction of workload is a top priority specifically for pilots. During a mission, pilots are bombarded with displays that must be monitored. They are continuously interpreting information and making decisions based on these interpretations. They communicate with their crew, other pilots and their crew, air traffic controllers, and etc. Additionally, they may also have to fire or drop weapons on a given target with perfect precision. That they also perform under hostile conditions can add to pilot stress (Wickens, Stokes, Barnett, and Hyman, 1981). Therefore, an effort to create systems that foster acceptable levels of workload in any given situation is imperative to enhancing performance. The relationship between performance and workload as described by Lysaght, Hill, Dick, et al. (1988) is, in short, extremely high or low levels of workload on the human operator results in poor performance. Acceptable performance can be expected at reasonable levels of workload.

Aviation companies are trying to develop and produce aircraft systems that will produce acceptable levels of workload (Hughes, 1989 and Kernstock, 1989). The wave of the future is incorporation automation into the overall design of the cockpit (Hughes, David, and Dornheim, 1995). Quantitative measures of pilot workload and performance are necessary to assist in selecting technologies for full-scale development and integration into new aircraft as well as for the upgrade of current aircraft.

There is no acceptable universal definition of workload, hence, there is not one for pilot workload. However, according to Gawron, Schiflett, and Miller (1989), many definitions contain two common elements. The first is "what the pilot is required to

---

† **Biography:** Alexandria R. Watson, a First Lieutenant in the U.S. Air Force, is a graduate of North Carolina A & T State University. She earned her commission as well as B.S. in Industrial Engineering May, 1993. She is currently pursuing a masters in I.E. specializing in Human-Machine Systems Engineering. She is expected to graduate May, 1996 upon which she will be stationed at Brooks Air Force Base, TX working in the Instructional Systems Branch.

accomplish with the aircraft." This element suggests workload increases with the difficulty of the task and the number of tasks being performed. The second element is "the conditions or circumstances under which the required operation is to be conducted." This element suggests workload increases with "the adversiveness of the conditions under which the task(s) are being performed." These conditions include degraded state of the pilot, such as inadequate training or fatigue, and hostile environments, such as extreme temperatures or high g. The ability of the pilot to perform tasks is affected by both elements (Harper and Cooper, 1984).

Yet, other studies suggest workload has only one main element. Sarno and Wickens (1992) state that the main contributor to pilot workload is task loading. This is the number of tasks required to be performed at a given instant in time. In system development, the tasks and functions for which the operator will be responsible for are defined. Decisions involving the definition of the tasks and functions are referred to as function allocation (North et al., 1982 and Kantowitz and Sorkin, 1987). According to Beevis (1989), the task load imposed on the operator is defined by the function allocation decisions. These decisions "require verification in terms of the operator's ability to perform the assigned tasks." Workload in this context refers to the load due to the combination of the characteristics of the system and the task demands.

Compensatory Tracking Task

One of the assumed requirement of the pilot in an automated cockpit environment is that the pilot should have the capability to compensate for the automated system failure. The goal of compensatory tracking model is typically to minimize the time-averaged delays between apparent task execution failures. For the most part, the pilot's perceptual and control strategies are focused primarily on how to compensate for this failure (Levison, 1979).

Many researchers (see e.g. Jex, McDonnel, and Phatak, 1966) have shown the performance of compensatory tracking is affected by the system complexity. A system complexity is defined by the characteristic function s by: $k(s = 1)$ for zero order, k/s for first order, k/s^2 for second order and etc. The consensus based on previous studies show:

(a) The system may or may not be stationary;
(b) Higher harmonics or oscillations are always present (e.g. the movement of the cursor);
(c) The operator responds only to the current state of the system.

The workload study by Levinson (1979) clearly indicates that, in compensatory control tasks, the ratio of the RMS(path) to RMS(v) can be used to define workload. The RMS(path) is the root means square of path error, and RMS(v) is the root means square for control velocity. Whether the nature of the systems stability defined by damping coefficients affect the workload index is another aspect of study yet to be investigated. This constitutes the premise of our experimental investigations.

A New Theory of Workload

The classic workload index based on time analysis is usually defined by

$$WL(t) = \frac{R^{*}TimeRequired}{TimeAvailable} = \frac{Rt}{T} \qquad (1)$$

where t is the time required to accomplish the task, T is the time available to finish the task, and R is a learning coefficient. From equation (1), there is no consideration for the system or task complexity. Our theory of workload considers the system complexity. We assume, according to Gheorghe's (1979) observations that as the system becomes complex, especially with respect to task demands on the human, there is a loss of human energy to the system environment, resulting in the ability to accomplish the desired task. Gheorghe (1979) notes that this energy loss is an exponential decreasing function of the systems complexity. This is defined by

$$\sigma = e^{-a} \qquad (2)$$

where σ is the energy loss and a is the complexity measure.

Hypothetically, assume that a system is known with a defined work content c. This may be a function of time available to accomplish a task. As the human operator energy is dissipated, the amount of work accomplished in terms of the energy level is

$$c - \sigma = c - e^{-a} \qquad (3)$$

Similarly, the system is gaining on energy level defined by the initial work content plus the energy loss by the human operator. That is

$$c + \sigma = c + e^{-a} \qquad (4)$$

If we consider the impact of system dynamics in equation (3), we can define a normalized workload index by

$$WL = \frac{c - be^{-a}}{c + e^{-a}} \qquad (5)$$

here, b is the sytem damping coefficient defined by

$$b = \begin{cases} > 1; & \text{overdamped system} \\ = 0; & \text{normal system with no oscilation} \\ = 1; & \text{critically damped system} \\ < 1; & \text{underdamped system.} \end{cases}$$

$0 \le WL \le 1$

By the above condition, we expect

$$c - be^{-a} \ge \quad \Rightarrow \quad c \ge be^{-a}$$

The complexity parameter a can be defined in various ways. For our compensatory task experiment,

$$a = \frac{RMS(patherror)}{RMS(v)} \qquad (6)$$

If c = 1 from equation (5), we have

$$WL = \frac{1 - be^{-a}}{1 + e^{-a}} \qquad (7)$$

Sample Experiment

In the simulation experiment, the subjects perform tasks using the Manual Control Laboratory (MCL). The MCL produces multivariate statistics on human and system behaviors. For each experiment, the values of a are computed, and the WL determined in equation (7). Figure 1 shows sample results for b = 0, b = 1, and b = 2. As shown, WL increases as a increases. Especially as the damping coefficient (a) increases, the system complexity (b) increases.

Conclusion

The success of a mission, either in a stressful or peaceful environment, depends on the interaction between the pilot and the aircraft. Focusing on the pilot, their functions involve problem-solving, planning, and decision-making. They must be able to perform all tasks required to operate the aircraft while handling changes in mission objectives, weather, threats, and malfunctions. Factors effecting pilot performance under such condition include personal characteristics such as experiences, capabilities, biases, and skills. This combined with mission demand, which includes goals, environment, and aircraft design, will determine the workload. The physical and psychological state of the pilot will also contribute to the levels of workload. We hope to extend this study to address the effects of motion during task performance.

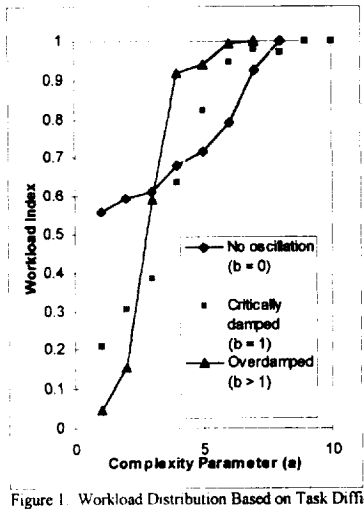


Figure 1. Workload Distribution Based on Task Difficulty

Bibliography

1. Beevis, D. (1989). Introduction: Task allocation and workload analysis models. In G.R. McMilliam, D. Beevis, E. Salas, M.H. Strub, R. Sutton, and L. VanBreda (Ed.) Applications of Human Performance Models to Systems Design, 2. New York: Plenum Press, 17-20.
2. Gawron, V.J., Schiflett, S.G., and Miller, J.C. (1989). Measures of in-flight workload. In R.S. Jensen (Ed.) Aviation Psychology. Aldershot, England: Gower Technical, 240-287.
3. Gheorghe, A. (1977). Applied Systems Engineering. New York: John Wiley & Son.
4. Harper, R.P. and Cooper, G.E. (1984). Handling qualities and pilot evaluations. Wright Brothers Lectureship in Aeronautics.
5. Hughes, D. (1984, August). Glass cockpit study reveals human factors problems. Aviation Week & Space Technology, 32-34.
6. Hughes, D. and Dornheim, M.A. (1995, January). Automated cockpits special report, Part I: Accidents direct focus on cockpit automation. Aviation Week & Space Technology, 52-54.

7. Jex, J.R., McDonnel, J.D., and Phatak, A.V. (1966). A "critical" tracking task for manual control research. IEEE Transaction on Human Factors in Electronics, 7(4), 138-145.
8. Kantowitz, B.H. and Sorkin, R.D. (1978). Allocation of functions. In G.Salvendy (Ed.) Handbook of Human Factors. New York: John Wiley and Sons.
9. Kernstock, N.C. (1989, February). Notar reduces pilot workload, improves responses in OH-6A. Aviation Week & Space Technology, 44-47.
10. Lysaght, R.J., Hill, S.G., Dick, A.O., Plamondon, B.D., Wherry, R.J., Zaklad, A.L., and Bittner, A.C. (1988, September). Operator Workload: Comprehensive Review & Evaluation of Operator Workload Methodologies. Technical Report 2075-3.
11. McCloy, T., Derrick, W., and Wickens, C. (1983, November). Workload assessment metrics - What happens when they dissociate? In Second Aerospace Behavioral Engineering Technology Conference Proceedings. Warrendale, PA: Society of Automotive Engineers, 37-42.
12. North, K., Stapleton, C., and Vogt, C. (1982). Ergonomic glossary: Terms commonly used in ergonomics. Bohn, Scheltema, Holkema, Ultrecht/Antwerp.
13. Sanders, M.S. and McCormick, E.J. (1993). Human Factors in Engineering and Design (7th ed.). New York: McGraw-Hill, Inc.
14. Sarno, K.J. and Wickens, C.D. (1992). Predictive workload models and multiple-task performance. In Proceedings of the Human Factors Society 36th Annual Meeting, 1, 12-16.
15. Wickens, C., Stokes, A., Barnett, B., and Hyman, F. (1989, February). The effects of stress on pilot judgement in a MIDIS simulator. Final Report for Period December 1987 - August 1988. Wright-Patterson AFB, OH: Armstrong Aerospace Medical Research Laboratory.

A Mechanism for Crossover Control in Genetic Algorithms

Ann Cheri Foxx[1], and Abdollah Homaifar
Autonomous Control Engineering, North Carolina A&T State University, Greensboro,NC

Keywords: Genetic Algorithms, Crossover Control, Metabits, Introns,Artificial Evolution, Endogenous Crossover

Abstract: This paper explores the use of a crossover control mechanism in genetic algorithms called metabit endogenous crossover control. This paper also shows how metabit crossover produces more highly fit individuals more often and more rapidly.

I. INTRODUCTION

It has been said that art imitates life. The same can be said about genetic algorithms (GA). Genetic algorithms are search algorithms based on the mechanics of natural genetics [1]. Simple genetic algorithms often perform three main operations: reproduction, crossover, and mutation. Reproduction is the process of creating a mating pool of copied individuals by some selection method. One selection method is the biased roulette wheel that has slots whose sizes are proportional to the fitness values of the individuals in the population. The higher the fitness of the individual, the larger its slot on the roulette wheel, and the greater its chance of being selected. This is a natural extension of the survival-of-the-fittest principle, which implies that those individuals with the highest fitnesses possess characteristics that are more likely to survive. The second operator, crossover, is the copying of genetic material from the parent to the child. Once again there are different methods of crossover. Finally, the mutation operator is the spontaneous change of an allele in a string. Mutation occurs with a small probability. The focus of this paper will be the crossover operator.

Crossover aims to create new structures with a minimum of disruption. This leads to exponentially increasing (or decreasing) proportions of schemata [1]. Crossover always occurs with some probability since artificial evolution is random in nature. Furthermore, there are several types of crossover. One method is one-point crossover in which a common point along the two parent strings is randomly selected and genetic material after that point is transposed.

An adaptation of this type of crossover is used in this work. The following is an example of one-point crossover:

Parent 1: **1001^100**
Parent 2: **0101^001**
Child 1: **1001001**
Child 2: **0101100**

A new generation is created by reproduction, crossover, and mutation and the old generation is replaced so that the population size remains constant. The new generation consists of the more highly fit individuals created by crossover. A fitness function is used to determine the fitness of an individual.

As a GA continues the reproduction and crossover process, highly fit individuals become more and more abundant. A problem that may arise as a result of this is premature convergence. This implies that there is one dominant individual. This individual and its offspring may eliminate desireable alleles which were developed in other individuals. Premature convergence also implies that because of this dominant individual, other individuals are never chosen for the mating pool. This would create a population with little diversity. Diversity often provides the key for finding a truly optimum string since some alleles found in less successful strings may be desireable. In order to tackle this dilemma, many strategies have been suggested. For example, adaptive selection methods are one way to overcome rapid convergence. Adaptive selection methods are means of limiting the participation of an individual in the mating pool. One such selection method, called ranking, causes each individual to receive an expected number of offspring according to the rank of its performance compared to other individuals and not its magnitude. Another method used to preserve diversity is the use of introns. In biological terms, an intron is a location within a chromosome which does not affect the chromosome's function, but does facilitate crossover. The insertion of introns increases the number of crossover points that preserve genes [2]. However, all locations are not equiprobable for crossing in a chromosome; various preferred locations in particular segments of chromosomes, called "hot spots", have been identified [3].

[1] Ann Cheri Foxx is a research assistant in Autonomous Control Engineering at North Carolina A&T State University. She is an undergraduate Electrical Engineering major who will graduate on May 11, 1996. She has research interests in genetic algorithms and control engineering.

These areas of high crossover probabilities cause a phenomena to occur called punctuated crossover. Holland was the first to conceive the notion of punctuated crossover, but he proposed it for purposes other than the one presented [4]. Punctuated crossover as proposed by Schaffer can be used as a method of crossover control. Punctuated crossover is similar to traditional crossover in that points along the parent strings are chosen between which crossover is to occur. However, parents need not have the same crossover points for punctuated crossover to occur. The following is an example of punctuate crossover.

BEFORE CROSSOVER

parent 1: a a a a a a a!b b b b b bb
parent 2: c c c c!d d d d d d!e e e e

AFTER CROSSOVER

child 1: a a a a d d d b b b e e e e
child 2: c c c c!a a a!d d d! b b b b

The advantage of punctuated crossover is that highly fit parts of a string are most likely preserved in addition to the crossover site markings for those "functional" parts of the string. If an offspring does not survive, then the punctuations that led to its production are also destroyed. This mechanism gives the GA a method for locating ideal places to perform crossover.

The punctuations are caused by the aforementioned hotspots. It is proposed that these hotspots do not occur within strings at a constant probability. In fact, each possible site for crossover has a randomly selected crossover probability of its own. These crossover probabilities are marked by nonfunctional mechanisms that do not affect the fitness or function of the string. The fact that "hot spots" occur at different crossover probabilities implies that the mechanisms represent different crossover probabilities. That is to say, some regions of crossover are more likely to produce more highly fit individuals and are, therefore, more likely to survive.

Since this occurs within the string, it is referred to as endogenous which literally means proceeding from within. The fact that the crossover probability at any two given points may be different leads to a so called differential crossover probability. The nonfunctional mechanism which facilitates this endogenous differential crossover is known as metabits[5]. Therefore, a metabit is a mechanism located within the chromosome which may represent different probabilities for crossover.

Metabits are located before the alleles of a chromosome. For example, a chromosome eeeeee and metabits (m) becomes mememememe [5]. Crossover must occur between metabit-allele (me) pairs. If the metabit is 1 at a crossover point within both of the two parent binary strings, then the standard crossover probability will be used at that crossover point. If either of the parent strings contain a metabit that is 0 at that crossover point, then a reduced probability will be used. Because the chances of having two metabits of 1 is less than the other possible combinations (i.e., 1 and 0, 0 and 0, or 0 and 1), a metabit of 1 is considered recessive and a metabit of 0 is considered dominant.

It is proposed by [5] that metabit crossover improves the performance of GAs searching for strings composed of multiple building blocks. Note that a building block is low-ordered with above-average fitness. Metabits improve the performance of strings containing these building blocks by (1) reducing the schema disruption by reducing crossover probabilties between the desired genes, and (2) generating highly fit strings faster and more frequently by causing punctuated crossover [5].

The remainder of this paper details the construction of a GA within which metabit crossover is implemented. It also presents the results of comparing a GA with endogenous crossover control to one that does not have endogenous crossover control. The conclusions are promising and provide some insight into future applications of this mechanism.

II. SIMULATION

The GA being presented is one which includes endogenous crossover control. First, the initial population is selected by flipping an unbiased coin, using a random number generator, to create binary strings. The user defines the population size, string length, crossover probability, metacrossover probability, mutation probability, and number of generations to be iterated. Second, one-point crossover is used to exchange genetic information between parents. Mutation occurs randomly and with a very small probability. The binary string is then decoded and evaluated for its fitness. Finally, the new population is produced by roulette wheel selection.

In order to initialize the random number generator for the selection of the initial population, a seed number is entered by the user. The random number generator then fetches a number (one or zero) which is placed in a string of some specified length. A string that contains metabits is actually two times the length of a string of equal magnitude that does not contain metabits. A metabit is inserted before each allele of a chromosome string. This must be taken into account when entering the desired string length.

One-point crossover is performed by first randomly selecting a crossover site within the parent strings and

then by crossing over. Mutation also occurs randomly during the operation of crossover with a specified probability. In order to determine the probability of crossover that will be used, standard crossover probability or metacrossover probability, a crossover site is chosen randomly by the random number generator. The metabits located within each parent string at that crossover site are tested. If both of the metabits are one, then the standard crossover probability (usually the inverse of the string length) is used. If either parent has a zero metabit at that crossover site, then the metacrossover probability is used. The metacrossover probability is substantially smaller than the standard crossover probability. Metacrossover probabilities of at least one hundred times smaller than the standard crossover probabilities have shown the most significant results as far as average fitness is concerned. The advantage of the differential crossover probability is that portions of strings that help to produce highly fit individuals are more likely to survive. Furthermore, these highly fit portions are less likely to be disrupted since the crossover probability is greater before or after the portion than within it. Thus, far less disruption of highly fit schema occurs. It should also be reiterated that crossover must occur after a metabit-allele pair and not in between since the metabit is a nonfunctional part of the string that marks a good or bad location for crossover.

After crossover and mutation, the offspring are decoded. The decoding process occurs within two steps for strings that include metabits. First, the metabits are removed from the string so that only the functional (contributing) alleles remain. Second, the binary string is converted to decimal. The decoded string then undergoes evaluation. For this GA the fitness function is:

(decimal value of string+ (2^(string length)-1))^10

The final phase of the GA is roulette wheel selection. In this technique, those individuals with a higher fitness will have a greater probability of being selected. The number of individuals selected for the population remains constant so that the population size remains the same.

III. RESULTS

The results of the experiment are favorable. In each case, the genetic algorithm with metabit crossover found more highly fit individuals more rapidly than the genetic algorithm without the endogenous crossover control. However, the gains from the use of metabits were not as significant as was hoped. Consider Figure 1. It is obvious that the genetic alogorithm with endogenous crossover control (solid line) outperformed the genetic algorithm without endogenous crossover control (dotted line); though, the gains were not extremely significant. Figure 2 shows how much better the genetic algorithm with endogenous crossover control was capable of operating comparatively.

Tables 1 and 2 show a comparison of a GA with and without endogenous crossover control. They describe the parameters of the GA and compare fitness values over a certain number of generations. This sample is representative of the many trials run.

Table 1: Genetic Algorithm

Population Size	100
Chromosome Length	75
Maximum Number of Generations	100
Crossover Probability	0.70
Metacrossover Probability	0.0003
Mutation Probability	0.0001

Table 2: Average Fitness Comparison

Gener-ation	Meta-GA Fitness	Control GA Fitness
1	0.905537	0.894908
2	0.945017	0.938747
3	0.968870	0.946905
4	0.976589	0.955593
05	0.975144	0.948098
10	0.981224	0.970586
15	0.982037	0.968735
20	0.982912	0.992237
25	0.996931	0.993295
30	0.999112	0.999590
35	0.999969	0.999677
40	0.995720	0.999608
45	0.999981	0.999758
50	0.999986	0.999783
60	0.999979	0.999897
70	0.999986	0.999795
80	0.996564	0.999738
90	0.999991	0.999800
100	0.999988	0.999734

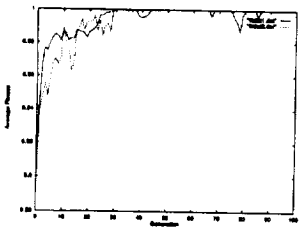


Figure 1: Average Fitness vs. Generation.

Population Size: 100, String Length: 50, Number of Generations: 100

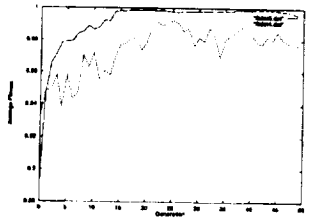


Figure 2: Average Fitness vs. Generation.

Population Size: 100, String Length: 75, Number of Generations: 100

The difference between the two GAs is that the GA without crossover control tends to converge to a solution quickly. It often wipes out less highly fit individuals in favor of more fit individuals. This decreases diversity to a great extent. The GA with endogenous crossover control is able to produce more highly fit individuals faster because it allows the desireable characteristics of the less fit individuals to presist while discarding the undesireable characteristics.

IV. CONCLUSION

This study has shown that the genetic algorithm with endogenous crossover control produces more highly fit individuals more rapidly than the genetic algorithm without endogenous crossover control. Metabits tend to create a type of punctuated crossover in which ideal crossover sites are chosen along with the genetic material that is being passed on from parent to child. Furthermore, highly fit schemata are allowed to remain in the population without dominating it [1]. This is because the differing crossover probabilities that a schema may encounter curbs the domination of a particularly highly fit schema. Thus, the endogenous crossover control mechanism serves as a means for protecting diversity while allowing the genetic algorithm to produce more highly fit strings.

The results shown in this paper may have been more conclusive had the overall GA been less convergent in nature. Metabit crossover works best in non-converging GA's because these types of GA's allow for continuous diversification. Continually diversifying the population is key in metabit crossover. Since one-point crossover was used and a simple genetic algorithm process of roulette-wheel selection, mutation, and crossover were chosen, the results were less than monumental. However, the gains achieved by metabit crossover are in no way negligible since it proves that this type of crossover does indeed work.

Applications for this modified approach to crossover are numerous. The most significant of these applications would be to control the overall convergence of a GA which would operate most efficiently if convergence is relatively slow. The procedure for implementing it into most simple genetic algorithms is not overwhelmingly difficult, though adjustment to decoding strings and evaluating fitness functions must be dealt with. Therefore, there are many possible uses for this generic crossover control mechanism.

ACKNOWLEDGMENT

This work is partially funded by grants from NASA-Autonomous Control Engineering Center under grant number ACE-48146. The authors wish to thank them for their financial support.

REFERENCES

[1] Goldberg, D.,(1987). *Genetic Algorithms in Search, Optimization, and Machine Learning.* Reading, MA: Addison-Wesley.

[2] Levenick, S.,(1991). Inserting Introns Improves Genetic Algorithm Success Rate: Taking a Cue from Biology. *Proceedings of the Fourth International Conference on Genetic Algorithms,* 123-127. San Mateo, CA: Morgan Kaufmann Publishers.

[3] Grimer, T.,(1989). Hot Spot Recombination within DXS164 in the Duchenne Muscular Dystrophy Gene. *American Journal of Human Genetics,* 45-47. University of Chicago Press.

[4] Holland, J., (1975). *Adaptation in Natural and Artificial Systems.* University of Michigan Press, Ann Arbor. Washington, National Academy of Sciences.

[5] Levenick, J.(1995). Metabits: Generic endogenous crossover control. *Proceedings of the Sixth International Conference on Genetic Algorithms,* 88-95. San Mateo, CA: Morgan Kaufmann Publishers, Inc.

Construction and Modeling of a Controls Testbed

James C. Nagle, Abdollah Homaifar, Ahmed A. Nasser, Marwan Bikdash
nagle@ncat.edu, homaifar@ncat.edu, bikdash@ncat.edu, ahmada@ncat.edu

ABSTRACT

This paper describes the construction and modeling of a control system testbed to be used for the comparison of various control methodologies. We specifically wish to test fuzzy logic control and conpare performance of various fuzzy controllers, including Hybrid Fuzzy-PID (HFPID) and Hierarchical Hybrid Fuzzy-PID (HHFPID) to other controllers including localized rate feedback, LQR/LTR, and H_2/H_∞. The control problem is that of vibration suppression in a thin plate with inputs coming from accelerometers and outputs going to piezoelectric actuators or "patches". A model based on experimental modal analysis of the plate is conducted and compared with an analytical model. The analytical model uses a boundary condition which is a mix of clamped and simply supported.

KEYWORDS: Smart Structure, fuzzy logic, vibration, control, modeling

1 Introduction

Over the past several years the use of distributed sensors and actuators to control vibration in flexible structures has emerged as a viable alternative to passive techniques[1-4]. Such structures are often labelled "Smart Structures" or "Active Structures". The increasing use of active vibration suppression is primarily due to the developement of materials capable of converting mechanical deformations into voltages or converting voltages into mechanical deformations. Perhaps the most ubiquitous of these materials is the piezoceramic. The use of such materials makes possible a lightweight, low volume, active alternative to passive vibration damping technology. Applications of these active techniques are expanding daily, and include vibration suppression in launch vehicles, aircraft, and automobiles[1-4]. The application of distributed sensors and actuators presents new challenges to control system designers. Several design methodologies, including H_2/H_∞ have been developed for or applied to this problem cope with these new demands. The application of these methodologies to simple geometries (plate, beam, etc.) is fairly well documented and allows a benchmark for comparison of new control designs, specifically Hybrid Fuzzy-PID and Hierarchical Hybrid Fuzzy-PID, to existing strategies. These new controllers have been tested extensively in simulation and have demonstrated excellent performance and robustness characteristics [5,6]. This paper reports on the work-in-progress and describes the construction of a simple plate structure with distributed sensors and actuators, the modeling of this structure, and the application of various controllers to this structure.

2 Construction

The rectangular plate is .050" thick aluminum with edges of 28" and 21". The frame of the structure clamps the plate between two 1" square aluminum members. The frame is supported at the corners with legs of angle aluminum which are anchored to a benchtop, preventing any rigid body motion. While constructing the plate/frame assembly it was noted that a buckle was present in the plate. A stretching jig was fabricated and the plate stretched until the buckle was not present after releasing tension from the jig.

Control input to the system is accomplished through the use of piezoelectric actuators. Each actuator is composed of two rectangular piezoelectric elements above and below the plate with the poling direction for each element oriented in the same direction. Each piezoelectric element was attached using a cyanoacrylate adhesive. The plate became the negative terminal connection with the positive terminal being the outer side of each of the two elements. When a voltage is applied to the patches one will attempt to expand along its major axis while the other will contract, producing a bending moment on the plate (see Fig.1).

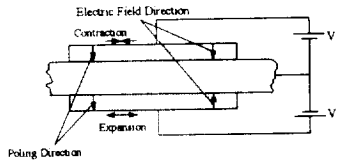


Figure 1: Relation between Applied Electric Field, Poling direction, and mechanical motion

A total of five actuators are used in this system. Sensing is accomplished through the use of accelerometers colocated with each actuator. Small lightweight accelerometers were chosen so as to minimize any structure/sensor interaction.

Locating the sensors and actuators is a significant problem in itself. Because the objective of this work is not to design the structure itself but rather to compare controller performance, it was decided to roughly place the sensors/actuators to maximize controllability for the first three modes (see Fig. 2).

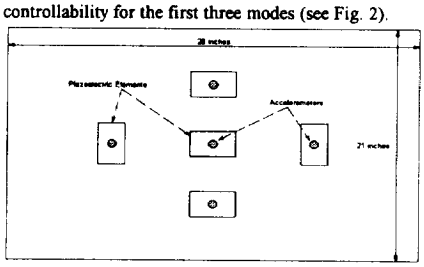


Figure 2: Sensor/Actuator placement on plate

The piezoceramic elements are Model T107-A2-501 sheets from Piezo Systems. The accelerometers are PCB Piezotronics model 352B22.

Figure 3 illustrates the signal flow for the closed loop control system using rate feedback.

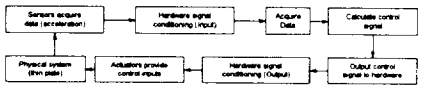


Figure 3: Signal Flow in control loop

The accelerometers are connected to a power supply/conditioner (PCB Piezotronics Model 483B07). The signal then flows to a custom-built lowpass filter bank (antialiasing filters). These filters are eighth-order Butterworth lowpass switched-capacitor filters. The switched capacitor design allows for a variable cutoff frequency to be implemented using a single potentiometer. The filters are self clocked using the optional Schmitt trigger configuration. The signal is then acquired by the National Instruments AT-MIO-16F5 board. All control algorithms are realized through National Instruments LabView environment. After computing the control signal, it is output through a National Instruments AT-AO-6 board (again 12 bits) and on to a smoothing filter that is identical to the antialiasing filter described earlier. The signal finally passes through a PCB Piezotronics 790 series power amplifier before arriving at the actuators.

3 Modeling

The structure was modeled from experimental data as well as analytically from first principles. The complete derivation of the equations of motion of the plate are too lengthy to be derived in this paper, but we present here an synopsis of the steps. It can be shown that the equation of motion of the plate is described by the following equation:

$$D\left\{\frac{\partial^4 w}{\partial x^4}+2\frac{\partial^4 w}{\partial x^2\partial y^2}+\frac{\partial^4 w}{\partial y^4}+c\alpha\psi\left[\frac{\partial^2}{\partial x^2}\left(\chi\left(\frac{\partial^2 w}{\partial x^2}+\nu\frac{\partial^2 w}{\partial y^2}\right)\right)\right.\right.$$

$$\left.\left.+2(1-\nu)\frac{\partial^2}{\partial x\partial y}\left(\chi\frac{\partial^2 w}{\partial x\partial y}\right)+\frac{\partial^2}{\partial y^2}\left(\chi\left(\frac{\partial^2 w}{\partial y^2}+\nu\frac{\partial^2 w}{\partial x^2}\right)\right)\right]\right\}+\rho_s h_s\ddot{w}$$

$$=\sum_{q=1}^{N}F_q+\frac{Y_{pz}}{1-\nu}\sum_{p=1}^{M}h_t d_{31}\left(\frac{\partial^2}{\partial x^2}+\nu\frac{\partial^2}{\partial y^2}\right)\left(\chi_p V_p\right)$$

Equation 1

where h_t is the total thickness of the plate and the patches at that location, $\psi = Y_{pz}/Y_p$, and V_p is the voltage applied to the p^{th} patch. Let the solution of the free vibration problem ($V_p = 0$ and $F_q = 0$) be of the form: $w(x,y,t) = W(x,y)e^{j\alpha t}$

Equation 1 then can be written as:

$$D\left\{\frac{\partial^4 W}{\partial x^4}+2\frac{\partial^4 W}{\partial x^2\partial y^2}+\frac{\partial^4 W}{\partial y^4}+c\alpha\psi\left[\frac{\partial^2}{\partial x^2}\left(\chi\left(\frac{\partial^2 W}{\partial x^2}+\nu\frac{\partial^2 W}{\partial y^2}\right)\right)\right.\right.$$

$$\left.\left.+2(1-\nu)\frac{\partial^2}{\partial x\partial y}\left(\chi\frac{\partial^2 W}{\partial x\partial y}\right)+\frac{\partial^2}{\partial y^2}\left(\chi\left(\frac{\partial^2 W}{\partial y^2}+\nu\frac{\partial^2 W}{\partial x^2}\right)\right)\right]\right\}-\rho_s h_s\alpha^2 W=$$

Equation 2

$W(x,y)$ is assumed to be separable in the x and y directions. That is: $W(x,y) = X(x)Y(y)$

Assuming that the plate is homogeneous such that the elastic effect of the piezo patches is distributed over the entire plate, an approximate solution for the eigenvalues and eigenfunctions of equation (1) with the appropriate boundary conditions are given by (see Elishakoff [7]):

$$\omega_{rsn} = \left(\frac{D}{\rho}\right)^{1/2}\left[\left(\frac{\bar{r}\pi}{a}\right)^2 + \left(\frac{\bar{s}\pi}{b}\right)^2\right]$$

$$X_{rn}(x) = \cosh\frac{\theta\pi x}{a} - \cos\frac{\bar{r}\pi x}{a} - \frac{C_1 - c_1}{S_1 - s_1\theta/\bar{r}}\left(\frac{\theta\pi x}{a} - \frac{\theta}{\bar{r}}\sin\frac{\bar{r}\pi x}{a}\right)$$

$$Y_{rn}(y) = \cosh\frac{\alpha\pi y}{b} - \cos\frac{\bar{s}\pi y}{b} - \frac{C_2 - c_2}{S_2 - s_2\alpha/\bar{s}}\left(\frac{\alpha\pi x}{b} - \frac{\alpha}{\bar{s}}\sin\frac{\bar{s}\pi y}{b}\right)$$

$$\Psi_{rs}(x,y) = X_{rs}(x)Y_{rs}(y)$$

Equation 3

where

$c_1 = \cos\bar{r}\pi$ $\quad s_1 = \sin\bar{r}\pi$ $\quad C_1 = \cosh\theta\pi$ $\quad S_1 = \sinh\theta\pi$

$c_2 = \cos\bar{s}\pi$ $\quad s_2 = \sin\bar{s}\pi$ $\quad C_2 = \cosh\alpha\pi$ $\quad S_2 = \sinh\alpha\pi$

$\theta = \sqrt{\frac{2\bar{s}^2}{\delta^2} + \bar{r}^2}$ $\quad \alpha = \sqrt{\frac{2\bar{r}^2}{\delta^2} + \bar{s}^2}$ $\quad \delta = \frac{a}{b}$

Equation 4

For a plate clamped all-around, $\bar{r}$ and $\bar{s}$ which correspond to mode r in the x-direction and mode s in the y-direction are the roots of the following pair of transcendental equations:

$$1 - c_1 C_1 + \frac{\bar{s}\delta}{\bar{r}}\left(2 + \frac{\bar{r}^2}{\bar{s}^2\delta^2}\right)^{-1/2} s_1 S_1 = 0$$

$$1 - c_2 C_2 + \frac{\bar{r}}{\bar{s}\delta}\left(2 + \frac{\bar{s}^2\delta^2}{\bar{r}^2}\right)^{1/2} s_2 S_2 = 0$$

Equation 5

For a simply-supported plate, the values of $\bar{r}$ and $\bar{s}$ are simply r and s respectively. After applying orthogonality conditions and expressing the soultion to equation (1) as: $w(x,y,t) = \sum_{j=1}^{\infty}\Psi_j(x,y)\eta_j(t)$ we obtain the following infinite dimensional system of equations:

$$\ddot{\eta}_j + 2\varsigma_j\varpi_j\eta_j + \omega_j^2\eta_j = \sum_{q=1}^{N}F_{jq} + \sum_{p=1}^{M}Q_{jp}V_p$$

where

$$F_{jq} = \int_{\Omega}\Psi_j F_q \,dx\,dy$$

$$Q_{jp} = \frac{Y_{pc}h_t d_{31}}{1-\nu}\int_{\Omega}\left(\frac{\partial^2\Psi_k}{\partial x^2} + \nu\frac{\partial^2\Psi_k}{\partial y^2}\right)\chi_p\,dx\,dy$$

Equation 6

and a modal damping term has been added to account for energy dissipation.

Experimental modal analysis:

The values of ω_j and ς_j are determined experimentally using an impulse hammer modal test. A Kistler 9722A200 instrumented hammer with a 2 PCB 353B65 shear-mode quartz accelerometers were used. A PC-based Data Physics DP420 analyzer was used for data acquisition and signal processing to generate frequency response functions(FRFs). Two sets of FRFs were obtained after taking an average of 10 measurements for each position on the grid. These FRFs were then exported to STAR Modal to obtain the Modal parameters. The advanced curve fitting technique of STAR-Modal was used to calculate the natural frequency, damping ratios and mode shapes. The results of the experimental modal analysis are then used to fit the mathematical model derived earlier. For the experimental study we used a $28"\times21"\times0.05"$ aluminum plate clamped all around. A mesh of 11×11 nodes was used as shown in Figure 2. (1.a)

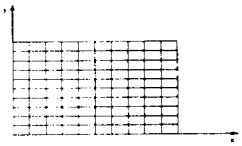


Figure 4: Plate geometry and coordinate system

It was noticed that the theoretical natural frequencies were not a close match to those obtained experimentally. The natural frequencies obtained experimentally were lower than those predicted by theory. This lead to an investigation of the boundary conditions. The narrow clamp made of two aluminum beams with $1"\times1"$ cross-section and constrained on four legs of the same cross-section cannot in reality produce perfectly clamped boundaries. We did not attempt to change the boundary conditions because in a real life situation the boundary conditions on a panel are neither perfectlyly clamped nor perfectly simply-

supported. Instead we modeled the boundary conditions as a combination of the clamed and simply-supported boundary conditions using a weighting function for each mode to account for the contribution of the constrained rotations on the boundaries. A MATLAB program was written to solve the transcendental equation 5 for $\bar{r}$ and $\bar{s}$ using the built-in functions of MATLAB's Optimization Toolbox. Equation (4) is then modified by substituting the weighted values (r', s') for $(\bar{r}, \bar{s})$ and (r, s) as follows:

$$(r', s') = (f_x \bar{r}, f_y \bar{s}) + ((1 - f_x)r, (1 - f_y)s)$$

The constant values of f_x and f_y are found by matching the natural frequencies of modes 1,2 and 2,1 from experiment and theory. Equation (4) is rewritten as:

$$X_n(x) = \cosh\frac{\theta' \pi x}{a} - \cos\frac{r' \pi x}{a} - \frac{C_1 - c_1}{S_1 - s_1 \theta / r'}\left(\frac{\theta \pi x}{a} - \frac{\theta}{r'}\sin\frac{r' \pi x}{a}\right)$$

$$Y_n(y) = \cosh\frac{\alpha' \pi y}{b} - \cos\frac{s' \pi y}{b} - \frac{C_2 - c_2}{S_2 - s_2 \alpha / s'}\left(\frac{\alpha' \pi y}{b} - \frac{\alpha'}{s'}\sin\frac{s' \pi y}{b}\right)$$

Equation 7

The values of the parameters in equation (7) are calculated using equation (5) by replacing $(\bar{r}, \bar{s})$ by (r', s'). The following table includes a comparison between the experimental and theoretical values of the natural frequencies.

Table. 1: Comparison of the experimental and theoretical natural frequencies of the plate

	Natural Frequencies		
Mode #	Experimental	Theoretical	% Diff.
1(1,1)	26.73	26.73	1.982
2(2,1)	47.11	47.11	1.995
3(1,2)	70.37	70.73	2.003
4(3,1)	81.92	82.15	2.296
5(2,2)	101.23	89.96	-9.349
6(4,1)	129.91	131.2	3.017
7(3,2)	137.37	123.7	-8.129
8(1,3)	140.15	138.0	0.439

Although the natural frequencies from the mathematical model are calculated, the experimental values of these frequencies are used in the modal equations. The mathematical model with the wieghting functions is used to find a good smooth estimation for the mode shapes. When constructing the modal equations these mode shapes are easier to differentiate and integrate than the experimental mode shapes because of the limited number of points in the experimental model.

4 Control

We are currently at the stage of implementing the various control algorithms. Localized rate feedback, (i.e. $u = -k * \frac{dx}{dt}$) will be implemented first, with an H_2/H_∞ and Hybrid Fuzzy-PID Controller following. Ultimately we will be comparing the performance of each controller when faced with model parameter uncertainty.

Acknowledgements:

We would like to thank Dr. Frank Pai, Structures Group coordinator, and the Structures Group of NASA Core at North Carolina A&T State University0 for their assistance and involvement. This work is partially funded by grants from the NASA Center of Research Excellence at North Carolina A&T State University under grant number NAGW-2924. The authors wish to thank them.

References

[1] Collins, S.A., Padilla, C.E., Notestine, R.J., von Flotow, A.H., Schmitz, E., and Ramey, M., "Design, Manufacture, and Application to Space Robotics of Distributed Piezoelectric Sensors", *J. Guid. Cont. & Dyn.*, Vol. 15, No. 2, 1992, pp. 396-403

[2] Crawley, E.F., and de Luis, J., "Use of Piezoelectric Actuators as Elements of Intelligent Structures", *AIAA Journal*, Vol. 25, No. 10, 1987, pp. 1373-1385

[3] Crawley, E.F., and de Luis, J., " Experimental Verification of Distributed Piezoelectric Actuators for Use in Precision Space Structures", *Proc. 27th AIAA Struct. Dyn. Conf.*, 1986, pp. 116-124

[4] Hanagud, S., Obal, M.W., Calise, A.J., "Optimal Vibration Control by the Use of Piezoceramic Sensors and Actuators", *J. Guid. Cont. and Dyn.*, Vol. 15, No. 5, 1992, pp. 1199-1206

[5] Homaifar, A., Sayyarrodsari, B., and W.E. Snyder, "A Theoretical Justification for Nonlinear Controlling Property of a Class of Fuzzy Logic Controllers", *Third IEEE International Conference on Fuzzy Systems*, pp. 382-386, 1994

[6] A. Homaifar, B. Sayyarrodsari, J.E. Hogans, "Fuzzy Controller for Robot Arm Trajectory", *Journal of Information Sciences*, 1994

[7] Elishakoff, I., "Random Vibrations of Orthotropic Plates Clamped or Simply Supported All Round", *Acta Mechanica*, 1977, Vol. 28, pp. 165-176

560-63

29074 P4

DESIGN OF GENERALIZED SUGENO CONTROLLERS BY APPROXIMATING HYBRID FUZZY-PID CONTROLLERS

Charles Clifton
clifton@ncat.edu

Abdollah Homaifar
homaifar@ncat.edu

Marwan Bikdash
bikdash@ncat.edu

Department of Electrical Engineering
North Carolina A&T State University
Greensboro, NC 27411

ABSTRACT

This paper proposes a simple technique to design a generalized Sugeno controller (GSC). A hybrid fuzzy-PID (HFPID) controller is approximated using recursive least squares by a Sugeno-type controller. The method used to generate the approximating data is shown to affect the approximation. A comparison of the performance between a Sugeno controller and a hybrid fuzzy-PID controller when they are both applied to a two-degree-of-freedom robot manipulator arm is shown. Furthermore, the number of terms needed to produce accurate performance of the generalized Sugeno controller is discussed.

I. INTRODUCTION

Fuzzy logic controllers (FLCs) have been successfully implemented in the automation of control strategies for complex, ill-defined, uncertain control systems in which traditional quantitative approaches have often failed to properly model the system. In [Sayyarrodsari and Homaifar, 1997], the hybrid fuzzy-PID controller was proposed in which a fuzzy engine interprets the knowledge of a human expert in the "design" of a PID controller. This approach is more "linguistic" , and therefore closer to the spirit of fuzzy logic.

It was shown in [Buckley, 1993] that the Sugeno-type controller (or generalized Sugeno controller) has the universal approximation property meaning that it can approximate any control law mapping the measurements (or states) into control actions within an arbitrary accuracy. However, design procedures for the Sugeno-type controller are difficult. In this paper, we propose a procedure to design a GSC by approximating the hybrid fuzzy-PID controller. The HFPID has many desirable properties: (1) it extends the application of conventional PID controllers to a more sophisticated, nonlinear control problem, (2) it overcomes the performance deficiencies in PID controllers when parameter uncertainties occur in the system or when the extension of the operation range requires a compromise in PID coefficient settings to avoid undesirable performance in one area at the expense of lower control quality in another, (3) it has excellent tracking ability, (4) it has small overshoot, fast rise time, and minimal steady state error, and (5) it is quite robust to parameter uncertainties.

This paper is organized as follows. Section II provides the necessary background information on the HFPID and the GSC. Section III discusses the methodology used in designing the GSC from the data obtained from the HFPID. Section IV compares the performance of the generalized Sugeno controller as compared to that of a hybrid fuzzy-PID controller when they are both applied to a two degree of freedom robot arm. Finally, section V gives a brief summary and the conclusions.

II. BACKGROUND INFORMATION

II-a. Robot Manipulator Arm

We apply our methodology of approximating the hybrid fuzzy-PID controller by a Sugeno-type controller to a two-degree of freedom robot manipulator arm. The dynamic behavior of two arms is described by the following nonlinear ordinary differential equation

$$M(\Theta)\ddot{\Theta} + n(\Theta,\dot{\Theta},t) = \tau + f(\Theta,p,t) \qquad (1)$$

where M is the 2×2 inertia matrix, n is a 2×1 matrix representing the sum of the centrifugal, coriolis, and gravitational forces, τ is the control torque vector, f is the collection of all disturbances, and p is a parameter vector. The two arm positions θ_1 and θ_2 are assumed to be with accurate optimal encoders. This system is highly nonlinear, strongly coupled, but is a prototypical problem in literature because many nonlinear systems have similar dynamical equations. For further details of the system equations and a summary of the existing control schemes, please refer to [Sayyarrodsari and Homaifar, 1997]

II-b. The Hybrid Fuzzy-PID Controller

The hybrid fuzzy-PID controller was developed as a step toward a systematic procedure for design and

analysis of a class of fuzzy-logic controllers. In this controller, the fuzzy inference engine is placed above a conventional PID controller within the control hierarchy [Sayyarrodsari and Homaifar, 1997]. The system under control is the robot arm discussed earlier. The two arm positions θ_1 and θ_2 are measured and then, compared with desired positions θ_1^d and θ_2^d. The errors ($e_1=\theta_1-\theta_1^d$ and $e_2=\theta_2-\theta_2^d$) are obtained, and their derivatives and integrals are numerically estimated. The errors, along with its time derivatives and integrals are fed into the FLC to produce the proportional, derivative, and integral coefficients P_1, P_2, I_1, I_2, D_1, and D_2. These coefficients are fed into the conventional PID controller to produce the torques for the robot arm.

The fuzzy engine in a HFPID simulates the role of the designer of a conventional PID controller. It sets the values of the proportional, integral, and derivative coefficients. The HFPID has, therefore, two layers; the lower PID controller with variable coefficients and the higher fuzzy inference engine setting on-line the values of these coefficients. The required fuzzy control rules are linguistic, quite easy to design, and are of the from: << **If** *error* is *big* and *derivative_of_error* is *small* and *integral_of_error* is small **then** *proportional_coefficient* should be *big* and *derivative_coefficient* should be *small* and *integral_coefficient* should be *medium* >>. The fuzzy engine can easily implement the knowledge of a human expert in the design of a PID controller. The outputs of the fuzzy part of the HFPID are determined through a nonlinear time-invariant mapping performed by the fuzzy engine from the system's state space to corresponding proportional, integral, and derivative coefficients based on the conditions the system under control actually experiences.

II-c. Generalized Sugeno Controller

A generalized Sugeno controller [Buckley, 1993] is a fuzzy engine mapping $x\in\Re^n$ into $y\in\Re$ where x is interpreted as being a state or a measurement and y as a control action. The inference is of the form:

$$\boldsymbol{R}_i\text{: IF } x_1 \text{ is } A_1^i \text{ and } x_2 \text{ is } A_2^i \text{ and ... and } x_n \text{ is } A_n^i \text{ THEN } y = P_i(\alpha_1^i, \alpha_2^i, \ldots \alpha_n^i);\ 1\le i \le k \qquad (2)$$

where x_j is the jth component of the input vector x and is a crisp value, A_m^i specifies which among the fuzzy attributes of x_m tested by rule i, and P_i is a polynomial in $x_1, x_2, \ldots, x_n$ given by:

$$P_i = \sum_{a_1=0}^{\alpha_1^i} \cdots \sum_{a_n=0}^{\alpha_n^i} C_i(a_i,\ldots\alpha_n)(x_1)^{a_1}\cdots(x_n)^{a_n} \qquad (3)$$

where $C_i(a_1, a_2, \ldots, a_n)$ are real numbers and $\alpha_1^i, \alpha_2^i, \ldots\alpha_n^i$ are constant positive integers. In order to design a GSC, we need to determine the consequence coefficients, $C_i(a_1,\ldots,a_n)$, and the order of each input variable, $\alpha_1,\ldots, \alpha_n$, for the output polynomial P_i. Note that this Sugeno controller is a generalization of the original Sugeno controller (OSC) [Takagi and Sugeno, 1985] where the consequence, y, is a linear function of the input variables. The rules of the original Sugeno controller take on the following form:

$$\boldsymbol{R}_i\text{: IF } x_1 \text{ is } A_1^i \text{ and } x_2 \text{ is } A_2^i \text{ and ... and } x_n \text{ is } A_n^i \text{ THEN } y = c_0^i + c_1^i x_1 + \ldots + c_n^i x_n \qquad (4)$$

where $c_0^i, c_1^i, \ldots, c_n^i$ are the consequence coefficients of the ith fuzzy rule. For further discussion of Sugeno-type controllers see [Buckley, 1993]. Buckley proved that a Sugeno-type controller can approximate any continuous real valued function in the output space to any degree of accuracy if the following were true: (a) the input fuzzy sets have continuous membership functions and (b) a continuous T-norm is being used in the rule evaluation process. This is the universal approximating property of the Sugeno-type controller.

III. METHODOLOGY AND PROCEDURE

This section proposes a technique that will use the data generated from the hybrid fuzzy-PID controller to identify the coefficients of the Sugeno-type controller. Also, this section describes how the HFPID and the resulting Sugeno-type controller will be implemented in a two-degree of freedom robot manipulator arm.

III-a. Determining OSC Consequence Coefficients

First we consider approximating the hybrid fuzzy-PID controller by the original Sugeno controller. To do so, we need to determine the coefficients c_0^i, c_1^i, etc. In this paper, we use subscripts to index vectors and superscripts to identify components within a vector. The output y for the inputs $x_1,\ldots,x_n$ is obtained by the following equations:

$$y = \frac{\sum_{i=1}^{k}\left(A_1^i(x_1)\cdots A_n^i(x_n)\right)\left(c_0^i + c_1^i x_1 + \cdots + c_n^i x_n\right)}{\sum_{i=1}^{k}\left(A_1^i(x_1)\cdots A_n^i(x_n)\right)} \qquad (5)$$

Let $\beta_i(x)$ be the truth value of the antecedant of the ith rule evaluated with input x or

$$\beta_i(x) = \frac{A_1^i(x_1)\cdots A_n^i(x_n)}{\sum_{i=1}^{k}\left(A_1^i(x_1)\cdots A_n^i(x_n)\right)} \qquad (6)$$

As a result, y can be simplified to:

$$\hat{y} = \sum_{i=1}^{k}\beta_i c_0^i + x_1\sum_{i=1}^{k}\beta_i c_1^i + \cdots + x_n\sum_{i=1}^{k}\beta_i c_n^i \qquad (7)$$

Let (x^j, y^j) be the jth pair of input/output vectors where m is the number of test points. In this paper, these test data are obtained from simulations of the HFPID controller. Then the consequence parameters $c_0^i, c_1^i, \ldots, c_n^i$ (i = 1...k) can be obtained by solving a recursive least squares parameter identification problem [Takagi and Sugeno, 1985] where we determine the unknown coefficients by minimizing the error index $J=\sum_i(y_i-y'_i)^2$where y is the output of the HFPID and y' is the output of the Sugeno controller given by (7).

The necessary conditions satisfied by the solution is XC=Y where $Y=[y_1 \ldots y_m]^T$, $C=[c_0^1\ c_1^1 \ldots c_n^1 \ldots c_0^k\ c_1^k \ldots c_n^k]^T$, and X is a (m×k(n+1)) matrix defined by:

$$X=\begin{bmatrix} \beta_1^1 & x_1^1\beta_1^1 & \cdots & x_n^1\beta_1^1 & \cdots & \beta_k^1 & \cdots \\ & \vdots & & & & \vdots & \\ \beta_1^m & x_1^m\beta_1^m & \cdots & x_n^m\beta_1^m & \cdots & \beta_k^m & \cdots \end{bmatrix} \quad (8)$$

where

$$\beta_i(x^j)=\frac{A_1^i(x_1^j)\cdot\ldots\cdot A_n^i(x_n^j)}{\sum_i\left(A_1^i(x_1^j)\cdot\ldots\cdot A_n^i(x_n^j)\right)} \quad (9)$$

is the truth value of the ith rule evaluated at the vector x^j [Takagi and Sugeno, 1985; Ljung and Sodestrom, 1986]. Then, dentoe the ith row vector of matrix X defined in (8) by x, and the ith element of Y by y_i. C can be then computed by the iteration:

$$C_{i+1}=C_i+S_{i+1}\cdot x_{i+1}^T\cdot(y_{i+1}-x_{i+1}\cdot C_i) \quad (10)$$

$$S_{i+1}=S_i-\frac{S_i\cdot x_{i+1}^T\cdot x_{i+1}\cdot S_i}{1+x_{i+1}\cdot S_i\cdot x_{i+1}^T} \quad (11)$$

where S_i is a (n(k+1)×n(k+1)) square covariance matrix and C_m at the final iteration is the sought solution. The initial estimates, C_0 and S_0, are chosen as $C_0=0$ and $S_0=\alpha I$ where α is a large number and I is the identity matrix [Takagi and Sugeno, 1985].

Note that if rule 1 never fires (i.e. $\beta_1^j=0\ \forall j$), X is not full rank and XC=Y has no unique least squares solution. In other words, we obtain the same J for any arbitrary values of $c_0^1, c_1^1, \ldots, c_n^1$. For the recursive least squares, these coefficients are dependent on the initial estimates of C_0 and S_0. The rank deficiency of X will be compounded by numerical difficulties, etc. If a rule never fires for the test data given, then this rule should be eliminated to make the solution of the least squares problem unique. Also, this rule will not be applicable or relevant in all trajectories similar to the test data.

In order to approximate the hybrid fuzzy-PID controller, input/output training data must be obtained. We approximate the controllers for each arm separately. For one arm, we define $x_1=e$, $x_2=de/dt$, and $x_3=\int edt$ and $y=\tau$, the torque at the joint of the arm. The input variables are fuzzified using the same fuzzy sets and membership functions defined for the HFPID for use in (9). These values along with the crisp input data sets are used to calculate the truth values. Finally, the recursive least squares approach defined in (10)-(11) is used to calculate the consequence coefficients.

III-b. Determination of the GSC

For this controller, nonlinear polynomials of x_1, x_2, and x_3 are added to the original Sugeno controller in (4) to yield the controller in (2) and (3). If we consider only quadratic terms (i.e. $\alpha_1^i=\alpha_2^i=\alpha_3^i=2$), then there would be a total of 3^3 or 27consequence coefficients for every rule. Determining all 27 of these coefficients for the GSC would provide a more accurate approximation of the HFPID, but many of these terms may not significantly affect the performance of the system. Therefore, only the significant nonlinear terms need to be added. That is, only the nonlinear terms that significantly decrease the error index when of the OSC are kept and the other terms are discarded.

To determine which nolinear terms are significant, we first select the nonlinear terms for test according to expert knowledge of the system. These terms are added to the rule base of the GSC in (2) and (3). Then, the consequence coefficients are determined using the approach defined in equations (8)-(11). The error index of the GSC controller with the added terms is compared to that of the OSC. If there is a vast improvement in the error index, then the terms are kept. This procedure is repeated until there is no significant improvement in the error index.

IV. SIMULATION RESULTS

This section presents how well the generalized Sugeno controller was able to approximate the hybrid fuzzy-PID controller. Two areas are considered in this comparison: (1) the error index J and (2) how close the trajectory of the Sugeno-controlled robot is to the trajectory of the HFPID-controlled robot.

The input/output test data was gathered from the hybrid fuzzy-PID controller by simulating the robot motion for a maneuver from 0 to 3.0rad. To obtain rich enough test data, a disturbance of 0.008cos(150t) was added to the error measurement. The derivative of error was estimated by a numerical differentiaiton followed by band pass filtering. The HFPID had a rule base consisting of 10 rules. For the maneuvers of interest, 2 rules never fired, and 3 rules fired very lightly. Therefore, to make the least squares problem well posed, we eliminated the above 5 rules.

The next step is to calculate the consequence coefficients of the OSC and GSC using the recursive

least squares method. Note that the domain range of the input data, the membership functions, the fuzzy sets, and the if-sides of the applicable control rules are the same for both controllers. The torques are calculated using the generated fuzzy control rules for the Sugeno controller.

Finally, nonlinear terms were incorporated into the GSC and tested for their significance. The following terms were added: (1) an $(\int e)^2$ term, which should decrease any steady-state error that would occur and (2) a $(\int e)^2 + e^{\prime}\int e$ term, which should, in addition, increase damping. For both cases, a significant decrease in the error index was achieved when compared to that of the OSC, but case (2) provided more improvement. This is shown in Table 1. The GSC accurately approximated the HFPID for both cases, but all controllers seem to have similar but satisfactory performance. This is shown in Fig. 1. However, the additional terms did allow for the GSC to be more robust than the OSC. Robustness was tested by changing the mass, inertia, and length of the robot manipulator arms by 100% without changing the controller. Under these conditions, the OSC had a smaller overshoot in error when compared to the HFPID. The GSC showed even more improvement. See Fig. 2 for details.

	Arm 1		Arm 2	
Sug. Ctrl.	J	%	J	%
Original	1.1392×10^{4}	--	1.0081×10^{4}	--
$+(\int e)^2$	1.0462×10^{4}	8.2	0.8895×10^{4}	11.8
$+(\int e)^2 + e^{\prime}\int e$	1.0426×10^{4}	8.5	0.8053×10^{4}	20.1

Table 1: Error Indexes

V. CONCLUSIONS

This paper has presented a simple approach to designing a generalized Sugeno controller through a hybrid fuzzy-PID controller using recursive least squares.We have applied this approach to design a generalized Sugeno controller for a two-degree of freedom robot manipulator arm. The performance of this controller approximated very well the performance of the HFPID controller. Moreover, the generalized Sugeno was found to be the most robust.

ACKNOWLEDGMENT

This work is partially funded by grants from NASA-Autonomous Control Engineering Center under grant number ACE-48146. The authors wish to thank them for their financial support.

REFERENCES

Buckley, J.J. "Sugeno-type Controllers Are Universal Controllers." *Fuzzy Sets & Systems*, no. 53, pp. 299-303, 1993.

Sayyarrodsari, B. and Homaifar, A. "Hybrid Fuzzy-PID Controllers Are Nonlinear System Controllers." (to be published in the *International Journal of Fuzzy Sets & Systems*).

Takagi, T. and Sugeno, M. "Fuzzy Identification of Systems and Its Applications to Modeling and Control." *IEEE Trans. on Systems, Man, and Cybernetics*, vol. smc-15, no. 1, pp. 116-132, 1985.

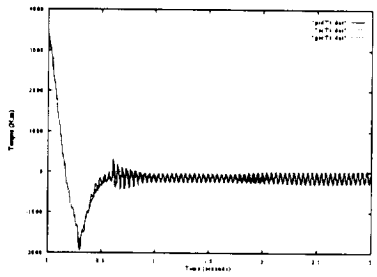


Fig.1: Arm 1 trajectories obtained by HFPID, OSC, and GSC.

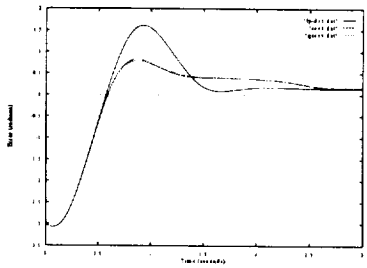


Fig. 2: Errors of Arm 1 for HFPID, OSC, and GSC when operating under mass and length uncertainty.

Supplemental Control for Aircraft Riding Qualities

Rolanda S. Williams[1] CR. Ashokkumar Abdollah Homaifar
NASA-CORE
North Carolina A&T State University
Greensboro, NC 27411

Abstract. The concept of preview control has been applied and proven to be successful in the automotive vehicle. These same concepts are now applied to an aircraft under the assumption that exogenous inputs (wind gust, turbulence, etc.) can be measured. A supplemental control law for surface deflection is designed to compensate for the loss in performance in the presence of atmospheric disturbances. Fuzzy logic control is employed to handle the nonlinear, time varying characteristics of the disturbance. A methodology to tune the outer loop control parameters is presented.

Keywords: supplemental control, exogenous inputs, flight control law, preview control, fuzzy logic control

I. Introduction

One in four Americans have a fear of flying. Even though air travel is the safest form of flying, the fear of crashing overwhelms that fact. Numerous recent aircraft crashes have certainly added to this fear. One of the most common passenger complaints is the effects of turbulence in passenger ride comfort. The study of flight in atmospheric turbulence is one of the most important areas in active control systems. It is difficult for a pilot to foresee exogenous inputs such as gust, turbulence, and other atmospheric disturbances and supplement a suitable control surface deflection to meet the standard flying qualities. Aircraft flying qualities are in general evaluated for both passenger's ride comfort and for ease of piloting. The pitch movements of the plane captures ride comfort. Human tolerance to normal accelerations unlike those felt on Earth is very poor. Thus it is imperative to limit vertical accelerations in order to maintain passenger and crew comfort [Mengali 1994].

Problem Definition

In this paper we investigate an aircraft control law based on knowledge of exogenous inputs. There is a proposed sensor that can detect and measure atmospheric disturbances. Given a stable flight control law, a supplemental control law can be designed to augment loss ride quality in the presence of disturbance. In this study, we must first determine if a control action can even be designed to minimize the effects of these inputs based on the information about exogenous inputs before applying preview control actions.

Due to the nonlinear and time varying nature of these disturbances, a fuzzy logic rule base will be used to design the previously mentioned controller. Fuzzy logic is also a favorable choice due to its ability to adapt to a wide range of parameters.

II. Controller Structure

This investigation will determine the effectiveness of supplemental control in the presence of the measurable exogenous inputs. The exogenous inputs were taken to be one dimensional. Although this may seem to be taken to be one dimensional. Although this may seem to be an oversimplification, this is a significant approach. Due to the isotropic nature of atmospheric disturbance, the concepts learned in this one dimensional problem can be applied to a more realistic three dimensional problem. This case study includes the feedforward block. The study of this response will determine the applicability of preview control based in the knowledge of exogenous inputs before they actually occur. Although preview control is not incorporated in this particular model, it is relevant to investigate the effectiveness of having the disturbance measurements. If this knowledge proves to be beneficial to the problem, then preview control can be applied.

Given the state equations for a unity feedback system representing a Boeing 777 aircraft we will evaluate this problem. For the ride quality problem, the emphasis will be put on studying the response of normal acceleration. A normal acceleration of 1 g (9.71 m/s^2) is desirable, as this is the normal acceleration felt on Earth. Using the previous state equations, a step response is taken.

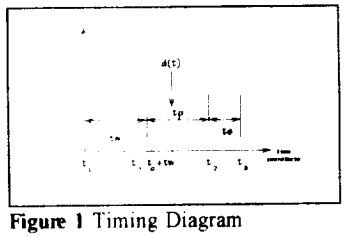


Figure 1 Timing Diagram

[1] Rolanda S. Williams is a graduate student in the Department of Electrical Engineering. Her research interests include the application of fuzzy logic control to automotive and aerospace systems. She has been a member of NASA-CORE since her senior year. Her expected graduation in May 1996.

The atmospheric disturbance is applied after the steady state has been reached. The disturbance acts on the system over a given time window (see figure 1) where

Table I Timing Variables

t_0	time that system reaches steady state
t_w	time window in that the system is at steady state before d(t)
t_1	time that d(t) acts on the system
t^p	time window that d(t) prevails
t_2	time that disturbance subsides
t_3	time at which the system returns to steady state

Next, the disturbance and supplemental controllers must be added to the mathematical models. The rusuting equations are

$$\dot{x}(t)=A_c x(t)+Br(t), \quad x(0)=x_0, \ t\geq 0 \qquad (1)$$

$$\dot{x}_d(t)=Ax(t)+B\delta_{ac}(t)+\Gamma d(t) \quad \forall t\in[t_1,t_2] \qquad (2)$$

$$\dot{x}(t)=A_c x(t)+Br(t), \ x(t_2)=x_2, t\in[t_2,\infty) \qquad (3)$$

where

Table II Controller Nomenclature

k	existing feedback vector
L(t)	feedfoward compensator
r(t)	reference input

Atmospheric disturbance is highly nonlinear and time varying. It has been established that adaptive proportional control and feedforward control will be utilized to minimize its effects. Now, a technique for developing these controllers must be established. A rule based logic seems to be most appropriate due to the nature of the disturbance.

IV. Fuzzy Logic Control

Fuzzy logic controllers have become very popular in recent years. The attractiveness is in alleviating the need for the designer to form an exact mathematical model of a system in order to design the controller. As systems become progressively more complicated, the ability to construct an exact system model becomes nearly impossible. The designer is able to design the controller through experience, intuitiveness, and approximate reasoning using as much exact data that can be found or formulated. This is especially helpful in processes that are normally controlled by human operators. A fuzzy controller can accept linguistic inputs and produce a crisp control action that can be understood by the system.

This logic, largely based on the spirit of human thinking, seems to be most appropriate in the ride quality problem. Although "good" ride quality is traditionally based on the eigen values of a given system, it may also be taken to be a subjective measurement. Ride quality is highly subjective in relation to passenger comfort and ease of piloting. This, along with the nonlinear and time varying nature of this problem, motivates the application of fuzzy logic [Lee 1990].

V. The Application of Fuzzy Logic

In this particular problem, fuzzy control was chosen for several reasons. The first being that there is no set model for random turbulence and wind gusts. They can come in the form of an exponential, sinusoidal, step function, or undefined function.[Holbit 1989] The FLC can be designed to handle the nonlinear, time varying disturbances of wind gusts and random turbulence. In common conventional control such as the PD controller, the gains are constant in all cases, whereas in a fuzzy system these constant gains can vary due to the different types of disturbances. The PD constants would have to be regenerated each time that a different type of disturbance occurred as opposed to an FLC that is designed to adjust to changes. This leads to a more flexible, efficient, and robust system that can operate under a wider range of conditions.

Fuzzy Feedforward Control

This component of the controller uses a feedforward path that considers the dynamic aspects of the aircraft control problem. The fuzzy feedforward path is turned on in the presence of exogenous inputs. This is analogous to a conventional feedforward loop. The inputs are the present and past error caused by the exogenous inputs. The rules decide on the most appropriate action based on these inputs. The measure of passenger comfort is normal acceleration. When one

is on the ground, one feels a normal acceleration of 1 g (9.71 m/s^2). When in an aircraft, passenger comfort is felt when the aircraft has a normal acceleration of 1 g. The error between this normal acceleration and the normal acceleration caused by the disturbance and the derivative of this error are the input of this system. The resulting command is the elevator deflection. The elevator command alters the position of the nose of the aircraft, thus changing the angle of attack. This directly affects the pitch rate of the system. The elevator command is a supplemental deflection that works in parallel with the primary control. Now that the fuzzy logic is in the control technique of choice, the inputs and outputs must be divided into fuzzy sets.

Designing the Membership Functions

Good membership functions are the key to effective fuzzy logic control [Sabharwal and Rattan 1992]. Guassian membership functions are used to divide the inputs into five fuzzy sets. These five fuzzy sets are PB (Positive Big), PS (Positive Small), Z (Zero), NS (Negative Small), and NB (Negative Big). The inputs are the error and change in error of the normal acceleration.

The range for the error input was chosen in order to achieve fine control close to zero and coarse control away from zero. In conventional PD controller, it has been shown that a small proportional gain when error is big and a large proportional gain when the error is close to steady state can lead to a more stable system. This can be accomplished by making the membership around zero more narrow than the other fuzzy sets.

As with automobile suspension systems, this control system was designed to handle the small discomforts of flying, not the huge potholes on the sky. A change in normal acceleration of one-half of a g is considered to be the largest reasonable error that should be handled by this controller. The support for each fuzzy set are not evenly spaced. The fuzzy sets close to zero are more narrow and closer together than the sets further from zero. This lends for finer control close to zero.

5.b.ii. Designing the Rule Base

The rule base consists of a set of feedforward rules using linguistic terms. The rules take the form:

If ***error*** *is* A_{error} *and* ***rate*** *is* A_{rate} *then loft is* A_{loft}

where A_i represents linguistic variables i

$A_{error} \in$ {NB, NS, Z, PS, PB}

$A_{rate} \in$ {NB, NS, Z, PS, PB}

$A_{loft} \in$ {NBI, NMD, NSM, NIL, PSM, PMD, PBI}

(For example: NB⇒Negative Big, Z⇒Zero, PS⇒Positive Small)

There are a possible 5^2=25 rules (2 inputs with five attributes each). Each possible rule combination was derived. In this feedforward path, the gain is applied to the exogenous input. It has been shown that the control is most effective when a large control action is applied when the error is close to zero and a small control action when the error away from zero [Sabharwal & Rattan 1992]. The rules were designed to be anticipatory, although no preview is involved.

VI. Experimental Results

A mathematical model is used to simulate the cruising aircraft (see equations). The duration of the simulation is 12 seconds with 50 samples per second taken. A step input is given and the disturbance is applied after steady state is reached. The disturbance is applied for 3 seconds and then released. The disturbance that is applied affects the normal acceleration at a maximum of ±0.5 g's. Various disturbances of different magnitude and functional characteristics are applied.

In analyzing the various data, there are several aspects that are looked upon as being desired. The first objective is to observe that the controller minimized the affects of the exogenous inputs. Also, it is desirable that the transition into and out of the disturbed atmosphere is smooth and continuous. The longitudinal vibrations and range of the elevator deflection are desired to be minimal. The desired response is the standard step response that setttle quickly to its steady state value.

In this simulation, the only the feedforward block of the controller is active. It is turned on the presence of disturbance. A "big" exponential disturbance is applied:

The abbreviated results are illustrated in figures 4-6. Figure (4) shows that the disturbance

$$w_g = 10e^{-0.447(t-5)}$$

causes the normal acceleration to decrease by 50%. Figure (5), the controlled response, show that this 50% decrease is corrected within 10% of the desired response. There is a 2% change in normal acceleration when the disturbance is applied. Once the disturbance is applied, the deflection is smooth, ranging between 0 and 3 degrees. After the disturbance is released at the 8 seconds, the signal returns to steady state within 2 seconds.

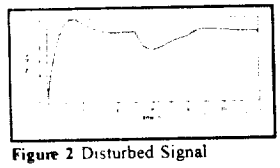


Figure 2 Disturbed Signal

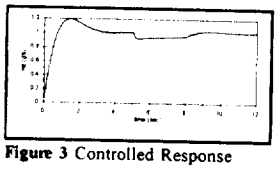


Figure 3 Controlled Response

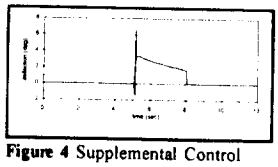


Figure 4 Supplemental Control

In order to test the robustness of this controller, different types of disturbances were investigated. The response in the presence of a sinusoidal disturbance beginning at 4 seconds and lasting for 4 seconds is shown in figures 7-9. In this scenario, the disturbance causes a 50% decrease and increase in normal acceleration. The controlled response shows the signal is corrected within 7.5% of the desired response, with a few minor bumps encountered. There is a 0.04% change in normal acceleration when the disturbance is applied. Once the disturbance is applied, the deflection is smooth, ranging between 0 and 3 degrees. After the disturbance is released at the 8 seconds, the signal returns to steady state within 2 seconds.

Note that the deflection is positive when the error is negative as well as when the error is positive. In this system the feedforward compensator L(t) is multiplied times the disturbance d(t) (see equation 16). When the error is negative (the reference is less than the output), the system needs to be pushed down. This requires a positive supplemental control because a negative (the error) times a positive (the control) results in a negative. When the error is positive (the reference is greater than the output), the system needs to be pushed up. This requires a positive supplemental control as well because a positive (the error) times a positive (the control) results in a positive.

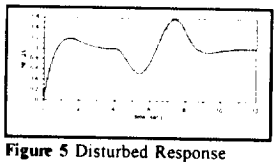


Figure 5 Disturbed Response

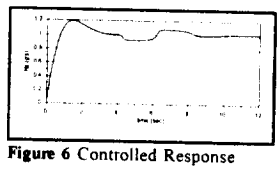


Figure 6 Controlled Response

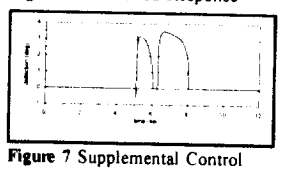


Figure 7 Supplemental Control

VII. Conclusions & Future Works

A properly designed supplemental controller is a valid approach to maintaining "good" aircraft ride qualities. The use of fuzzy logic control creates a robust controller that can perform well in a variety of wind disturbances. The controller was able to maintain a normal acceleration within 7% for both an exponential and sinusoidal disturbances without changing the membership functions or rule base.

While the feedforward controller produces good results, the stability can not be guaranteed. For future work, the attributes of optimal control will be exploited to design an adaptive proportional controller to work along with the feedforward controller.

Acknowledgements

This work is partially funded by grants from the NASA Center of Research Excellence at North Carolina A&T State University under grant number NAGW-2924. The authors wish to thank them.

VIII. Bibliography

Holbit, Frederic M., *Gust Loads on Aircraft Concepts and Applications*, AIAA Education Series, 1988.

Lee, C.C., "Fuzzy Logic In Control Systems: Fuzzy Logic Controller - Part I" IEEE Trans. Syst. Man Cybern., vol. 20, no. 2, pp 404-418, March-April 1990.

Mengali, Giovanni, "Ride Quality Improvements by Means of Numerical Optimization Techniques", Journal of Guidance, Control, and Dynamics, vol. 17, no. 5, pp. 1037-1041, September-October 1994.

Sabharwal, Deepak and Kuldip S. Rattan, "Design of Rule-Based Fuzzy Controller for the Pitch Axis of an Unmanned Research Vehicle", Proc. IEEE 1992 National Aerospace and Electronics Conference, Dayton Ohio, pp 449-455, May 1992.

290726
p.3

AIRCRAFT PITCH CONTROL WITH FIXED ORDER LQ COMPENSATORS

James Green* CR.Ashokkumar† A.Homaifar‡
The North Carolina A & T State University
Greensboro, NC 27411

ABSTRACT

This paper considers a given set of fixed order compensators for aircraft pitch control problem. By augmenting compensator variables to the original state equations of the aircraft, a new dynamic model is considered to seek a LQ controller. While the fixed order compensators can achieve a set of desired poles in a specified region, LQ formulation provides the inherent robustness properties. The time response for ride quality is significantly improved with a set of dynamic compensators.

I. Introduction:

While designing a feedback control, ride and handling qualities are major performance objectives in aircraft control problems. Such objectives are normally achieved by closed closed loop pole assignment [1]. Preserving these closed loop poles (within the desired regions) in the presence of perturbations is another requirement [2]. LQ problems have inherent stability margins to tolerate unstructured uncertainties. LQ design techniques with regional pole constraints have been studied extensively in the literature[see [3], and its references]. Similar approach, but with dynamic compensators, have been investigated for automotive applications [4]. The compensators given in [5] for aircraft control problem are considered in LQ problem setting. The objective of this approach is to improve aircraft ride quality defined in [1].

*Undergraduate Student, Dept. of Electrical Engineering.

†Post Doctoral Research Associate, NASA Center of Research Excellence.

‡Associate Professor, Dept of Electrical Engineering.

II. Problem Formulation:

An aircraft model in pitch plane [with normal acceleration (n_z), pitch rate (q) and elevator deflection (δ_e) as state variables and command input (δ_c) as control variable], is given by [2]:

$$\dot{x} = \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & -14 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} n_z \\ q \\ \delta_e \end{bmatrix}}_{x(t)} + \underbrace{\begin{bmatrix} b_1 \\ 0 \\ 14 \end{bmatrix}}_{b} u \tag{1}$$

It is well known that the control law

$$u(t) = -R^{-1}b'Px(t) + r(t) \tag{2}$$

minimizes the performance index

$$J = \int_0^\infty \{x'Qx + u'Ru\}dt \tag{3}$$

and satisfies the algebraic riccati equation

$$A'P - PbR^{-1}b'P + PA + Q = 0 \tag{4}$$

Selection of weighting matrices to achieve a controller in equation 2 for exact pole assignment has been extensively investigated in reference [3]. Suppose, we choose a set of dynamic compensators given in [5] for the control law structure [2] (see Figure-1), then the state equations for the compensators are:

$$\dot{z}_1 = -\tau_1 z_1 + n_z \tag{5}$$

$$\dot{z}_2 - q = -\tau_3 z_2 + \tau_2 q \tag{6}$$

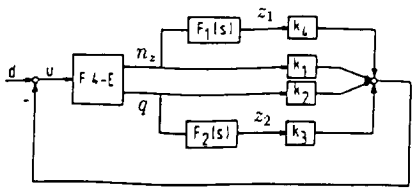


Figure 1: Control Law Structure [Ref 2] with Filters $F_1(s) = \frac{1}{s+\tau_1}$ and $F_2(s) = \frac{s+\tau_2}{s+\tau_3}$

From the aircraft dynamical equations 1, substituting for $\dot{q}$, we have

$$\dot{z}_2 = a_{21}n_z + (a_{22} + \tau_2)q + a_{23}\delta_e - \tau_3 z_2 \quad (7)$$

For the new state vector $\bar{x}(t)$,

$$\bar{x}(t) = [x(t), z_1(t), z_2(t)]'$$

the state space equations become,

$$\dot{\bar{x}}(t) = \bar{A}\bar{x}(t) + \bar{b}\bar{u}(t) \quad (8)$$

where,

$$\bar{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ 0 & 0 & -14 & 0 & 0 \\ 1 & 0 & 0 & -\tau_1 & 0 \\ a_{21} & (a_{22}+\tau_2) & a_{23} & 0 & -\tau_3 \end{bmatrix}$$

$$\bar{b} = \begin{bmatrix} b_1 \\ 0 \\ 14 \\ 0 \\ 0 \end{bmatrix}$$

It can be verified that for these dynamic compensators, the system in equation 8 is completely controllable. Thus the control law

$$\bar{u}(t) = -\bar{R}^{-1}\bar{b}'\bar{P}\bar{x}(t) + r(t) \quad (9)$$

minimizes the performance index

$$\bar{J} = \int_0^\infty \{\bar{x}'\bar{Q}\bar{x} + \bar{u}'\bar{R}\bar{u}\}dt \quad (10)$$

and satisfies the algebraic riccati equation

$$\bar{A}'\bar{P} - \bar{P}\bar{b}\bar{R}^{-1}\bar{b}'\bar{P} + \bar{P}\bar{A} + \bar{Q} = 0 \quad (11)$$

With the above formulations, we shall now present the closed loop eigenvalues for various values of the design parameters. The design parameters for J are obviously the weighting matrices Q and R. However, note that the performance index $\bar{J}$ is significantly influenced by the other design parameters τ_1, τ_2, and τ_3, in addition to $\bar{Q}$ and $\bar{R}$. The next section presents the simulation results.

II. Simulation Results:

For F-4 aircraft model at Mach=1.5, Altitude=35,000ft, the system dynamic matrices are given by:

$$A = \begin{bmatrix} -0.5162 & 26.96 & 178.9 \\ -0.6896 & -1.225 & -30.38 \\ 0 & 0 & -14 \end{bmatrix}$$

$$b = \begin{bmatrix} -175.6 \\ 0 \\ -14 \end{bmatrix}$$

The matrices $\bar{A}$ and $\bar{b}$ for the state vector $\bar{x}(t) = [x(t), z_1(t), z_2(t)]'$ are

$$\bar{A} = \begin{bmatrix} A & 0 & 0 \\ A_1 & -\tau_1 & 0 \\ A_2 & 0 & -\tau_3 \end{bmatrix}$$

$$\bar{b} = \begin{bmatrix} -175.6 \\ 0 \\ -14 \\ 0 \\ 0 \end{bmatrix} \quad \text{where,}$$

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} -0.6896 & (-1.225+\tau_2) & -30.38 \end{bmatrix}$$

At this flight condition, the short period damping (ζ_{sp}) and frequency (ω_{sp}) requirements are:

$$0.35 \le \zeta_{sp} \le 1.3 \quad (12)$$

and

$$3.29 \le \omega_{sp} \le 11.8 \quad (13)$$

In complex plane, these constraints impose regional pole constraints shown in Figure 2.

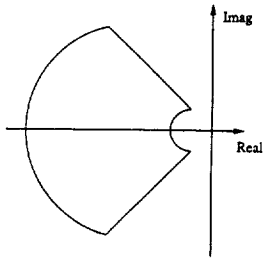


Figure 2: Regional Constraints for Aircraft

Table-1

Design Variables	ω_{sp}	ζ_{sp}
Q, R	4.5078	0.4789
$\bar{Q}$, $\bar{R}$	6.4458	0.5316
Desired	[3.29, 11.8]	[0.35, 1.3]

The weighting matrices $Q = I_3$ and $R = 10^4$ as well as the weighting matrices $\bar{Q} = I_3$ and $\bar{R} = 10^4$ provide the acceptable closed loop poles [see Table 1].

However, what needs to be observed is the time response plots (due to step input) shown in Figure 3. We observe that the normal acceleration at the sensor location is nonminimal. Moreover, the peak accelerations are siginificantly reduced with dynamic compensators (about 50%).

Acknowledgements:

This work is partially supported by grant from the NASA Center of Research of Excellence at NC A&T State University under grant # NAGW-2924. The authors wish to thank the NASA-CORE administration.

References:

1. "Flying Qualities of Piloted Airplanes," *MIL-F-8785B(ASG)*, Aug 7, 1969

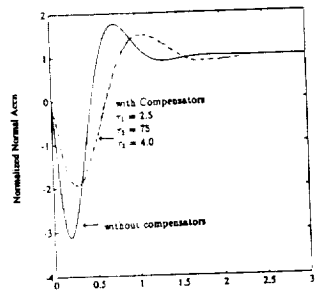


Figure 3: Time response polts due to step input

2. S.N. Franklin and J. Ackerman, "Robust Flight Control:A Design Example," *Journal of Guidance, Control, and Dynamics*, v 4, n 6, 1981, pp 597.

3. Y. Ochi and K. Kanai, "Pole Placement in Optimal Regulator by Continuous Pole-Shifting," *Journal of Guidance, Control, and Dynamics*, v 18, n 6, 1995, pp 1253-1258

4. H. Peng and M. Tomizuka, "Preview Control for Vehicle Lateral Guidance in Highway Automation," *Journal of Dynamic Systems, Measurement and Control*, v 115, 1993, pp 679-686

5. CR.Ashokkumar,"Robust Optimal Compensators with Tight Control Philosophy," *Report No. GCD/CRA/02*, NASA Center of Research Excellence, North Carolina A & T State University, Greensboro, NC

563-61

290727 p4

David versus Goliath: Should We Evolve Small or Large Populations?

Dean Walker
Autonomous Control Engineering
Department of Computer Science
North Carolina A&T State University
Greensboro, NC 27411
email: dwalker@ncat.edu

Gerry Dozier
Autonomous Control Engineering
Department of Computer Science
North Carolina A&T State University
Greensboro, NC 27411
email: gvdozier@ncat.edu

Abstract

Evolutionary Algorithms (EAs) are problem-solving methods based on natural selection. Unlike many search algorithms, which operate on a single candidate solution (CS), EAs operate on a population of individuals where each individual represents a CS. EAs select parents based on their fitness and allow them to create offspring. The offspring are then evaluated and are added to the population, usually replacing lesser fit individuals to keep the population size constant. This process is repeated until an individual representing a satisfactory CS has been evolved.

A Microevolutionary Algorithm (MEA) is a small population EA (population size < 20). Interest in MEAs has been increasing because of their ability to solve some problems quicker and more efficiently then EAs with larger population sizes.

In this paper, we will compare the performances of four MEAs and two EAs that use a conventional population size on two popular optimization problems.

Keywords: Evolutionary Algorithm, Microevolutionary Algorithm, Steady-State, Generational, Radcliffe's Crossover

1 Introduction

Evolutionary Search is a problem solving paradigm based on simulated evolution (Back 1991, Fogel 1966, Holland 1975). Search algorithms based on this paradigm are commonly referred to as Evolutionary Algorithms (EAs). Unlike many search algorithms, which operate on a single candidate solution (CS), EAs operate on a population of individuals where each individual represents a CS. The population of individuals maintained by a 'typical' EA may be as small as 20 or possibly as large as a few thousand.

After an initial population of randomly generated individuals has been produced, each individual in the population is assigned a fitness by an evaluation function. Individuals within the population are selected to be parents based on their fitness. The parents create offspring (which are also individuals that represent CSs) by either mutating themselves slightly, mating with other parents or both. The offspring are then

evaluated to determine their fitness and are added to the population, usually replacing lesser fit individuals in order to keep the population size constant. This process of selecting parents and allowing them to procreate based on their fitness is repeated until an individual representing an optimal or near optimal CS has been evolved.

EAs have been applied successfully to a variety of areas such as function optimization (Back 1991), scheduling (Corne 1994), aircraft design (Bramlette 1991), Constrained Optimization (Homaifar 1994) and Constraint Satisfaction Problems (Dozier 1995), among others.

A Microevolutionary Algorithm (MEA) is a small population EA (population size <20)(Karr 1991). Interest in MEAs has been increasing because of their ability to solve problems quicker and more efficiently than EAs with conventional population sizes.

In this paper, we will compare the performances of four MEAs and two EAs that use a conventional population size on two popular optimization problems. In Section 2, we describe the six EAs that will be compared, and in Section 3, we describe the two optimization problems that will be solved by our EAs. In Sections 4 and 5, we present our results and our conclusions.

2 The Six EAs

The six EAs that will be compared in this paper use tournament selection (Goldberg 1991) to select individuals to be parents, Radcliffe's crossover operator (Radcliffe 1991) and no mutation.

In tournament selection, two individuals are selected from the population at random and the individual with the better fitness is selected to be the first parent. This process is repeated to get the second parent. The four MEAs will incorporate a modified tournament selection method in which the best individual in the current population is always selected as one of the two parents to undergo Radcliffe's crossover.

Using Radcliffe's crossover, two parents $\langle x_1, x_2 \rangle$ and $\langle y_1, y_2 \rangle$ create a child $\langle random(x_1, y_1), random(x_2, y_2) \rangle$ where $random(x, y)$ is a random floating-point value between x and y.

The six EAs are of the form $EA(p, i)$ where p represents the population size and i represents the number of individuals in the current population that are replaced by i offspring to create a new population of individuals[1]. The six EAs are as follows: $EA(100,99)$, $EA(5,4)$, $EA(3,2)$, $EA(100,1)$, $EA(5,1)$, and $EA(3,1)$. The first three EAs are called generational EAs and the last three EAs are called steady-state EAs (Davis 1991).

Because of the small population size of the MEAs, they are more prone to converge prematurely to a sub-optimal solution. To overcome this problem, we use a restart method similar to the one described by Goldberg (Goldberg 1989, Krishnakumar 1989). Our restart method is as follows. We evolve our population of individuals until the difference between the best fit individual and the worst fit individual is less than a pre-specified threshold (we use 0.0000001 as our threshold value). Once the difference between the best and the

[1] The worst i individuals are replaced by the i offspring.

worst individual of the current population is below this threshold, we save the individual with the best fitness and randomly generate the other p-1 individuals.

3 The Test Suite

To compare the performance of the six EAs, we used two popular problems $f6(x, y)$ and $f7(x, y)$ where x and y can be assigned values from -100 to 100 (Schaffer 1989). These problems are defined as follows:

$$f6(x,y) = 0.5 + \frac{\sin\sqrt[2]{x^2+y^2} - 0.5}{[1.0+0.001(x^2+y^2)]^2}$$

$$f7(x,y) = (x^2+y^2)^{0.25}[\sin^2(50(x^2+y^2)^{0.1})+1.0]$$

4 Results

Figure 1 shows the performance of the six EAs on $f6$ and $f7$. Each of the EAs was run 100 times on $f6$ and on $f7$. For each run, a maximum of 4000 function evaluations were allowed. If an EA did not find a solution that was better than a pre-specified near-optimal solution (0.99754414 for $f6$ and 23.309239 for $f7$) within 4000 function evaluations, the search was terminated and the run was counted as a failure.

In Figure 1, the average number of function evaluations needed to find a solution that was better than the pre-specified near-optimal solution is recorded along with the percentage of the 100 runs that were successful.

In Figure 1, EA(5,1) has the best performance on $f6$ and $f7$. All of the MEAs seem to outperform EA(100,99) and EA(100,1). The MEAs seem to have better performance because their small population size allows them to converge quickly and then restart the evolutionary process while keeping the best solution found so far. Also in Figure 1, the steady-state MEAs seem to outperform the MEAs.

EAs	f6	f7
EA(100,99)	2343.0 (0%)	- (0%)
EA(5,4)	1265.7 (96%)	1862.6 (29%)
EA(3,2)	1155.5 (98%)	2261.1 (32%)
EA(100,1)	- (0%)	- (0%)
EA(5,1)	806.5 (100%)	1842.3 (55%)
EA(3,1)	1077.0 (97%)	1891.6 (24%)

Figure 1:A comparison of the six EAs

5 Conclusions

In this paper, we have shown that for two function optimization problems, MEAs are more efficient search procedures than EAs that evolve large population sizes. Our ongoing research is devoted to further comparisons of MEAs and EAs on a variety of problems including those which have computationally expensive evaluation functions.

References

Back, Thomas (1991). A survey of Evolution Strategies, *Proceedings of the Four International Conference on Genetic Algorithms*

Bramlette, Mark f., and Bouchard, Eugene E. (1991). Genetic Algorithms in Parametric Design of Aircraft, in: L. Davis (ed.) *Handbook of Genetic Alogrithms*. Van Nostrand Reinhold, New York.

Corne, D., Ross P., and Fang, H-L(1994) Fast Practical Evolutionary Timetabling, *Proceedings of the AISB Workshop on Evolutionary Computation, Springer-Verlag.*

Davis, L. (1991). *Handbook of Genitic Algorithms.* Van Nostrand Reinhold, New York.

Dozier G., Bowen, J, and Bahler, D. (1995). Solving Randomly Generated Constraint Satisfaction Problems Using a MicroEvolutionary Hybrid That Evolves a Population of Hill-Climbers, *Proceedings of the 1995 IEEE International Conference on Evolutionary Computation.*

Fogel, L. J., and Walsh, M. J. (1966). *Artificial Intelligence Through Simulated Evolution*, John Wiley & Sons, Inc.

Goldberg, D. E. (1989). Sizing Populations for Serial and Parallel Genetic Algorithms, *Proceedings of the Third International Conference on Genetic Algorithms.*

Goldberg, D. E., and Deb K. (1991). A Comparative Analysis of Selection Schemes Used in Genetic Algorithms, Foundations of Genetic Algorithms, Gregory J. E. Rawlins (ed.)

Holland J. H. (1975) *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Harbor, MI.

Homaifar, A., Qi, C. X., and Lai, S. H. (1994). Constrained Optimization Via Genetic Algorithms, *Simulation* 62:4, 242-254.

Karr, C. L. (1991). Air-Injected Hydrocyclone Optimization via Genetic Algorithms, in: L. Davis (ed.) Handbook of Genetic Algorithms. Van Nostrand Reinhold, New York.

Krishnakumar, K. S. (1989). Microgenetic Algorithms for Stationary and Non-stationary Function Optimization, *Proceedings of the SPIE Intelligent Control and Adaptive Systems Conference.*

Radcliffe N. J. (1991). Forma Analysis and Random Respectful Recombination, *Proceedings of the Fourth International Conference on Genetic Algorithms.*

Schaffer, J. D., Caruna, R. A., Eshelman, L. A. and Das, A. (1989). A Study of Control Parameters Affecting Online Performance of Genetic Algorithms for Function Optimization, *Proceedings of the Third International Conference on Genetic Algorithms.*

Multi-Axis Fluidic Thrust Vectoring of a Supersonic Jet using Counterflow

E. L. Barber[1], F. S. Alvi, A. Krothapalli

CeNNAs, Florida A&M University and Florida State University, Tallahassee, Florida 32310

Abstract

The most common techniques currently used to efficiently vector supersonic jets require external flaps and/or pivoting devices. Fluidic thrust vectoring using counterflow eliminates the need for such complex hardware. Thus, the promise of decreases in both weight and drag as well as increased maneuverability makes this technique an attractive alternative. This technique has been successfully employed to achieve single axis fluidic thrust vectoring of a Mach 2 rectangular jet. To better compete with contemporary systems the current study extends this technique to multi-axis thrust vectoring of a Mach 2 diamond-shaped jet by applying counterflow to one of its four sides. To evaluate the performance of this technique the Planar Laser Scattering (PLS) technique is used to show the continuous vectoring of the diamond jet up to 20 degrees. Also, cross-stream PLS images are acquired to show the vectoring can be achieved off all four surfaces of the diamond jet.

Introduction

Flight tests have demonstrated the performance advantages of using thrust vector control over more traditional approaches which rely on aerodynamic surfaces for maneuverability. Thrust vector control of high performance fighter aircraft has been shown to improve control and agility at high angle-of-attack conditions as well as at low speeds where aerodynamic surfaces are relatively ineffective. Furthermore, it is envisioned that thrust vector control may eliminate much of the vertical and horizontal tail structure of the aircraft, significantly reducing weight and drag and thereby increasing aircraft range, as well as reducing takeoff and landing distances providing functional operation on damaged airfields. These structural alterations would, in addition, provide the benefit of a considerably reduced radar cross signature.

Current thrust vectoring technologies have focused on both two-dimensional pitch vectoring systems as well as axisymmetric multi-axis vectoring nozzles. However, these promising concepts require complicated hardware to redirect the engine exhaust, adding considerable weight to the aircraft and imposing limitations in terms of jet dynamic response and nozzle cooling. Although recent advances have been made in developing light weight, multi-axis vectoring nozzles with low external drag[1,2] design challenges at the system integration level still exist which must overcome the actuation and control hardware requirements needed to redirect the engine exhaust by mechanical means.

A different approach to jet thrust vectoring — commonly referred to as fluidic control — uses one or more secondary air streams to redirect the primary jet thereby achieving thrust vector control in the absence of moving surfaces. Early studies of fluidic control[3,4] demonstrated that small secondary jets could be used effectively to control the vector response of the primary flow. Although large vector angles could be achieved in this fashion, nonhysteretic jet response was possible only over a very limited portion of the operating domain, seriously reducing the likelihood that the technology would be viable in aircraft applications. At present, most fluidic control schemes, including some recent advances based on boundary layer control[5,6], suffer from similar stability issues which must be addressed before these non-mechanical approaches can be moved out of the laboratory.

A relatively new fluidic control technique, which employs secondary flow traveling in a direction opposite to the primary jet — namely

[1] Edward Barber received his BS degree from Florida A&M University in 1994, and is presently enrolled in the Florida A&M University-Florida State University College of Engineering pursuing his Masters degree in Mechanical Engineering. Edward will obtain his MS in 1996, and then pursue his Ph.D. Edward's primary studies are in the field of fluids and his work has focused on the building of a supersonic shock tunnel for the investigation of hot jet screech tones.

counterflow control — was recently explored by Strykowski and Krothapalli[7]. Counterflow Thrust Vectoring *(CFTV)* has been shown to be very effective in achieving pitch vectoring of subsonic[8] and supersonic[9] rectangular jets up to Mach 2. These studies conclusively show that the primary jet can be continuously vectored up to angles approaching 20 degrees. Furthermore, if implemented correctly, *CFTV* does not suffer from the bistability problems encountered with earlier fluidic control schemes. In this note we will describe the results of a study in which a Mach *2* diamond shaped jet was used to extend the single-axis vectoring concept of rectangular jets to multi-axis operation. As the following results will demonstrate, we were able to continuously vector the diamond jet in two mutually perpendicular planes to achieve pitch and yaw control. In addition to the inherent advantages of fluidic control techniques — stationary control surfaces and relatively simple hardware requirements — *CFTV* has the added advantage that the control surfaces are cooled by the counterflowing air entrained from cold surroundings.

Facilities

The experiments were conducted in the blow-down compressed air facility of the Fluid Mechanics Research Laboratory located at Florida State University which is capable of supplying hot and cold air to drive the primary jet. The Mach 2 diamond-shaped nozzle used for the experiment was designed to have the same throat and exit area as the rectangular nozzle used for the supersonic study[9]; details of the diamond jet can be found in reference 10. The stagnation to ambient pressure ratio across the nozzle was fixed to achieve isentropically expanded flow at Mach 2. The jet was operated cold at a stagnation temperature of 300 K with Re = 1.23 x 10^6, based on the short dimension of the jet.

To achieve thrust vector control using counterflow a secondary air stream needs to be established along the outer surface of one of the jet shear layers. This secondary stream is created by connecting a vacuum pump to a manifold and suction chamber placed along the periphery of the jet; details of this arrangement are discussed below. The thrust vector response using counterflow was quantified using planar laser sheet (PLS) imaging of the flow. The laser sheet was oriented along the diamond jet axis in the plane in which the jet was vectored. Sub-micron ice crystals are formed due to the condensation of the moist ambient air entrained into the shear layers at the periphery of the cold jet. These ice crystals scatter the light from the laser sheet thus rendering the jet shear layers visible.[11]

Results

Recent studies of pitch vectoring using counterflow indicate that the amount of secondary flow required to achieve a desired thrust vector angle is inversely proportional to the area of the shear layer along which the secondary flow is applied.[9] For this reason the diamond nozzle was selected as a candidate geometry with which to demonstrate multi-axis thrust vectoring capability. In contrast to the 4:1 aspect ratio rectangular jets studied previously, which produce only two shear layers of significant size along which counterflow can be established, the diamond jet is bounded by four shear layers of equal size. Accordingly, by creating a secondary counterflowing stream along the outer periphery of any of the four shear layers, the jet can, in principal, be vectored in the direction normal to the shear layers, providing multi-axis control.

Side-view and end-view schematics of the counterflowing nozzle-collar assembly and the diamond nozzle are shown in Fig. 1.

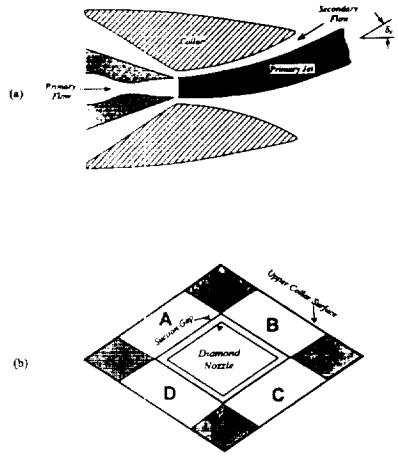


Fig. 1: Schematic of multi-axis fluidic thrust vector nozzle and collar geometry showing (a) side view and (b) end view.

The side view also indicates the conditions necessary to achieve jet vectoring at an angle δ_V as a result of

the secondary counterflow which is established along the upper shear layer. The collar provides four gaps at the nozzle exit, between the nozzle and the collar, through which suction can be applied. When the vacuum system connected to a particular chamber of the diamond nozzle (labeled A-D in Fig. 1) is activated, a secondary counterflowing stream is created within the gap defined by the collar surface and the corresponding shear layer of the jet. The collar was designed such that the counterflow stream is applied, in principle, to only one of the shear layers and isolated from the other three, a requirement of the *CFTV* technique. (The results presented below will indicate that some leakage occurs between neighboring chambers affecting the thrust vector response to counterflow.) Nozzle-collar design parameters, such as gap height, collar length and collar curvature chosen in the present study of the diamond jet, were taken from the performance predictions based on earlier studies of single-axis *CFTV* nozzles.[9] The diamond jet collar design was chosen to produce a jet deflection angle δ_V between 15 and 20 degrees at maximum suction. As the subsequent discussion will show, this design approach proved to be reasonably successful.

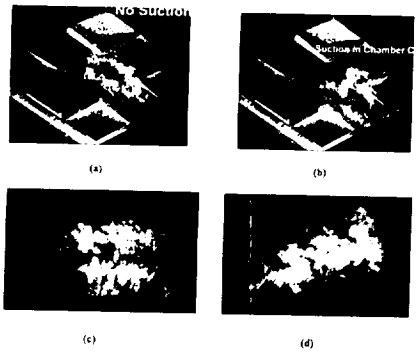


Fig. 2: PLS images showing the vectored and vectored diamond jet. a) Cross-sectional view, unvectored jet. b) Cross-sectional view, jet vectored into chamber C. c) Side view, unvectored jet. d) Side view, jet vectored to 16 deg.

Instantaneous PLS images showing cross-sectional views of the unvectored and vectored diamond jet are presented in Fig. 2a and 2b, respectively. No suction has been applied for the unvectored case shown in Fig. 2a, while Fig. 2b corresponds to the case with maximum suction/counterflow in chamber C. As the image clearly shows, the jet has been vectored into chamber C. Note that the unvectored jet in Fig. 2a appears to be off-center due to the oblique viewing angle of the camera. The large vectoring angles which can be achieved using multi-axis *CFTV* are perhaps more dramatically displayed in the side view images shown in Figs. 2c and 2d. Fig. 2c depicts the unvectored jet with no suction and Fig. 2d shows the jet at maximum deflection ($\delta_V \sim 15$ degrees).

Similar to the results of the single-axis study,[9] we were able to continuously vector the diamond jet into any of the four chambers by establishing counterflow via suction in the desired chamber. The relationship between the degree of counterflow and jet deflection was explored by varying the suction pressure and measuring the jet deflection angle from side view PLS images. A summary of the diamond jet vectoring angles, δ_V, as a function of the counterflow static pressure parameter is provided in Fig. 3. The static pressure parameter is the leading order term obtained from a control volume analysis of the nozzle-collar system for ideally expanded jets, details of which can be found in reference 9. The parameter is essentially a non-dimensional ratio of the side force acting on the jet, represented by $\Delta P_{exit} \cdot A_{side}$, and the axial force imposed by the jet $\rho_1 U_1^2 \cdot A_{jet}$. The pressure ΔP_{exit} represents the vacuum pressure established in the secondary stream as measured in the jet exit plane, and increases as the counterflow is increased.

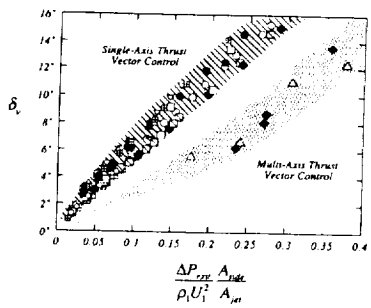


Fig. 3: Thrust vector response curve. Upper curve shows data for single-axis vectoring (ref. 9), lower curve shows data for the multi-axis thrust vector study.

For completeness, data from the single-axis rectangular jet experiment (ref. 9) is also included in Fig. 3. The upper curve represents the single-axis *CFTV* results and includes data for hot and cold jets measured over a range of collar design parameters,

i.e. various collar lengths and suction gap heights. Estimates of δ_V for the single-axis study were obtained using a control volume analysis including integrated pressure profiles measured along the collar surface as well as flow visualization images using the PLS technique employed in the present study. Within experimental uncertainty the single-axis data essentially collapse onto a single curve displaying an approximately linear dependence on the nondimensional pressure parameter.

The good collapse of the single-axis thrust vector data over a wide parametric range indicates that the scaling in Fig. 3 captures the essential features of the pressure forces and momentum fluxes contained in a complete control volume analysis applied to the jet-nozzle system. Furthermore, the agreement between the vectoring angles obtained using the control volume analysis and those estimated using PLS images indicates the appropriateness of employing optical techniques for determination of δ_V.

The data for the current multi-axis study is shown in the lower curve, where the scatter and the uncertainty in the measurements is roughly indicated by the shaded region. As in the single-axis study, the jet thrust vector angle is almost linearly proportional to the amount of counterflow. At the maximum suction pressure of approximately 51.7 kPa — a pressure limitation primarily imposed by the efficiency of the vacuum system — a maximum deflection angle of approximately 15 degrees was obtained. Although the multi-axis data show reasonably linear behavior, similar to the performance of a single-axis system, the curves for the two cases are distinct and the single-axis geometry appears to be slightly more efficient.

We believe that the primary reason for the differences in thrust vector performance of the single- and multi-axis systems is due to a noticeable leakage between adjacent chambers of the collar assembly of the diamond jet. This was determined by monitoring the static pressure distributions in the four chambers of the diamond nozzle. Leakage effects indicate that the counterflow is not confined to a single shear layer of the diamond jet, resulting in a degradation of performance due principally to the transverse pressure gradient across the shear layer. Although these three-dimensional effects will undoubtedly influence the collar performance, we believe that they can be minimized by improvements in collar design, which in the present study was not optimized, but used to demonstrate proof of concept of a multi-axis thrust vector system based on fluidic control.

Conclusions

Fluidic thrust vector control was examined in a Mach 2 diamond nozzle to determine the efficacy of multi-axis control using counterflow. The results clearly demonstrate that a secondary counterflowing stream applied along the periphery of the primary shear layer can be used effectively to achieve nonhysteretic thrust vector control in multi-axes and up to angles approaching at least 15 degrees. These results are an extension of previous studies where Counterflow Thrust Vectoring was applied to single-axis pitch control at Mach numbers up to 2, and indicate the robustness of the *CFTV* concept, which has many attractive features including the elimination of mechanical control surfaces and the inherent film cooling made possible by the ambient counterflow.

Acknowledgments

Support for this research was provided by the Office of Naval Research (contract N00014-92-J-1406) and NASA (contract NAG-2930) and is gratefully acknowledged.

References

[1] Snow, B.H., "Thrust Vectoring Control Concepts and Issues," SAE Technical Paper Series, 901848, 1990.

[2] Capone, F.; Smereczniak, P.; Spetnagel, D. & Thayer, E., "Comparative Investigation of Multiplane Thrust Vectoring Nozzles," AIAA Paper 92-3264, July 1992.

[3] Comparin, R.A.; et al., "On tne Limitations and Special Effects of in Fluid Jet Amplifiers," ASME Symposium on Fluid Jet Control Devices, 28 November 1962.

[4] Warren, R.W., "Some Parameters Affecting the Design of Bistable Fluid A\mplifiers," ASME Symposium on Fluid Jet Control Devices, 28 November 1962.

[5] Fitzgerald, R.E. & Kampe, R.F., "Boundary Layer TVC for Missile Applications," AIAA Paper 83-1153, June 1983.

[6] Carrol, G.R. & Cox, H., "A Missile Flight Control System Using Boundary Layer Thrust Vector Control," AIAA Paper 83-1149, June 1983.

[7] Strykowski, P.J. & Krothapalli, A., "The Countercurrent Mixing Layer: Strategies for Shear-layer Control," AIAA 3rd Shear Flow Control Conference, paper 93-3260, Orlando, FL, 6-9 July, 1993.

[8] Van der Veer, M. R. "Counterflow Thrust Vectoring of a Subsonic Rectangular Jet," M.S. Thesis, University of Minnesota, Minneapolis, Minnesota, 1995.

[9] Strykowski, P.J. & Krothapalli, A., Forliti, D. and Van der Veer, V. M. "Counterflow Thrust Vectoring of Supersonic Jets," AIAA 34th Aerospace Sciences Meeting, paper AIAA-96-0115, Reno, NV, 15-18 Jan., 1996.

[10] Krothapalli, A., Alvi, F. S., Washington, D. and King, C. J., " Aeroacoustic Characteristics of a Supersonic Diamond-Shaped Jet," First Joint CEAS/AIAA Aeroacoustics Conference, Munich, 1995. Submitted to the AIAA Journal.

[11] Clemens, N.T. & Mungal, M.G., "A planar laser Mie scattering technique for visualizing supersonic mixing flows," *Exper. in Fluids*, Vol. 11, No. 2/3, 1991, pp. 175-185.

S65-34

290729

p5

THE FLOWFIELD CHARACTERISTICS OF A MACH 2 DIAMOND JET

Donnell Washington[1], Farrukh S. Alvi, Anjanevulu Krothapalli
Department of Mechanical Engineering
CeNNAs, Florida A&M University and Florida State University
Tallahassee, Florida

ABSTRACT

The potential for using a novel diamond-shaped nozzle which may allow for superior mixing characteristics of supersonic jets without significant thrust losses is explored. The results of flow visualization and pressure measurements indicate the presence of distinct structures in the shear layers, not normally observed in shear layers of axisymmetric and rectangular jets. The characteristics of these features suggests that they are a manifestation of significant streamwise vorticity in the shear layers. Despite the distinct nature of the flowfield structure of the present shear layer, the global growth rates of this shear layer were found to be very similar to its two-dimensional and axisymmetric counterparts. These and other observations suggest that the presence of streamwise vorticity may not play a significant role in the global development of a compressible shear layer.

INTRODUCTION

Stealth capabilities of modern fighter aircraft require the use of complex geometries, such as the one considered in this paper. These nozzles generally provide not only a reduced radar signature, but also a much reduced infrared (IR) signature. The reduction in IR signature is attributed to enhanced mixing of the hot exhaust gases with the cold ambient air. This effect also frequently leads to reduced far-field noise. However, the detailed flow characteristics of jets issuing from these nozzles are not yet available. It is the purpose of this paper to provide the basic properties of a jet issuing from a converging-diverging diamond shaped nozzle; a representative of a stealth nozzle. The noise characteristics of such a jet may also prove to be beneficial for the High Speed Civil Transport (HSCT) application. An investigation concerning this application is discussed in a recent paper by Krothapalli et al. (1995).

Mixing and acoustic characteristics of a supersonic jet are largely determined by the behavior of compressible shear layers in the jet near-field. Studies by Brown and Roshko (1974), Papamoschou and Roshko (1988) and recent work by Clemens and Mungal (1992) and Shau et al. (1993), among others, have afforded considerable insight into the behavior of compressible shear layers. To date, most researchers have focused on two types of jet geometries: rectangular or two-dimensional and circular or axisymmetric. Although rectangular jets may have somewhat better mixing properties than axisymmetric jets their higher aspect ratios lead to greater frictional losses in the nozzle, and hence reduced thrust.

Diamond Jet

It is clear that axisymmetric and rectangular jets do not have all the desirable characteristics needed for civilian and defense applications. The motivation behind the present research is to study novel jet geometries which may potentially have superior mixing and stealth characteristics coupled with minimal thrust loss. One such novel geometry, which to the authors' knowledge has not yet been investigated, is a diamond shaped jet. Such a jet is the focus of the present study. A secondary reason for choosing this particular shape was based on the work of Pannu and Johannesen (1976) who investigated the flowfield generated by notched axisymmetric nozzles.

[1] Donnell Washington II received his BS in Mechanical Engineering from Washington State University in 1992, and his MS degree from Florida A&M University in 1995. He plans to graduate with his Ph.D. in 1997 or 1998 from Florida A&M University. In his graduate career Donnell has co-authored several journal publications. In addition he has presented papers at various conferences including ASME, APS, and several NASA meetings. The focus of Donnell's research is the characteristics of supersonic jets.

the rectangular nozzle whose data is shown in Fig. 5 has an aspect ratio of four which makes a comparison with this data somewhat suspect.

Nevertheless, for the sake of completeness, the rectangular nozzle data has also been shown, to at least provide some qualitative insight. In this non-dimensionalized space, the rectangular jet has the shortest potential core and the axisymmetric jet has the longest, while the diamond jet is between the two. This agrees with our earlier discussion, where we submitted that the mixing characteristics of the diamond jet shear layers are expected to lie between the two-dimensional and axisymmetric shear layers.

Shear Layer Growth Rates

Though the relative lengths of the potential cores yield some insight into the mixing behavior of the three different types of shear layers, our caveat regarding the appropriateness of D_{eq} for the high aspect ratio rectangular nozzle, makes such a comparison somewhat qualitative. To obtain a more reliable estimate of the shear layer growth, detailed pitot surveys, across the shear layer were also conducted at several streamwise locations. At each location the shear layer pitot thickness, δ_{pit} , was estimated using the well known criterion of $0.05P_{pit}$ - $0.95P_{pit}$ (Papamoschou and Roshko, 1988). A plot of the δ_{pit} as a function of streamwise distance is shown in Fig. 6. The error bar, indicated at only one of the points is representative of uncertainties at all the points. The shear layer growth rate, δ_{pit}/dx, estimated from a linear fit to the data and normalized by the corresponding incompressible value, provides a normalized growth rate of 0.27. This value lies within the data scatter of published data for both axisymmetric and two-dimensional shear layers at comparable M_c' s. This measurement validates the somewhat unanticipated results of the visual growth rate - the present shear layer has a growth rate comparable to axisymmetric and two-dimensional shear layers. One final note regarding the growth rates. Although the growth rates from visual and pitot measurements lie within the rather large scatter of published data, the 'pitot' growth rate extracted from visual data was approximately 30% higher than the corresponding rate deduced from pitot surveys. This discrepancy is likely due to the fact that the scaling factor of 0.72, used to convert visual thickness to pitot thickness, is primarily based upon visual measurements using *schlieren* images, whereas the current visual data relied on *PLS* images. Consequently, 0.72 may not be the suitable scaling factor. Since to our knowledge no such scaling factor exists for PLS images, we are forced to use this value. With the increasing use of laser imagery and a larger set of PLS data available, it is likely that a more appropriate scaling factor will soon be available.

Shear Layer Vortical Structures

The presence of streamwise vorticity, strongly suggested in the PLS images, was further investigated via pitot measurements. This was accomplished by conducting pitot surveys in the shear layer in the shape of the nozzle contour (in a diamond-shaped locus of points) at several streamwise locations. A plot of such profiles, at three streamwise stations, is shown in Fig. 7. The four corners of the diamond have been marked as A, B, C, and D, and correspond to the locations in the sketch of nozzle contour, shown in the inset. To accentuate the fluctuations in the pitot pressures, the mean pitot pressure, P_{pit} (the average value of the pitot pressure along the nozzle contour), is subtracted from the local pitot pressure, P_{pit} and the reduced pressure (P_{pit} - P_{pit}) is expressed as percentage of the local pitot pressure. The value $((P_{pit} - P_{pit})/P_{pit}) \times 100\%$, which represents the percentage variation of the local pitot pressure, is plotted in the ordinate. The abscissa is simply a number representing the measurement location along the nozzle contour.

The data shown in Fig. 7, corresponds to surveys conducted in the middle of the shear layer at each x location. The same trends are observed in the surveys in other regions of the shear layer as long as one does not come too close to the shear layer edges. As expected, four distinct peaks are observed in the vicinity of the jet corners. Moreover, the shape of these perturbations, the minima located at the corners and the maxima on either side, suggest the presence of two counter-rotating vortices at each corner. If this is indeed the case, we expect the strength of these vortices should be *inversely* proportional to the included angle at the diamond jet corners. The profiles in Fig. 7 clearly confirm this hypothesis, pressure difference between the local maxima and minima is significantly higher at A and C.

A more intriguing feature of these profiles is the presence of localized maxima and minima in the shear layers in regions away from the corners. The location of these peaks and valleys roughly coincides with the location of the ripples observed in the PLS images. More significantly, the overlap of the profiles from three streamwise stations strongly suggests the stationary nature of the features responsible for this pressure distribution. This evidence, together with the visual features discussed previously, provides further support for our hypothesis regarding the presence of streamwise vortices. A more rigorous proof of streamwise vorticity requires direct velocity measurements of the shear layer. Such measurements constitute the next

stage of this ongoing study and will be conducted in the near future.

CONCLUDING REMARKS

The flowfield of a novel diamond-shaped nozzle is explored in this paper. The results of the pitot and flow visualization studies show compelling evidence of substantial streamwise vorticity in the shear layer, a feature not observed in two dimensional shear layers. Interestingly enough, though the turbulent and mean flow field *structure* of the present shear layer is very different than shear layers generated by axisymmetric and rectangular jets, it does not appear to have a significant effect on the *global growth rates* of the shear layer. The growth rate of this shear layer was found to be very similar to its two-dimensional and axisymmetric counterparts. The acoustic properties of the diamond jet have also been studied, the results of which are presented in a separate publication (Krothapalli et al., 1995).

Given the fact that this geometry has never been studied in the past, the measurements presented herein are valuable in their own right, but many issues remain unresolved. Issues such as the strength of the streamwise vortices and why they have an apparently insignificant effect on the global shear layer growth. To answer some of these questions, we plan to make detailed velocity measurements using the Particle Image Velocimetry (PIV) technique in the immediate future.

ACKNOWLEDGEMENTS

We would like to thank E. Barber and J. King for their assistance in conducting the experiments. We also gratefully acknowledge the support of NASA (NAG 2930) in conducting this research.

REFERENCES

Arnette, S. A., Samimy, M. and Elliott, G. S., "On Streamwise Vortices in High Reynolds Number Supersonic Axisymmetric Jets," *Physics of Fluids A*, Vol. 5, 1993, pp. 187.

Brown, G. L. and Roshko, A. R., " On Density Effects and Large Structure in Turbulent Mixing Layers," *Journal of Fluid Mechanics*, Vol. 64, No. 4, 1974, pp. 775-781.

Clemens, N. T. and Mungal, M. G., "Two-and-three-Dimensional Effects in the Supersonic Mixing Layer," *AIAA Journal*, Vol. 30, No. 4, April 1992, pp. 973-981.

Gutmark, E., Schadow, K. C. and Wilson, K. J., "Mixing Enhancement in Coaxial Supersonic Jets," AIAA Paper 89-1812, June 1989.

Krothapalli, A., Buzyna G. and Lourenco, L., "Streamwise Vortices in an Underexpanded Axisymmetric Jet," *Physics of Fluids A*, Vol. 3, 1991, pp. 1848.

Krothapalli, A., Washington, D. and Alvi, F. S., " Aerodynamic and Acoustic Characteristics of a Supersonic Diamond-Shaped Jet," First Joint CEAS/AIAA Aeroacoustics Conference, Munich, 1995

Monkewitz, P. A., Lehmann, B., Bariskow, B. and Bechert, D. W. "The Spreading of Self-excited Hot Jets by Side Jets," *Physics of Fluids A*, Vol. 1, 1989, pp. 446-448.

Pannu, S. S. and Johannesen, N.H., "The Structure of Jets from Notched Nozzles," *Journal of Fluid Mechanics*, Vol. 74, 1976, pp. 515-528.

Papamoschou, D. and Roshko, A. R., "The Compressible Turbulent Shear layer; An Experimental Study," *Journal of Fluid Mechanics*, Vol. 197, 1988, pp. 453-477.

Shau, Y. R., Dolling, D. S. and Choi, K. Y., "Organized Structure in a Compressible Turbulent Shear Layer,"*AIAA Journal*, Vol. 31, No. 8, Aug. 1993, pp. 1398-1405.

Strykowski, P. J. and Krothapalli, A., "The Countercurrent Mixing Layer: Strategies for Shear-Layer Control," AIAA Paper 93-3260, July 1993.

Strykowski, P. J. and Krothapalli, A., "The Effect of Counterflow on the Development of Compressible Shear Layers" submitted to the *Journal of Fluid Mechanics*, 1995.

Figure 1

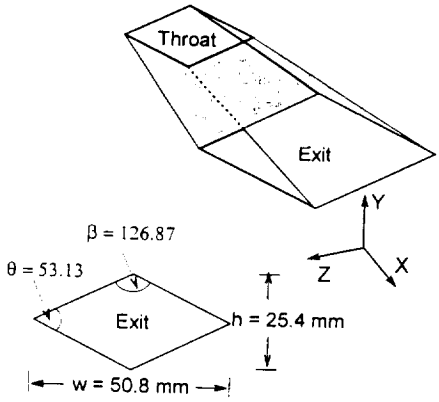


Figure 2

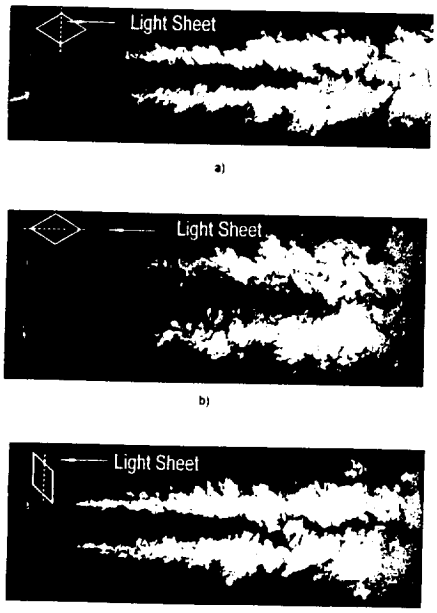


Figure 3

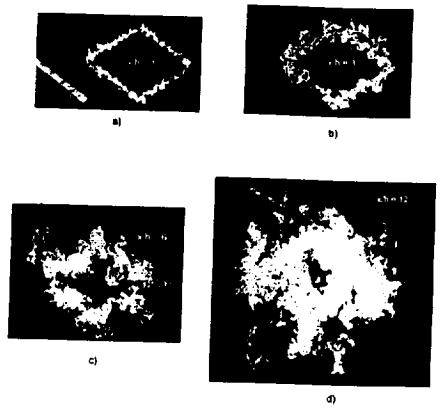


Figure 4

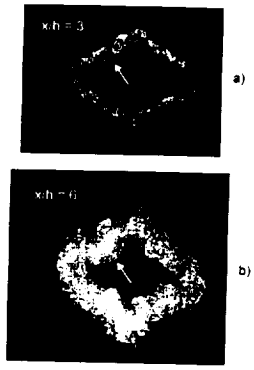


Figure 5

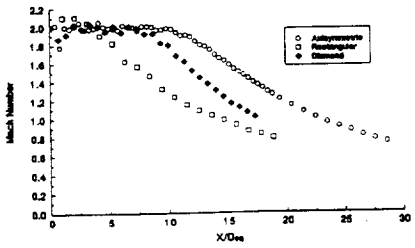


Figure 7

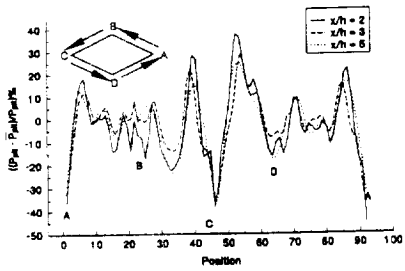
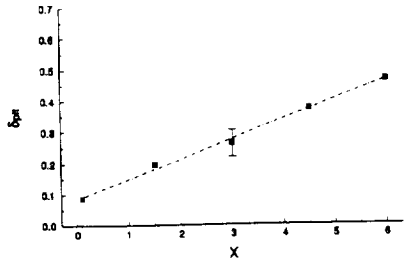


Figure 6



Condensate Accretion in Shock Tube's expansion Fan

Ephrem-Denis Mezonlin*, U. P. De Silva, F. Hunte and J. A. Johnson III
Center for Nonlinear and Nonequilibrium Aeroscience (CeNNAs)
Florida Agricultural and Mechanical University
Tallahassee, Florida

It has been shown that turbulence and temperature influence the droplet sizes in expansion fan induced condensation by studying the Rayleigh scattering from one port in our shock tube's test section. We have modified our set-up so as to allow, using two ports, the real time measurement of the influence of turbulence and temperature on the rate at which these droplets grow. To do this, we looked at the Rayleigh scattering from two different ports for ten Reynolds numbers at five different temperatures. We modeled the time of flight of droplets, using the equations of one-dimensional gas dynamics and the measured shock wave speed in shock tube's driven section.

Introduction

The condensation phenomena occur in many branches of sciences, in meteorology, geochemistry, biology, aerodynamics etc. For example, heterogeneous nucleation and the subsequent droplet and crystal growth play a critical role in the contribution from PCS and vapor trails to atmospheric environmental concerns and in the development of ice crystals on environmentally exposed aerofoils. Although in literature one could find several works on the theory of condensation (Lothe and Pound[1,2] Dunning[3] and Wegener[4]), no effort has been made to study the droplet growth, specially with taking in to account turbulence parameters. Among the recent experiments, work carried out by Britan et al.[5] and Schnerr and Bohning[6] measured several parameters in condensation successfully. In their work recently, U. P. De Silva, A. Gardner and J. A. Johnson III[7] have experimentally shown that the droplet size at a shock tube's contact surface significantly varies with the variation of Reynolds number of the flow, which is directly proportionate to the turbulent intensity. In addition, they also observed a pronounced dependence of the droplet size on the local temperature, independent of the Reynolds number variations; this was accomplished by observing the particle size changes at one optical port for different Reynolds numbers at three different temperatures. In this work, we take advantage of our existing capabilities and facilities for shock tube flow analyses and control; specifically, we used our 5'' inner diameter pressure ruptured shock tube, so that measurements could be made on the growth of droplets in real time at the shock tube's expansion fan. For this, we modified our driver-section-test-chamber insofar as to afford humidity and temperature control. On the test section, we add one more window, and by doing so, we were able to observe the

* Ephrem-Denis Mezonlin received his BS in physics from Université Nationale du Benin (UNB) in 1991, and his MS degree from Institute of Mathematics and Physical Sciences (IMSP) of UNB in 1993. He plans to graduate with his Ph.D. some time in 1998.

droplet at two different optical ports. With two pressure transducers set on the test section 2 feet apart, we could determine the shock wave velocity, and subsequently the droplet growth. We collected experimental data over ten different Reynolds numbers at five different temperatures.

Experimental set-up and procedures

We used a pressure driven shock tube (figure 1), consisting of a driver chamber of 60"x 5.2", a driven chamber containing a test section of 60"x5.2", and an extension tube of 76"x5.2". Aluminized Mylar sheets of two different thicknesses were used as the diaphragm-material which separates the driven chamber from the driver chamber. Compressed air is supplied to the driver chamber from a reservoir, where the relative humidity of the air could be controlled by setting the partial pressure in the reservoir (at 70°F). A heating tape is wrapped around the driver chamber, and a thermocouple probe is used in the driver chamber to measure the initial temperatures. The test section has three optical stations for diagnostics, two of which consist of four perpendicular glass ports, the other has two in-line glass ports. We used a 18 Watt, CW Argon Ion laser, running in the single-line, TEM_{00} mode, for the right angle scattering through two of these three optical stations. Two pressure transducers are mounted on either side of the optical station and the pressure traces and scatter signals at the right angle are recorded on a four channel, 500-MHz, Tektronix digitizing oscilloscope. All the signals are stored on floppy disks for further analysis. The receiving optics, including a natural density filter, transmit the received scatter signal from the water droplets to the photo multiplier tubes using fiber optics and also store the data on the same oscilloscope.

According to the references 7 and 8, at shock wave Mach numbers above 1.7, the contact surface demonstrates fully compressible turbulent behavior. Hence, to assure a fully compressible environment, the whole experiment is conducted at Mach number of 1.8. In order to achieve such condition, we maintain the shock tube pressure ratio (P_4/P_1) constant and introduce variations of Reynolds number by changing P_1 and P_4 individually while maintaining the same pressure ratio (18.5). After each run, we measure the shock wave velocity using the quartz pressure transducers. The scatter signals from the two photo multipliers are converted to the particle size by using Rayleigh's scattering theory[9-11].

Results and data analyses

We make Rayleigh scattering measurements, using our Argon Ion laser and appropriate analytical and calibration procedures, so as to provide continuous recordings of droplet size and density at the measuring stations. Off-line reconstructions are performed, using appropriate computer software and hardware, so as to provide histories in the droplet frame of reference using the velocity calculated from the assumptions of one-dimensional shock tube flow[12] and confirmed by direct LDV. From these results, we obtain droplet growth rate data and a droplet growth rate model which will determine the roles of temperature and suspensions in the evolution of droplet.

Influence of Reynolds number on the droplet size (figure 2) presents two dominant modes. This behavior was suggested by Tsuge[13]. In figure 3, we showed the influence of the temperature on the droplet size; it is also obvious that at a given Reynolds number the curve looks almost flat. Therefore, it is evident from figure 3 that the effect of temperature on droplet size is not as strongly marked as the influence of the Reynolds number. This behavior is clearly seen when we start to examine the droplet growth (figure 4) in space; here, the droplet growth is measured as the change in size of from one optical port to the

next port that is 12" away from the first port.

Conclusions

We have witnessed a new clear dependence of droplet size on the strength of turbulence. As one can see in figure 2, increasing Reynolds number produces an increase in relative sizes of the droplets that varies with temperature as well. But the rate of droplet growth measurements clearly show that Reynolds number affects the droplet growth to a much greater extent than the temperature.

Acknowledgements

This work was partially funded by NASA grant NAG-2930 to Florida A & M University.

References

1. Lothe, J., and Pound, G. M.," Reconsiderations of Nucleation theory".Journal of Chemical Physics, Vol. 36, No. 8 1962, pp. 2080-2085.

2. Lothe, J., and Pound, G. M.," Statistical Mechanics of Nucleation ." *Nucleation*, edited by A. C. Zettlemoyer, Marcel Dekker, New York, 1969, Chap.3.

3. Dunning, W. J. "General and Theoretical Introduction," *Nucleation* edited by A. C. Zettlemoyer, Marcel Dekker, New York, 1969, Chap.1

4. Wegener, P. P. "Gas Dynamics of expansion Flows with Condensation and Homogeneous Nucleation of Water vapor." *Nonequilibrium Flows*, Pt.1, Vol.1, edited by P. P., and Wegener, Marcel Dokker, New York, 1970, Chap.4.

5. Britan, A. B., Zuev, A. P., and Khmelevski, A. N., " The effect of Condensation on the Parameters of the Flow behind Shock Waves in Water Vapor," Zhurnal Tekhnicheskoi Fiziki, Vol. 62, Jully 1992, pp. 765-770.

6. Schnerr, G. H., and Bohning , R., "Compressible Turbulent Boundary Layers with Heat Addition by Homogeneous Condensation." AIAA Journal, Vol. 30, No. 5, 1992, pp. 1284-1289.

7. De Silva, U. P., Gardner , A., and Johnson, J. A. III, "Turbulent Manipulation of Condensation at a Shock Tube's Contact Surface" AIAA Journal, Vol. 33 . No. 2, 1995, pp. 368-370.

8. Johnson, J. A. III, and Ramaiah, R. I. L., "Reduced Molecular Chaos and Flow Instability" Stability in the Mechanics of Continua, edited by F. H. Schroeder, Springer-Verlag, Berlin, 1982, pp. 318-329

9. Stein, G. D., and Wegener, P. P. "Experiments on the Number of Particles formed by Homogeneous Nucleation in the vapor Phase" Journal of Chemical Physics, Vol. 46, 1967, pp. 3685-3686.

10. Kerker, M., *The Scattering of Light*, Academic Press, New York, 1969, Chap. 3

11. Shirinzadeh, B., Hillard, M. E., and Exton, R. J. " Condensation Effect on Rayleigh Scattering of Light Measurements in a Supersonic Wind Tunnel," AIAA Journal, Vol. 29, No. 2 1991, pp. 242-246.

12. Keenan, J. H., Chao, J., and Kaye, J., Gas Tables- Thermodynamic Properties of Air, Products of Combustion and Component Gases, Compressible Flow Function, 2nd Ed., Weiley, New York 1980.

13. Tsuge,S. "Approch to the Origin of Turbulence on the Basis of Two-Point Kinetic Theoty, " Physics of Fluifs, Vol. 17, Jan. 1974, pp. 22-33.

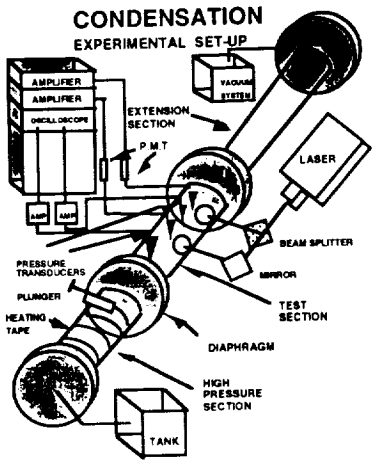


Figure 1. Diagram of the shock-tube and the experimental set-up (not to scale).

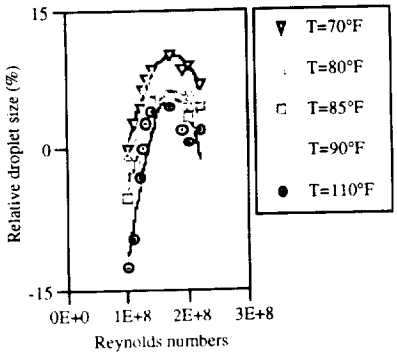


Figure 2. The effect of Reynolds number (turbulent intensity) on the droplet size.

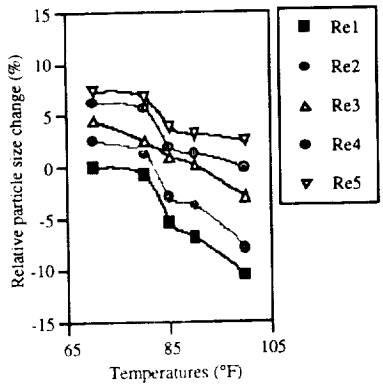


Figure 3. Effect of the starting temperature on relative droplet size at various Reynolds numbers.

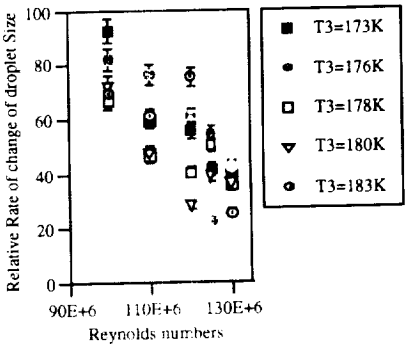


Figure 4. Effect of the turbulent intensity on the growth of droplet size in space.

TURBULENCE AND ICING PROCESSES IN SUBSONIC FLOW

Frantz Camulaire[1] , Upul P. de Silva and Joseph A. Johnson III

Laboratory for Modern Fluid Physics, CeNNAs, Florida A & M University, Tallahassee, FL 32310, USA

ABSTRACT There is a rapidly growing interest in the science and technology of ice accretion processes. These processes are particularly critical on the outer surfaces of an aircraft in flight; here the possibility of icing control is of utmost importance to the global aircraft industry and the associated regulatory agencies so that the effectiveness and reliability of anti-ice and deicing equipment can be improved. The three major areas of interest in ice accretion conditions on an aircraft are ground icing, inflight icing, and tail plane icing, which correspondingly support aircraft flight profiles of: takeoff; cruise and hold; and approach and land. Furthermore, there has been very little quantitative attention paid to the fundamental role of turbulence physics in icing processes. There is also great potential value indeed in icing experiments that simultaneously utilize more than one diagnostic technique, which then would independently measure the role of turbulence on relevant local quantities underlying the ice accretion processes. We propose such a series of experiments, exploiting the state-of-the-art diagnostic tools already available to us at LMFP and CeNNAs and the considerable experience we possess from similar experimentation in high speed flow environments in measuring turbulence distortion of condensate accretion rates and in characterizing turbulent parameters in nonequilibrium flow.

1. INTRODUCTION

There is a rapidly growing desire in science and technology to understand the physics of icing phenomena, especially during ice accretion processes. Since detailed experimentation on ice accretion is impossible to conduct under natural (flight) icing conditions, it is very important to understand scaling requirements and isolate similarities and differences between artificial (laboratory) icing and the flight icing environment. In addition, reliable data from laboratory experiments are required to verify the various computational models suggested by researchers (as, e.g., in Kirby & Hansmann 1988).

1.1. Historical background and motivation

Just during the last 15 years (1980-1995), there have been eight examples of aircraft catastrophes that involve or are attributed to ice accretion (Bergrun 1995). Therefore, it is of utmost importance to the global aircraft industry and its regulatory agencies to understand the ice accretion process well enough such that reliability and effectiveness of aircraft de-icing and anti-icing equipment could be enhanced. The

[1] Frantz Camulaire graduated with BS in Mechanical Engineering from The City College of the CUNY in 1995. He is currently persuing a MS in Mechanical Engineering at FAMU/ FSU College of Engineering. Frantz anticipates on graduating in 1997 and continue onto doctoral studies in Fluid Mechanics. He has presented papers at various conferences. Frantz presently works as a research assistant at CeNNAs and has research focuses on Turbulence related processes involving Aircraft Icing and Internal Combustion Engines.

three major areas of interest in ice accretion conditions on an air-craft are ground icing, inflight icing, and tail plane icing which correspondingly support aircraft flight profiles of: takeoff; cruise and hold; and approach and land (Reinmann 1994). From the limited number of experiments conducted so far, it has been shown that the effect of icing on various aerodynamic surfaces result in: reduction of flight control effectiveness by 8-12% (Ratavsky & Ranaudo 1993); sharp increase in drag (Shin & Bond 1992); reduced maximum lift coefficient; excessive vibration; and probable premature stall (Kirby & Hansmann 1988) in all stages of flight, -any of which could lead to disastrous results. As a complement to the experiments, there has been a fairly large number of numerical investigations involving ice shape, size, and effect of turbulence on accreting ice (Shin & Bond 1992; Potapzuk 1993, Bidwell 1993, etc.). In past, the majority of experiments carried out on ice accretion has been one-parameter studies, i.e., relying predominantly on one diagnostic technique at a time (e.g. ultrasonic pulse-echo - Kirby & Hansmann 1988; LDV - Bragg & Kerho- 1994; light scattering - Oldenberg & Ide 1990, etc.). Furthermore, there has been very little quantitative attention paid to the role of turbulence in the icing processes except for the work by Bragg & Kerho (1994). Thus, there is great potential value indeed in the icing experiments that simultaneously utilize more than one diagnostic technique, which then would independently measure the role of turbulence on relevant local quantities underlying the accretion process. We propose such a series of experiments, exploiting the state-of-the-art diagnostic tools that are readily available to us at the Florida A&M University (FAMU) Laboratory for Modern Fluid Physics (LMFP) and the NASA Center for Nonlinear and Nonequilibrium Aeroscience (CeNNAs) at FAMU and the considerable experience that we posses from similar experimentation in high speed flow environments in measuring water droplet sizes (De Silva et al 1995) and in characterizing and visualizing turbulent parameters (Johnson et al 1993; De Silva et al. 1996 and references therein).

1.2. Physics of ice accretion

When an aircraft encounters liquid water in the form of super cooled cloud droplets, or freezing rain, ice will form some place on the aircraft's exposed surfaces. The typical diameter of droplet in such a cloud could vary from less than 10 μm to over 50 μm. If the encountered water source is freezing rain, droplets could be several millimeters in diameter. Therefore, the shape of the ice accreted on the aerodynamic surfaces of the aircraft and their influence on the performance of the aircraft could depend on a number of parameters: cloud temperature; the average cloud temperature and the size-spectrum; the amount of liquid water per unit volume contained in the cloud; size, shape and airspeed of the accreting body. Two distinct icing regimes have been identified: rime - cold droplet clouds with temperatures below -10°C and small droplet sizes that freezes on impact with the accreting surface and typically protrude towards the airflow; and glaze-clouds with warmer

temperatures of 0°-10°C which would not freeze on impact, but instead coalesce and run back over the accreting surface and freeze with two pronounced growth peaks, or horns, on either side of the body stagnation line (Kirby & Hansmann 1988).

New turbulence physics based on modified kinetic theory (Tsuge & Sagara 1976 and references therein) has provided profound new insights on the role of turbulence in nonequilibrium processes. It has successfully predicted (Johnson & Chen 1978) the first experimental proof of the existence of coupling between turbulence and a molecular reaction rate (Johnson,Johnson, and Lu 1990). It has led to the first theoretical determination of a connection between entropy evolution, turbulence, and chemical nonequilibrium (Orou and Johnson 1994). More recently, it has motivated the first experimental evidence of a connection between turbulence and the evolution of droplet sizes in flow with condensation (De Silva, Gardner and Johnson 1995) and the first suggestion of natural turbulent closure with the vanishing triple correlation in compressible flow (De Silva et al., 1996). We have determined a theoretical model for condensation which uses Tsuge's approach (Hazoume and Johnson 1996) and also a very first indication that the theoretical formulation can predict the evolution of droplet sizes during condensation (Mezonlin et al., 1996). It remains to explore this approach and its derivative formulations for icing processes.

1.3 NASA activities on understanding physics of ice accretion

In 1928, the National Advisory Committee for Aeronautics (NACA, now NASA) initiated a fundamental research program to define the problem of aircraft icing and also to understand and investigate the means for preventing or reducing formation of ice that could cause dangerous or inefficient flight conditions (Bergrun 1995). Since then, NASA has always vigorously taken part in investigating and researching icing process under both artificial and natural icing conditions. The NASA Lewis Icing Research Tunnel (IRT) has been extensively used in such investigations with NASA Langley Low Turbulence Pressure Tunnel (LTPT) being used for certain calibration and model development work. Computationally, the variations of LEWICE code, developed by scientists affiliated with NASA, is the widely accepted code for simulating accretion process. One of the latest efforts in the NASA activities aimed at solving the ice accretion problem was the establishment of a long-term collaboration between McDonnell Douglas/ NASA Lewis/ NASA Langley in 1993 to improve understanding the ice accretion effects on multi-element airfoils (Shin et al. 1994).

1.4. Comparison between Theoretical and Experimental results

There are many different types of turbulence models that offers numerical solutions to compressible flows over aerodynamic components. Van Driest formulated a model for smooth wall geometries in which the turbulent shear stress near the wall

may be expressed as function of the mixing length :

$$\tau_t = \rho \cdot F^2 \cdot l_m^2 \cdot \left[\left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial w}{\partial y} \right)^2 \right]$$

where the mixing length after Michel et al. is

$$l_m = \delta \cdot k_r \cdot \tanh\left(\frac{k}{k_r} \cdot \frac{y}{\delta} \right)$$

with kr=0.085 and k=0.41 and

$$F = 1 - \exp\left(\frac{l_m \sqrt{\tau \cdot \rho}}{26 \cdot k \cdot \mu} \right)$$

where F is the turbulence damping factor near a smooth wall. The roughness effect is simulated by modifying F so as to annihilate its damping effect and increase the turbulent shear stress. Using the Van Driest roughness model F the becomes

$$F = 1 - \exp\left(\frac{l_m \sqrt{\tau \cdot \rho}}{26 \cdot k \cdot \mu} \right) + \exp\left(\frac{60 \cdot y}{26 \cdot k_s} \right)$$

where ks is the equivalent sand grain roughness.

Cebeci-Smith have also formulated a turbulence model which is based on an algebraic eddy-viscosity formulation. The mixing length expression in the inner eddy viscosity formula using roughness effects is expressed as:

$$l = 0.40(y + \Delta y)\left[1 - \exp\left[\frac{-(y + \Delta y)}{A} \right] \right]$$

where Δy is a function of an equivalent sand grain roughness ks.

In our efforts we will be able to evaluate most of the computational work that had been produced by NASA research centers and by other researchers in the recent past on the ice accretion phenomena. In addition, our simultaneous measurements of ice particle growth rate and shape-change using complementary diagnostic techniques will provide conclusive real-time parameters on icing phenomena. We propose to develop a LMFP Icing Physics Facility (IPF) which will allow us to measure all necessary icing parameters on a scaled down version of any aerodynamic model under laboratory conditions closely resembling the natural conditions. The IPF will provide a simple icing research facility which could be operated to provide multiple-runs in a cost effective, efficient manner thereby also establishing the reproducibility of all the measurements that IPF produce. Our past experience and expertise in measuring, as well as describing, turbulence in high speed flows will enable us to investigate in detail the role of turbulence in the ice accretion process as well as the onset of icing (in case of super cooled water).

2. EXPERIMENTAL FACILITIES, PROCEDURES AND TECHNIQUES

The proposed LMFP IPF will be used as the facility for our suggested icing experiments. The whole line of diagnostic techniques described below would be utilized in IPF.

2.1. LMFP Icing Physics Facility (IPF)

This icing research tunnel will have a variable temperature control from -22°F to approximately 70°F over a continuous velocity range of 150 to 400 mph. Spray droplets are provided by

precision, easily controlled, ultrasonic nozzles.

2.1.1. Wind Tunnel

The wind tunnel is a closed loop, continuous flow, epoxy coated, steel wind tunnel with approximately 10'x5' overall dimensions. Test section is 12" long with 4 square inch cross section, heated glass windows in three stations spaced by 4" apart. Flanged wind tunnel components are easily assembled and disassembled. When iced, the first corner will be easily removed and de-iced prior to continuing testing. However, according to the preliminary calculations, the tunnel would have to run for several hours, before a significant ice-build-up occurs on any internal components of the tunnel. It will have insulated sections, with corner turning vanes provided for high-quality flow. Especially, the first vane would have thick vanes with a lower droplet collection efficiency. We would provide heated glass viewing windows for undisrupted optical diagnostics throughout the operation of the tunnel. The tunnel also consists of: an integrated spray duct section which allows nozzles to be changed for varying droplet sizes; an integrated, interchangeable, duct section that allows easy addition and removal of screens and honeycomb structures for high quality aerodynamic testing; a flanged and hinged test section that allows easy access to the test models and the interior; static ports and total pressure probe with a high quality differential pressure transducer that enables tunnel calibration; and a support pallet for the tunnel sections.

2.1.2. Refrigeration System

We use a self contained brine, dual loop, refrigeration system which would enable a wide range of heat load conditions, i.e., approximately 3 tons of cooling across a range of -22°F-70°F. This system would also contain: a closed loop temperature control with stagnation chamber RTD feedback; compact, insulated, brine storage tank; and approximately 2'x2' ($4ft^2$) heat exchanger that will serve as a temperature and flow conditioner. The whole unit is palletized such that it could be removed from the wind tunnel via quick disconnects.

2.1.3. Drive System

The whole system would be driven by a pallet mounted blower, motor, and variable frequency drive system. The fan system will provide smooth velocity transition over a wide range from 150 mph to 400 mph; fan is powered by a 15 hp electric motor. The palletization of the system provides easy removal and installation from the system.

2.1.4. Water Spray System

The water spray system, which is one of the most important aspects of any icing wind tunnel, will provide properly filtered and de-ionized water in the desired droplet size and the tunnel liquid water content in our IPF. The water spray section is flanged and easily removed thereby allowing an aerodynamic flow conditioner section consisting of screens and honeycombs to be inserted for aerodynamic testing without spray if necessary. The de-

ionization and filtration of water will eliminate the possibility of unreliable data due to contaminated water. The droplet size is maintained constant by using ultrasonic nozzles which are readily interchangeable. Three droplet sizes are initially available at 15, 25 and 40 microns with additional particle sizes achieved if necessary by purchasing and installing similar nozzles of required particle-sizes.

2.1.5. Overall Control

The wind tunnel speed is changed and maintained via the variable speed VFD. The wind tunnel velocity will be set via a knob on the VFD panel. The VFD unit can receive and be controlled by a computer using a IEEE-488 interface. As mentioned earlier, the differential pressure transducer integrated with the static ports and the total pressure probe will provide the velocity readout. The refrigeration system has a self contained, programmable, single loop process controller. An RTD will provide the temperature feed back. Spray water pressure and flowrate will be supplied via gauges. Control will be manual.

2.2. Experimental Methods and Diagnostic Techniques

A typical experiment will be performed in the following sequence. First the proper water spray nozzle is installed in the spray unit and the target (scaled-down) aerodynamic model is installed in the test section, the IPF is then run until the required temperature and velocity are reached in the test section. After this, our multi-faceted diagnostic techniques are used in any necessary combination on points of a predetermined measuring grid through the optical ports at the observation stations in the test section. For example, if we are carrying out LDV measurements at points of a grid, we mount the LDV assembly on our computer controlled Ampro, heavy-duty, x-y-z traverse which routinely provides us very high spatial resolution.

A variety of diagnostic techniques are used to analyze the important parameters of the ice accretion processes and also independent records of flow velocities.

a. Growth rate Rayleigh/Mie scattering

This method was developed at LMFP and successfully utilized to measure particle-sizes of condensed water droplets traveling with the contact surface in a shock-tube (De Silva et al. 1995). When the IPF is operating at the desired parameters scattering signals will be obtained at different ports (different points in space) and at different angles at one station (different scatter angles from the same point in space) which will eventually provide us with two separate measurements of time-histories of the size of ice particles at the same point in space, and also provide the rate-of-change of ice particle size in space.

b. Videometric Schlieren Shadowgraphy

Our Xybion-ISG-350 high-speed CCD camera utilizing the state-of-the-art Videometric Analysis System will be used in conjunction with a laser sheet schlieren system to image multiple images of the growing ice particles onto the same frame. With controlled

time intervals between each exposure a well documented history of evolution of ice particles will be recorded. In addition, the process could be recorded as a continuous video signal and observed later to qualitatively appreciate the accretion process.

c. Optical Array Probe (OAS) Diagnostics

Here, a Ar+ CW laser beam would be projected across the open space (where we are interested in measuring the particle size of ice) between two probe arms, magnified by a set of lenses, and projected onto a 24-element linear photodiode array. Droplets crossing the laser beam will shadow one or more of the photodiode elements. The droplet size is determined by the number of photodiode elements shadowed; the resolution could be controlled by the coefficient of magnification of the lens system (Oldenburg & Ide 1990).

d. Phase Doppler Particle Analyzer (PDPA)

We would use the signal obtained from our 3-D LDV onto 3-different photodetectors and use the phase shift between the collected doppler signals from different detectors to calculate the particle size independently.

e. Hot-Wire Anemometry

We will use our TSI hot-wire anemometer with a 3-D probe to independently measure the velocities at the test section.

f. Laser Doppler Velocimetry (LDV)

We will also use our 3-D LDV system as a anemometer, to independently record velocity, thereby giving us three independent means of velocity measurement (differential pressure transducer, hot-wire anemometer, and Laser Doppler anemometer) of the flow.

g. Laser Induced Fluorescence (LIF)

If required, we would use our LIF and PLIF techniques which could also image and record spatial and temporal density fluctuations (when M>0.3) and will also provide 3-D velocity in conjunction with the newly developed Direct Estimate Velocimetry (DEV) (De Silva et al. 1996).

h. Shadowgraphy

In addition, ordinary shadowgraphy will be used to qualitatively understand the flow during flow conditions of M>0.3.

2.3. Data Analysis

In addition to analyzing the data to establish parameters related to icing particles, we will also subject the velocity data to a complete turbulence analysis. In order to find out any dependency of icing particles and the ice accretion process in general on turbulence the data will be collected at several different turbulence intensities. A detailed analysis of turbulence will be carried out including spectral. correlation and fractal analysis. Relationship between the turbulent parameters obtained after subjecting data to various analyses and the parameters critical in

describing the icing accretion (size, shape, growth rate, etc.) will be investigated.

3. PROPOSED ACTIVITIES

As indicated above, following activities will be carried cut under this proposed research:

1. A working closed loop icing wind tunnel (IPF) would be constructed with the ability of achieving temperatures down to -22°F with all the necessary controls and a velocity range of 150-400 mph.

2. Using the particle sizing techniques of Rayleigh/Mie angle scattering, Phase Doppler Particle analyzing, Optical Array Probe Diagnostics and Videometric

schlieren shadowgraphy the ice particle sizes will be measured.

3. Ice particle growth in time and in space will be calculated using time sequenced data.

4. The dependence of particle shape and size on the following parameters would be studied: influence of water droplet diameter; water number density in the flow; water temperature; flow velocity; flow temperature; and aerodynamic body shape.

5. The influence of turbulence on icing particle size and shape would be investigated in detail.

6. The flow velocities will be measured by three independent methods namely: 3-D LDV; Differential Pressure transducer coupled with static port and total pressure probe; and hot-wire-anemometer system with a 3-D probe.

7. If necessary, LIF technique in conjunction with the Direct Estimate Velocimetry (DEV) will be used to analyze the flow.

8. Obtained results will be compared with results from computational models by other researchers.

9. Comparison will also be made between our experimental data and existing flight icing data obtained by other researchers.

We expect to take approximately one year completing the facility, implementing calibration diagnostics, and developing the first set of data on the overall basic system. During the second year, we would expect first results on the role of turbulence physics in subsonic icing, the influence of droplets and icing on the local turbulence on aerofoils and first results on potentially portable diagnostics to the Icing Wind Tunnel Facility at Lewis. The third year would be focussed on reporting and validating the new insights derived on the topics outlined above and on developing, as appropriate, additional projects arising from the Icing Program needs at Nasa Lewis.

4. REFERENCES

Arrington, E. A., Pickett, M. T. & Sheldon, D. 1994 Flow quality studies of the NASA Lewis Research Center Icing Research Tunnel diffuser. NASA TM-106311

Bergrun, Norman 1995 A warming trend for icing research, Aerospace America, 8, 22-27

Bidwell, C. S. 1993 Ice accretion prediction for a typical commercial transport aircraft NASA TM-105976; also AIAA pap. 93-0174

Bragg, Michael B. & Kerho, Michael E. 1994 3-D LDV measurements on a 30-degree swept wing with a simulated ice accretion. NASA CR-195327

De Silva, U., Gardner, A. & Johnson III, J. A. 1995 Turbulent manipulation of condensation at a shock tube's contact surface. AIAA J. 33, 368-370

De Silva, U., Zhang, J., Orou, J. Chabi & Johnson III, J. A. 1996 First evidence of natural closure in turbulence. J. of Phys. D: appl. Phys,

Hazoume, R and Jonhson III, J. A. 1996 Condensate droplet growth with temperature fluctuations (Submitted to Phys. Letts.)

Johnson III, J. A., Johnson, L. E. & Zhang, J. 1993 An attractor driven approximation for turbulent burst dynamics in a supersonic shear layer. Proc. 19 th Intl. Symposium of Shock Tubes and Waves, Marseilles.

Johnson III, J. A., Johnson, L. E. & Lu, X. 1990 Turbulence in a reacting contact surface. Phys. Fluids A, 2, 2002

Kirby, Mark S. & Hansman, R. J. 1988 An experimental and theoretical study of the ice accretion process during artificial and natural icing conditions. NASA CR-182119

Mezonlin, E-D, De Silva, U. P., Hunte, F. and johnson III, J. A. 1996 Condensate accretion in a shock tube's expansion fan (Submitted to Applied Phys. Letts.)

Oldenberg, J. R. & Ide, R. F. 1990 Comparison of two-droplet sizing systems in an icing wind-tunnel NASA TM-102456

Potapzuk, Mark G. 1993 Navier-Stokes analysis of airfoils with leading edge ice accretions. NASA CR-191008

Ratavsky, T. P. & Ranaudo, R. J. 1993 Icing effects on aircraft stability and control determined from flight data preliminary results. NASA TM-105977; also AIAA paper 93-0398

Reinmann, John J. 1994 Icing: Accretion, detection, protection CASI accession number: 95N14897

Shin, J., Wilcox, P., Chin, V. & Sheldon, D. 1994 Icing test results on an advanced two-dimensional high-lift multi-element airfoil. NASA TM-106620; also AIAA pap. 94-1869

Shin, Jaiwon & Bond, Thomas H. 1992 Experimental and computational ice shapes and resulting drag increase for a NACA 0012 airfoil NASA TM-105743

Tsuge, S. and Sagara, K. 1976 Kinetic theory of turbulent compressible flow and comparison with classical theory. Phys. Fluids 19, 1478

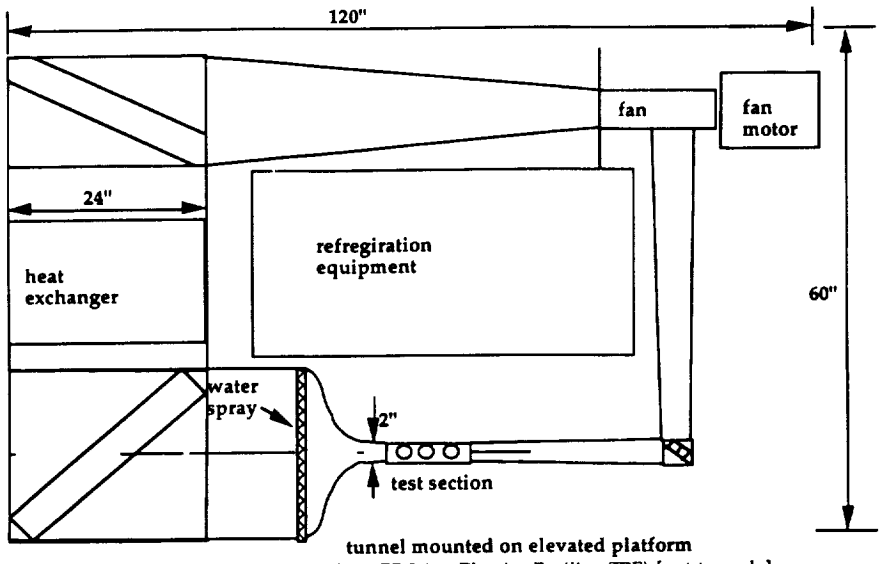


Figure 1. A sketch of the proposed LMFP Icing Physics Facility (IPF) [not to scale]

Turbulence in Weakly Magnetized Plasmas

F. Hunte[1], **R. Appartaim, J. A. Johnson III**
Florida A & M University, Tallahassee, Florida 32310

Abstract

An experiment has been designed which will allow for the study of turbulent effects on atomic and molecular processes in a weakly magnetized plasma environment. Spectral analysis of a turbulent argon plasma in an axial magnetic field will be performed at field strengths ranging from 0-0.2 Tesla. Correlations between emissions from atomic species and turbulence intensity indicators will be investigated in an effort to discover the effects of turbulence on reactive processes where they exist. This study also will provide insight into the heat and mass transport processes associated with collisional atomic and molecular processes.

1. Introduction

From the earliest history of gas discharge phenomena to the present, the behavior of ions and electrons has been widely studied and characterized. The instabilities and turbulent characteristics of plasmas has been treated for these particles [1]. What is lacking, however, is a treatment of the role of neutrals in collisional plasmas. Experimental evidence has already revealed a coupling between collisional dynamical processes and strong turbulence. Furthermore, local electric and magnetic field fluctuations associated with nonequilibrium molecular systems seem to cause high amplitude fluctuations and a manipulation of inherent complexity with broad spectral composition in parallel with (and perhaps independent of) the standard background turbulent behaviors. Therefore, the study of plasma turbulence in magnetized collisional plasma environments is of crucial importance in order to understand the coupling between turbulent fluctuations, anomalous transport and the distortion by turbulence of reactive processes.

2. Experiment

The emission spectra resulting from the excitation of neutrals in a low pressure glow discharge tube will be observed in the absence of a magnetic field. These observations will subsequently be compared to data obtained from emissions in the presence of magnetic fields of varying intensity. It is expected that such a comparison will reveal correlations between turbulent transport processes induced by the magnetic parameter and collisional processes which are the basis for interactions between neutral particles. It is further expected that this experiment will provide a basis for a series of experiments involving the injection of energetic neutrals into a plasma.

Included in this experiment is a mechanism by which the plasma will be driven to various states of turbulence which can be monitored and characterized. The focus of current investigation will be the effects of turbulence on transport and collisional processes which yield evidence of correlations between these phenomena. From these results, it is expected that it will be

[1] Frank Hunte graduated with his B. Sc. in physics from Florida A & M University in 1995. He is now studying for his Master's degree and plans to continue on to a doctoral degree in physics. Frank has participated in major scientific projects including the Superconducting Super Collider (1993) and the Princeton Plasma Physics Laboratory 's Tokamak Fusion Test Reactor (1995).

possible to deduce the implications of turbulence for reactions involving neutrals.

Experimental Setup

The discharge is created in a cylindrical Pyrex tube having a diameter of 2.5 cm and a fixed length of 60 cm. Two solenoidal field coils are mounted coaxial with the discharge tube such that an axial magnetic field (0-0.2 T) is produced in the positive column of the glow discharge. The discharge is initiated by a tungsten filament at the cathode powered by a 50V-50A DC power supply and sustained by a 2 kV DC power supply (Fig. 1).

The vacuum system currently consists of a VARIAN SD1400 mechanical pump capable of an ultimate pressure of $\sim 10^{-4}$ Torr. A turbomolecular pump will shortly be installed in order to attain lower pre-discharge pressures ($\sim 10^{-6}$ Torr). After the tube has been evacuated, it is filled to the desired pressure by means of a fine control needle valve. The gas pressure is monitored with a thermocouple gauge.

The collision induced light emission from neutrals is observed with a f/3.8 Jarrell-Ash, defraction grating spectrometer. The optical setup consists of two plano-convex lenses which collect the light from the discharge through a flat window on the positive column and focuses it onto the entrance slit of the spectrometer. The line of sight is set and constantly checked with a 632.8 nm He-Ne laser beam (not shown in the figure). The observations are currently limited to the spectral range of 3500Å-5000Å by the glass optics used to monitor light emissions from neutral atoms. The light source is 1 mm above the tip of the probes located in the positive column of the discharge.

3. Discussion

For an argon discharge at an average electron temperature T_e of 2eV, light emissions occur mostly in a spectral band ranging from 350Å to 5000Å [2]. Measurement of the absolute intensity of a line can be used to estimate the atomic species emitting the line. Doppler broadening measures directly the temperature of the atomic species.

The Schottky diffusion theory which ignores inertial effects is a collision dominated treatment that describes the positive column of a glow discharge at medium pressures (50 mTorr-10 Torr) relevant to this experiment [2]. The electrons in such a discharge will have a Maxwellian distribution of velocities such that only a small fraction at the tail end of the distribution will have enough energy to cause ionization from the ground state. Several excitations will be possible due to energy transfers by the other electrons [3]. Beam emission spectroscopy (spectroscopy on emissions from a neutral beam injected into a plasma column) will be performed. A second phase of experiments will introduce a beam of neutral atoms into the glow discharge. This tracer beam will allow spectral observations to be performed in highly localized spots in the positive column where certain effects of turbulence are known to exist. The results of these experiments should further characterize the correlations of turbulent interactions involving neutrals.

Magnetic Effects

The most significant effects of the magnetic field are twofold. It has been shown experimentally, and theoretically by Kadomstev, that a critical magnetic field B_C marks the onset of helical density oscillations in the positive column. As the axial field B_Z is increased above B_C, the fluctuation spectra of the helicon waves broaden [2], indicating a transition to turbulence. Fluctuations in the emission of light from neutrals are observed in parallel with this occurrence.

Further, an axial magnetic field inhibits lateral motion by way of diffusion or freefall. In this way, the confined plasma is prevented from losing energy to the walls. Reduced radial losses suggest that more energy is available within the discharge for collisional interactions between particles [3].

4. Summary

This experiment will be a precursor to a series of experiments using a glow discharge tube target for beams of photons, neutrals, ions and excited molecules in a weakly magnetized colliisional plasma. Diagnostic tools available at LMFP and CeNNAs and the considerable experience derived from similar experimentation in high density plasma flow environments in measuring the magnetic distortion of plasma turbulence and the role of such processes in the evolution of new diagnostic tools for high speed flow will be exploited.

References

[1] R. Ramaiah and J. A. Johnson III, Phys. Rev. A **36**, 774 (1987)

[2] A. Castellanos, 1994, Ph.D. Dissertation, City College, CUNY

[3] R. N. Franklin, *Plasma Phenomena in Gas Discharges* (Clarendon, Oxford, 1976), pp. 38-48

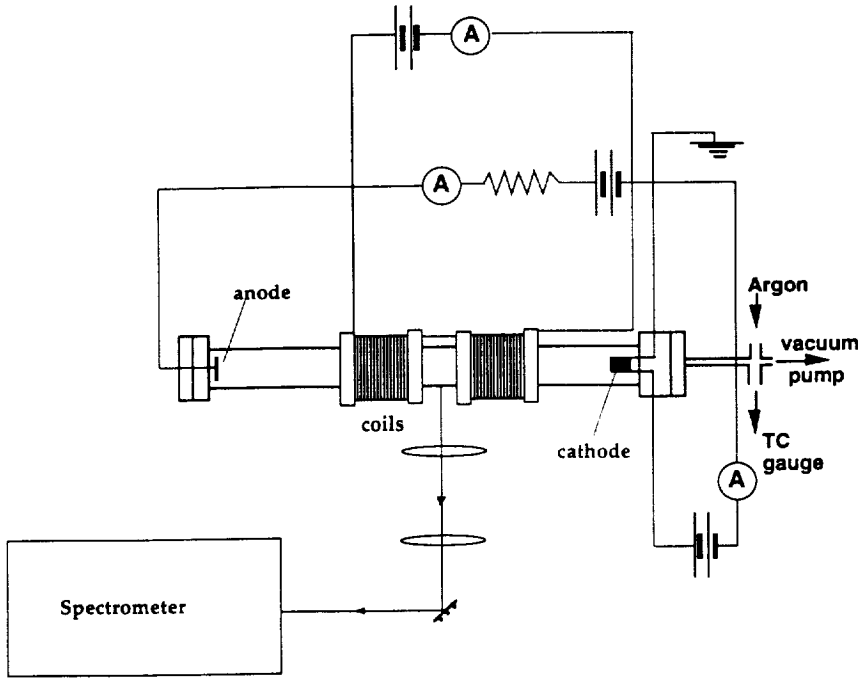


FIGURE 1. Schematic arrangement of apparatus for glow discharge in a magnetic field.

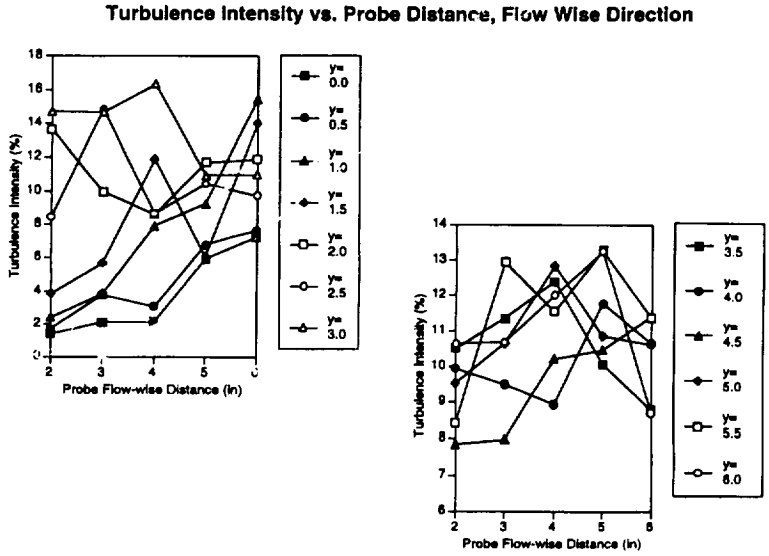


figure 5. Turbulent intensities along the streamwise direction at various transverse locations calculated from the measured velocity data.

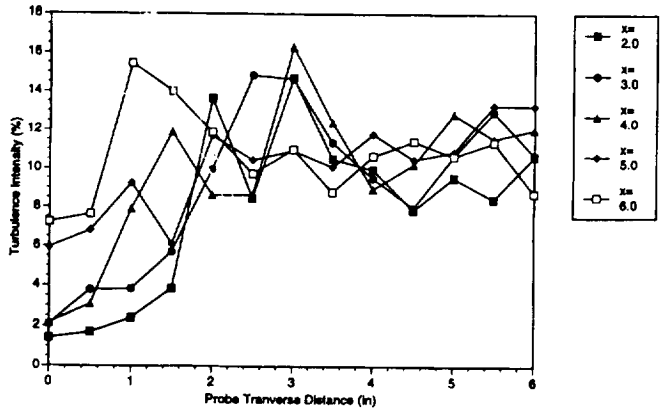


figure 6. Turbulent intensities along the spanwise direction at various streamwise locations calculated from the measured velocity data.

569-34

290733 P5

Calibration of Direct Velocimetry using Hot Wire Anemometry

Michael L. Holland*, U. P. De Silva, J. A. Johnson III

Center for Nonlinear and Nonequilibrium Aeroscience (CeNNAs)
Florida Agricultural and Mechanical University
Tallahassee, Florida

We studied the functionality of the IFA300 Hot Wore Anemometry system and its feasibility for the calibration of Direct Estimator Velocimetry (DEV) using Laser Induced Florescence (LIF). The experimental setup consisted of a single-wire hot wire probe measuring one component of velocity, the anemometry computational hardware and software, an HP oscilloscope, and a table fan to produce a simple flow with fluctuations. Measurements were taken at several points in the stream wise and transverse directions in the flow and various parameters recorded such as mean velocities, temperature, turbulence intensities, skewness coefficients and flatness coefficients. The IFA300 software also allowed us to perform statistical manipulations such as spectrum analysis on velocities samples and correlation. Utilization of data files, also produced by the anemometry software, and post analysis were performed to produce graphic al representations of turbulent intensity versus probe position and a flow field velocity profile. We concluded that the IFA300 Hot Wire Anemometry system is a reliable and functional method for calibration of DEV using LIF. Our future intentions are to set up a test chamber such that both velocity measurement techniques can be applied simultaneously, thus the calibration.

Introduction

There exist a number of techniques for measuring various parameters of fluids, such as velocity, temperature, pressure, and others. Flow diagnostic techniques, as they are known, may be categories in several general areas, two[1] of which are intrusive and optical (non intrusive). Non intrusive measurement techniques may be subdivided into two general methods, visualization (qualitative) or measurement (quantitative). The visualization methods include such techniques as shadowgraphy, Schlieren, and others while the measurement techniques include such techniques as Laser Doppler Velocimetry, Laser Induced Florescence (Direct Estimator Velocimetry), Particle Image Velocimetry and Laser Spectroscopy. Intrusive methods include Pitot Tube (Pitot Static), Hot Wire Anemometry and others.

In using these different fluid diagnostic techniques, the calibration of the method or techniques is crucial to its validity in research. Such a situation was present for the Direct Estimator Velocimetry (DEV), which is a unique method of converting Laser Induced Fluorescence (LIF) into velocities. This was the intent of our research, to calibrate the Direct Estimator Velocimetry using Hot Wire Anemometry, subsequent to the evaluation of the feasibility of the Hot Wire Anemometry as a calibration technique.

Background

Hot Wire Anemometry is a technique for measuring fluid velocity that has been used since the early 1800's. During early use they were typically made by hand in the laboratories in which they were intended to be used. At that time the problem with anemometers was that they had a low sampling rate, however, due to

* Michael L. Holland is a senior at Florida A&M University pursuing his B.S in Mechanical Engineering. He will complete his undergraduate studies in April 1996. He plans to continue his education for a M.S. degree and subsequently a Ph.D., also in Mechanical Engineering.

advances in technology, has been greatly improved by changes in materials and computer aided operations. Advances in technology and uses for hot wire anemometry, even though still named hot 'wire anemometry', have made this name obsolete. Not only are hot wires used but also hot films, or a combination of wire and film. Describing the system as anemometry is slightly improper as well, because the word anemometry means measuring air velocity, however, other fluids velocities can also be measured.

All hot wire anemometry systems have the same components, a probe and its cables and its electronics package. Generally there are three types of hot wire probes, constant temperature, constant current or pulsed wire probe. The types most commonly used are constant temperature and constant current. Constant temperature probes functions such that as the wire is cooled by convection in the flow the current is adjusted to maintain the constant temperature across the wire. From this compensating current the velocity of the flow can be measured. Similarly with the constant current probe the temperature of the probe adjust to an equilibrium temperature to compensate for the heat lost do to the convective flow. This equilibrium temperature is a function of the fluid velocity.

Experimental Setup and Procedure:

To determine if the hot wire anemometry was feasible for the calibration of the Laser Induced Fluorescence consisted of obtaining a hot wire anemometry system and measuring simple flow and noting the systems abilities to measure critical flow parameters. Our system was a TSI IFA 300 Constant Temperature Hot Wire Anemometer. Some of its features included[2], constant temperature wire probe, a sampling frequency of a maximum 300 kHz for a film probe and 500 kHz for wire probes, built in thermocouple circuit for fluid temperature measurements, data acquisition and analysis software, capacity for up to eight channels (eight probes), and a SMARTTUNE™ Bridge Compensation Technology which reduces the oscillations in signal frequency.

The experimental setup was the TSI IFA 300 Constant Temperature Hot Wire Anemometry system, a PC to run the data acquisition software, a four channel connector board, a Hewlett Packard oscilloscope, a thermocouple temperature sensor and the hot wire velocity sensor (probe) (figure 1)

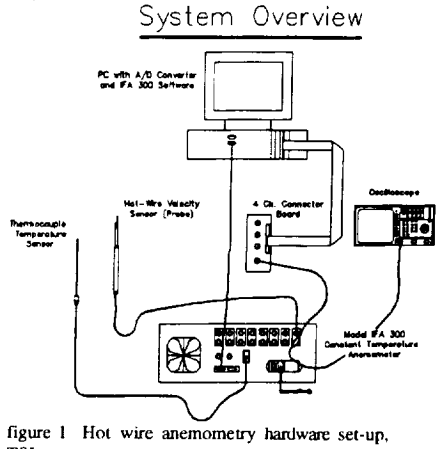


figure 1 Hot wire anemometry hardware set-up, TSI

300

A small table top fan was used to create the simple flow with fluctuations. With this setup, data point within the flow were taken both in the transverse and stream wise directions. (figure 2)

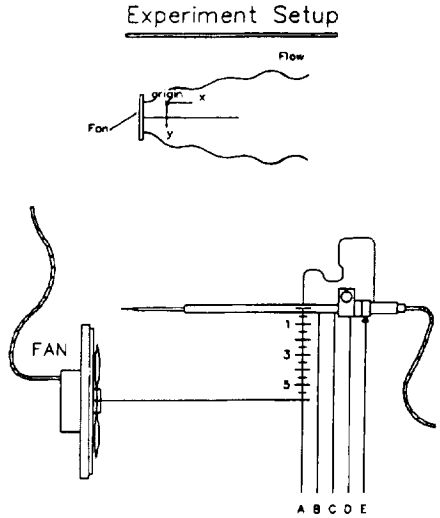


figure 2 Experimental set up of hot wire probe and table top fan

The IFA 300 has four main menus for setting up the probe and measuring the fluid velocities, a Diagnostic menu, a Calibration menu, a Data Acquisition menu and a Post Analysis menu. The diagnostics menu is used to determines several probe conditions such as probe resistance, cable resistance, operation temperature, overheat ratio, cable length, probe type and several other system conditions. Most of this information is provided by the manufacturer for the particular probe used. The calibration menu enables users to calibrate the probes relationship between bridge voltages and reference velocities. The probe data provided is entered into a probe data chart along with a selection of the method of calibration. Calibration methods consists of the option to enter probe data, voltage and/or pressure either manually or allowing acquisition by the system. In this experimental set up a calibration curve was provided that was very similar to the experimental probe which allowed for immediate measurement without requiring an initial calibration. Also from this menu, the calibration data, calibration coefficients and calibration curve may also be viewed. The calibration curve is a fourth order polynomial fit. The two main menus are the Data Acquisition menu and the Post Analysis menu. The Data Acquisition menu sets up several parameters and conditions for the collection of velocity data, such as a probe table consisting of the previously established probe data for each probe that is intended to record velocities. In this experiment, only one probe was used, however, numerous probes can collect data simultaneously. Other parameters can also be set such as the sample rate, atmospheric pressure, the units of velocity, temperature units, probe position, along with any comments desired. Once the acquisition is made, the system presents three components of velocity graphically simultaneously, along with the average velocity over the given sample range. the temperature measured by the thermocouple is also presented and the options to zoom and select different ranges to view on each plot is available via cursors.(figure 3) The final menu is the Post Analysis menu which allows further analysis of the data acquired from the data acquisition mode. Within this menu a time history of velocity of the flow can be viewed, which is a similar graphical representation of the flow as in the acquisition mode. Other options include a spectrum analysis plot, correlation plot (linear or logarithmic), and a flow field plot. There is also an option to save the data files so that they may be analyzed outside of the IFA 300 system.

The experiment consisted of measuring the flow of a simple desk top fan, 6 inch diameter, along a grid with 1 inch intervals stream wise and 0.5 inch interval transversely. Due to the symmetry about the fan's axis measurements were only taken along half of the flow area. The data was initially collected and stored in ASCII format in the PC and later transferred onto a Power Macintosh computer for further computations and analysis.

Results and Discussions:

The one dimensional velocity data were acquired over several runs. A sample of the velocity data as it appears on the IFA 300 computer screen is shown in figure 3.

From the averaged velocity data, velocity profiles[3,4] were constructed over the measuring grid for the fan flow field. The constructed velocity profile is shown in figure 4.

Turbulent intensities were measured from the measured velocity data after subjecting the data to Reynolds averaging. The measured turbulent intensities were plotted against the stream wise distance (figure 5) as well as the transverse distance (figure 6)

Conclusions and Suggestions:

The IFA 300 Hot Wire Anemometry system is a very advanced and versatile system. The range of data analysis and its ability to measure small fluctuation in flow velocity such as those seen in the flow produced by a table top fan proves its ability to measure key flow parameters for the intended calibration. The IFA 300 is very feasible for the calibration of the Direct Estimator Velocimetry.

Since calibration curves used in the feasibility study were only approximate curves, actually calibration curves will be made specifically for the experimental probe. The test section for calibration of the DEV must be fabricated then subsequent calibration of the DEV using LIF will be performed. After obtaining a three dimensional measuring probe to measure all three components of velocity, it is to be applied to a icing wind tunnel facility. Calibration of the three dimension probe must also be accomplished.

Acknowledgemdents

This work was supported in part by NASA grant NAGW 2930 to Florida A&M University.

References

1. Bradshaw, P., Experimental Fluid Mechanics, second edition, Pergamon Press, 1970

2. TSI Incorporated, IFA300 Constant Temperature Anemometer, Instruction manual, 1994

3. Fox, Robert W., McDonald, Alan T., Introduction to Fluid Mechanics, fourth edition, John Wiley & Sons, Inc., NY, 1992

4. Incropera, Frank P., Dewitt, David P, Fundamentals of Heat and Mass Transfer, third edition, John Wiley & Sons, Inc. NY 1990

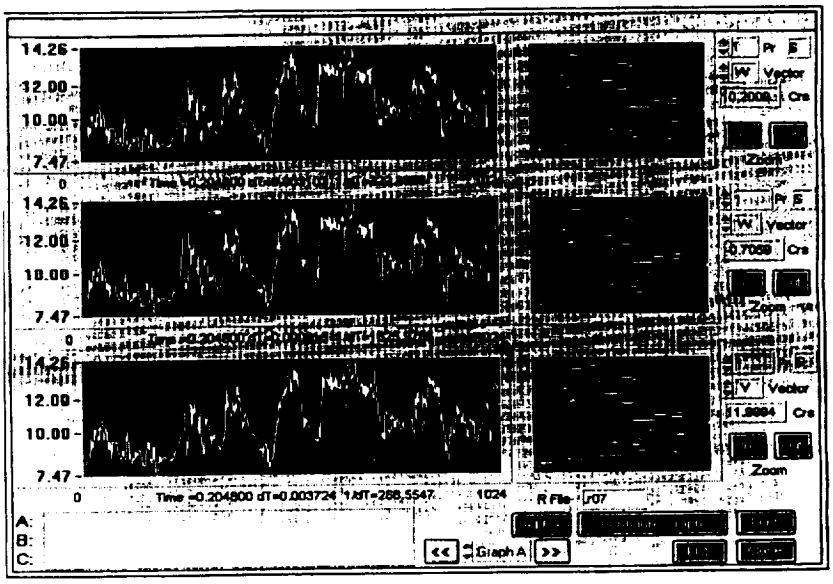


figure 3. A sample screen display of the IFA-300 velocity measurements.

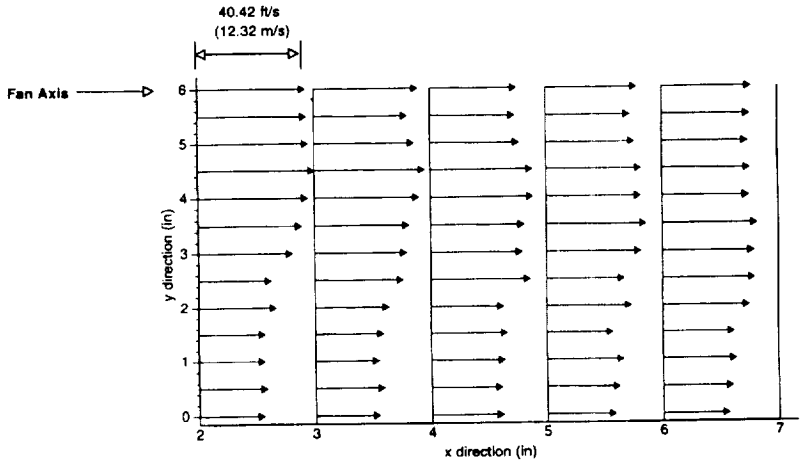


figure 4. Velocity profiles constructed from the hot-wire measurements.

Near Infrared Luminescence Properties of Mn^{5+}: $Ca_5(PO_4)_3F$

Valetta R. Davis* and Uwe Hömmerich
Hampton University
Department of Physics
Research Center for Optical Physics
Hampton, VA 23668

George B. Loutts
Norfolk State University
Center for Materials Research
2401 Corprew Avenue
Norfolk, VA 23504

Keywords: Spectroscopy, Luminescence, Solid-State Lasers, Crystal-field theory, Mn^{5+}

Abstract

We report a spectroscopic investigation of Mn^{5+} doped $Ca_5(PO_4)_3F$ or FAP. Mn^{5+} doped crystals have recently attracted world wide attention for potential solid-state laser applications. Following optical excitation of Mn: FAP with the 600 nm output of a Nd: YAG OPO laser system, we observed a strong near infrared luminescence centered at around 1150 nm. The room temperature luminescence decay time was measured to be approximately 635 μs. We attribute the infrared luminescence to the $^1E \rightarrow {}^3A_2$ transition of tetrahedrally coordinated Mn^{5+} ions located in a strong crystal field environment. Absorption, luminescence and lifetime data of Mn: FAP will be presented and discussed.

Introduction

In the last decade, the development of new tunable solid-state lasers has become an active area of research [1]. Most applications of tunable solid-state lasers remain in research and development. However, tunable lasers have emerged from the laboratory and are rapidly finding a variety of practical uses. Current applications of tunable solid-state lasers include basic scientific investigations in atomic, molecular, and solid-state spectroscopy, and laboratory studies of new semiconductor and fiber-optic devices. Future applications now under development include aircraft and based remote sensing lidar [1].

One of the most intensive trends in the development of broadly tunable solid-state lasers is transition metal ion lasers [1]. Transition metal ions with the $3d^2$ electronic configuration have special characteristics which make them potential active centers in solid-state laser materials. In a tetrahedral coordination environment, $3d^2$ ions exhibit strong visible absorption bands arising from the electric dipole-allowed transition $^3A_2 \rightarrow {}^3T_1$. The strong absorption is attractive for both flashlamp and diode pumping. Also important for laser material applications is the crystal field strength the $3d^2$ dopant ion experiences. The crystal field strength can be either weak or strong depending on whether the metastable level is the $^3T_2$ state or ^1E state. For the weak crystal field case the luminescence of the $3d^2$ ion is expected to be broad-band, whereas in the strong crystal field case the luminescence will be narrow-band [2].

Among the transition metal ions with a $3d^2$ electron configuration, Cr^{4+} has received the most attention. Laser emission has been demonstrated for the systems Cr^{4+}:forsterite [3,4], Cr^{4+}:YAG [5,6], and Cr^{4+}:Y_2SiO_5 [7,8]. The efficiency of these new laser materials is limited by nonradiative decay mechanisms which reduce the room temperature quantum efficiency of the system. Because of these drawbacks,

intensive research has begun to find host materials for other $3d^2$ ions such as Mn^{5+} [9,10,11]. Mn^{5+} is isoelectronic to Cr^{4+} and has the possibility of providing a higher laser efficiency. Laser action has been reported in Mn^{5+}: $Sr_5(VO_4)_3F$ [11,12] and Mn^{5+}: $Ba_3(VO_4)_2$ [11]. In this paper, we present the first spectroscopic study of Mn^{5+} doped into $Ca_5(PO_4)_3F$ (FAP).

Experimental Details

The $Ca_5(PO_4)_3F$ (FAP) crystal (apatite structure) was grown at the Center for Materials Research at Norfolk State University using the Czochralski method. The growth of FAP is well developed because of its potential as a host for rare earth based laser materials [13]. The nominal Mn concentration used for the sample in this study was 1%. The Mn: FAP crystal exhibits a blue color which changes slightly along its cross-section. This color change indicates an inhomogenous dopant distribution. The PO_4^{3-} group in FAP provides a tetrahedral site which allows the incorporation of Mn^{5+}.

The room-temperature absorption spectrum was taken with a Cary 5E Spectrophotometer. The 600 nm output of a SURELITE OPO laser was used as the excitation source for the room-temperature luminescence measurements. The SURELITE OPO is an optical parametric oscillator which can be manually tuned to produce coherent, broadband radiation between 410 nm and 2200 nm. It is pumped by 355 nm light from a Nd: YAG laser. The luminescence was dispersed by an Acton Research Corporation 1 meter monochromator using a grating blazed at 1.6 μm. Detection was achieved with a high sensitivity, liquid nitrogen-cooled Ge-detector (Applied Detector Corporation, Model 403L). The signal was digitized by a Stanford Research Systems Boxcar Averager and stored on a DOS computer.

The excitation source for luminescence lifetime measurements was again the SURELITE OPO laser, and the light was dispersed by the 1 meter monochromator. The high sensitivity Ge-detector was replaced with a faster responding, liquid nitrogen-cooled Ge-detector (Applied Detector Corporation, Model 403HS). The lifetime was averaged by a Tektronik TDS 420A digitizing oscilloscope interfaced with a DOS computer.

For low-temperature measurements, the sample was cooled using an Advanced Research Systems cryogenic refrigeration system. The refrigerator allowed the temperature to be varied between 12K and 300K. The experimental arrangement used is shown schematically in Figure 1.

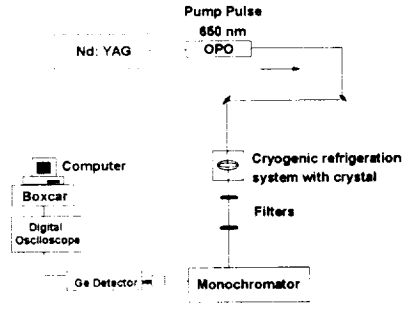


Figure 1. Schematic of arrangement for luminescence spectra and lifetime measurements.

Results and Discussion

Energy levels of tetrahedrally coordinated Mn^{5+} ions

The Mn^{5+} ion has two d-electrons in the unfilled (3d) shell. The ground state (lowest energy level) of the free ion is a ^3F state (S = 1, L = 3). The excited states (next higher energy levels) are ^1D, ^3P, and ^1G. Optical transitions of Mn^{5+} occur within the unfilled 3d shell [2].

When Mn is doped into FAP, the result is a tetrahedral oxygen configuration of Mn^{5+} ions. This oxygen configuration forms a crystal field which influences the free ion Mn^{5+} energy levels. Figure 2 shows an overview of the local Mn environment.

Utilizing crystal field theory [2,14], it can be interpreted that the free ion levels change energy positions depending on the strength of the crystal

field. In many cases, the free ion level may even split into several new levels.

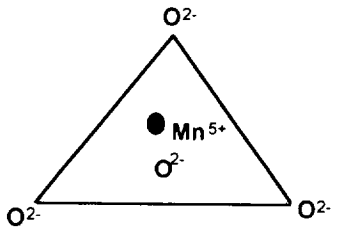


Figure 2: Tetrahedral oxygen coordination of Mn^{5+}.

The effect of the crystal field on the free ion energy level for a $3d^2$ system has been calculated by Tanabe-Sugano as a function of the crystal field strength parameter (Dq), and the Racah parameters (B and C) [2,14].

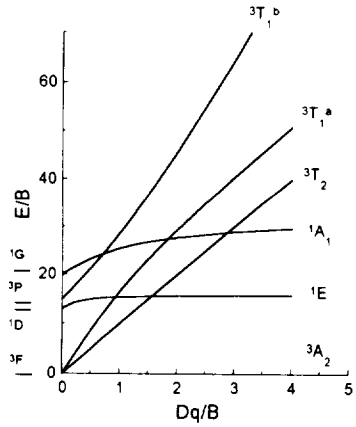


Figure 3: Tanabe-Sugano diagram for the electronic energy levels of a $3d^2$ ion in a tetrahedral crystal field [2,14].

The crystal field strength parameter Dq is directly related to the interatomic distance R between the Mn ion (central ion) and the four surrounding oxygen ions (ligands), i.e., $Dq \sim 1/R^5$ [2,14]. The shorter R is, the stronger the crystal field, and consequently, its effect on the free ion level. The Racah parameters B and C consider corrections on the Mn energy levels due to covalence effects, and are normally determined from the experiment [14].

The Tanabe-Sugano diagram in Figure 3 shows that the free ion energy levels are obtained for Dq/B = 0. This means the crystal field is "switched off". The effect of the crystal field (Dq/B > 0) results in the splitting of the ground state into three energy levels: $^3A_2$, $^3T_2$, and $^3T_1$. The order of the energy levels is $^3A_2 < {}^3T_2 < {}^3T_1$. The crystal field notation of the energy levels can be derived from group theory. The 'A', 'E', and 'T' states indicate orbital degeneracies of 1 (non-degenerate), 2, and 3 respectively. The orbital degeneracy of the E and T states can be removed if the tetrahedral site is distorted. The Tanabe-Sugano diagram will be used in the following discussions to identify the 5+ valence state of Mn in FAP, and to discuss the energy levels.

Absorption Spectrum of Mn: FAP

From group theory [2,14], it can be inferred that the only electric-dipole allowed transition of a $3d^2$ ion is the $^3A_2 \rightarrow {}^3T_1$ transition. Therefore, a characteristic absorption feature for $3d^2$ ions is a strong and broad visible absorption. Another "fingerprint" for $(3d)^2$ ions is a weaker infrared absorption which arises from the electric dipole forbidden transition $^3A_2 \rightarrow {}^3T_2$. Figure 4 shows the unpolarized absorption spectrum of Mn: FAP.

The main feature of the absorption spectrum is a broad band which spans the entire visible region. The strong visible absorption is extremely attractive for either flashlamp or diode laser pumping. A significantly weaker absorption spectrum is observed in the near infrared spectral region. Since the absorption spectrum of Mn: FAP exhibits the two characteristics discussed above, we conclude that Mn was successfully incorporated into the FAP lattice as Mn^{5+}. Further support of this conclusion will be

provided in the discussion of the luminescence properties of Mn: FAP.

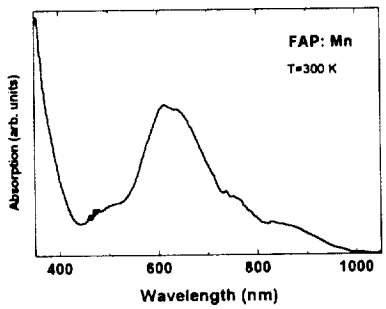


Figure 4: Room-temperature absorption spectrum of Mn^{5+}: FAP

Luminescence of Mn: FAP

Figure 5 shows the near infrared luminescence spectra of Mn: FAP at room and low temperature (12 K). The spectra span over the region from 1100-1300 nm, and exhibit a sharp peak at 1150 nm (zero-phonon line) and some weaker vibronic structure around 1200 nm.

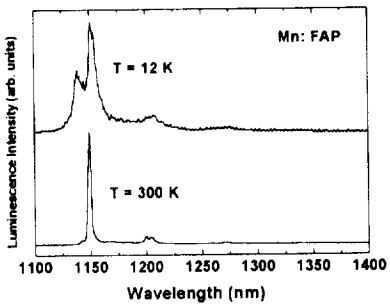


Figure 5: Near infrared luminescence of Mn^{5+}: FAP at low and room temperature

Transients of the luminescence decay time at 300 and 12 K are shown in Figure 6. The luminescence lifetime is approximately single exponential. The 1/e lifetime at T = 12 K and T = 300 K was 635 μs and 730 μs, respectively. The small change in lifetime with respect to temperature suggests a high quantum efficiency of the near infrared luminescence (~ 90%).

The infrared luminescence of Mn: FAP is similar to observations made for other Mn^{5+} doped oxide crystals. The energy position, narrow band shape, and lifetime of the luminescence is consistent with the $^1E \rightarrow {}^3A_2$ transition of Mn^{5+} ions in a strong crystal field environment ($Dq/B > 1.5$). The high luminescence quantum yield, coupled with the fact that Mn: FAP can be grown with a high optical quality, makes this system attractive as a potential solid-state laser material. Tunability of such laser system, however, will be limited by the narrow luminescence band.

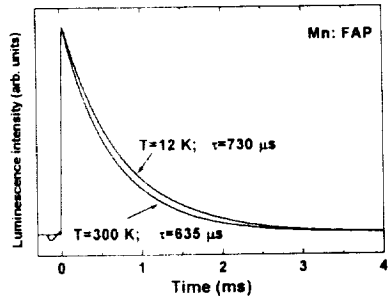


Figure 6: Luminescence lifetime of Mn^{5+}: FAP at 12 K and 300 K

Conclusion

The spectroscopic investigation of Mn doped FAP showed a broad absorption spectrum and a strong near infrared luminescence characterized by an intense peak at 1150 nm. The observed absorption and luminescence features reveal that Mn was successfully incorporated into the FAP lattice as tetrahedrally coordinated Mn^{5+}. A luminescence quantum efficiency of ~ 90% was estimated from lifetime measurements. The strong and efficient infrared luminescence of Mn^{5+}: FAP indicates the potential of this material as a new infrared solid-state laser

material. In the future, additional luminescence experiments at temperatures greater than 300 K will be performed in an attempt to induce broad band luminescence from the $^3T_2$ level. This would provide the opportunity to obtain tunable laser activity from Mn^{5+}: FAP.

Acknowledgements

Valetta Davis and Uwe Hömmerich would like to acknowledge the financial support of NASA through Grant NAGW-2929.

*Valetta Davis received her B.S. degree in Physics from Fisk University. She recently completed her first year of graduate study at Hampton University under the guidance of Uwe Hömmerich. Her research to date has primarily been in the area of Solid-State Laser Spectroscopy. Valetta anticipates graduating in the Spring of '97.

References

[1] Peter F. Moulton, Proceedings of the IEEE, **Vol. 80, No. 3**, (1992) 348.

[2] B. Henderson and G. F. Imbusch, Optical Spectroscopy of Inorganic Solids, (Clarendon Press, Oxford, 1989)

[3] V. Petricevic, S. K. Gayen, and R. R. Alfano, Appl. Phys. Lett. **52** (1988) 1040 and Appl. Phys. Lett. **53** (1988) 2590.

[4] H. R. Verdun, L. M. Thomas, D. M. Andrauskas, T. McCollum and A. Pinto, Appl. Phys. Lett **52** (1988) 2593.

[5] A. P. Shkadarevich, in OSA Proceedings on Advanced Solid-State Lasers, Optical Society of America, Washington DC (1989) **Vol 5**, 60.

[6] S. Kuck, J. Koetke, K. Petermann, U. Pohlmann, and G. Huber, in OSA Proceedings on Advanced Solid-State Lasers, Optical Society of America, Washington DC (1993), **Vol. 15**, 334.

[7] C. Deka, B. H. T. Chai, Y. Shimony, X. X. Zhang, E. Munin, and M. Bass, Appl. Phys. Lett. **61** (1992) 2141.

[8] J. Koetke, S. Kuck, K. Petermann, G. Huber, G. Cerullo, M. Danailov, V. Magni, L. F. Qian and O. Sevelto, Optics Comm. **101** (1993) 195.

[9] U. Oetliker, M. Herren, H. U. Gudel, U. Kesper, C. Albrecht, and D. Reinen, J. Chem. Phys. **100** (1994) 8656.

[10] U. Hömmerich, H. Eilers, and W. M. Yen, Chem. Phys. Lett. **213** (1993) 163.

[11] L. D. Merkle, A. Pinto, H. R. Verdun, and B. McIntosh, Appl. Phys. Lett. **61** (1992)

[12] L. D. Merkle, Y. Guyot, and Bruce H. T. Chai, J. Appl. Phys. 77 (1995) 474. 2386.

[13] L. D. Merkle, B. Zandi, and Bruce H. T. Chai, in Advanced Solid-State Lasers, Technical Digest, Optical Society of America, Washington DC (1996) 320.

[14] H. L. Schlafer and G. Gliemann, Ligand Field Theory, (Wiley, New York, 1969).

S71-27 290735 p. 7

Hyperthermal Ion Scattering from an Organic Thin Film

M. U. Spurlock, C. W. Lowe, S.C. Mathur
Hampton University, Department of Physics, Hampton, Virginia 23668

As an ion impinges on a surface of a film the ion transfer its energy by inelastic and elastic processes. The inelastic process occurs when the ion transfers its energy to the electron cloud of the target atom . The elastic process occurs when the ion transfers its energy to nuclei of the target atom. The dominate inelastic and elastic processes are determined by the ion's incident energies. We have showed in the hyperthermal ion scattering range that the inelastic process is dominant. Therefore, the dominate means for energy transfer in the hyperthermal ion range will be from the ion to the target atom. We have simulated the single impulsive scattering of a 50 eV Li^+ ion from a thin film of anthracene in a temperature range of 275 K to 50 K. We observed peak broadening with increasing temperature.

Introduction

There has been a gradually surge in the study of organic material in the recent years this is partly due to the unique properties of organic material. Increased research will lead to the development of organic devices that can have properties of integrating organic electronic devices to biological systems, compact, environmentally safe, inexpensive, durable, and enhance performances. Developments in organic material science research are fast pace and increasingly changing. Organic light emitting diodes (OLED) for flat panel displays have been developed. These OLED have maximum output of .3 mW and quantum efficiency of 5.5%.[1] There is considerable amount of advances towards organic semiconductors. Storage time effects have been studied in poly ethylene terphthalate for improve device applications.[2] Also, optical and electrical properties are enhance by the use for molecular organic heterojunction.[3] In the fields of laser spectroscopy and processing, optical and electrical properties were changed in 3, 4, 9, 10-perylenetetracarboxylic dianhydride (PTCDA) and 1, 4, 5, 8, naphthalenetetracarboxylic dianhydride (NTCDA) .[4]

Nonlinear optical (NLO) properties in organic crystals are primary for optical device application. There has been development of organic optical devices that have low signal to noise ratio.[5] In theory the nonlinear optical properties can be described by applying a uniform electric field to an organic crystal. The

electrical field distorts molecules resulting in a uniform polarization of the molecules occurs with in the organic crystal. Increasing the electric field leads to saturation of polarized molecule resulting in nonlinear changes within the polarization. First, second and third order generation of waves may occur in this process. Third order wave generation decreases the electronic bonding to nuclei that directly changes the electrical susceptibility of the crystal. The delocalization of these electrons may lead to the generation of photo-carries whereas the crystal has demonstrated nonlinear electrical and optical properties. In this process the crystal's permitivitty and index of refraction are change. The engineering of NLO properties allows the development of optical devices that have ideal properties of electrical field strength, high performance of photoconductivity, and high saturated current velocities.[6]

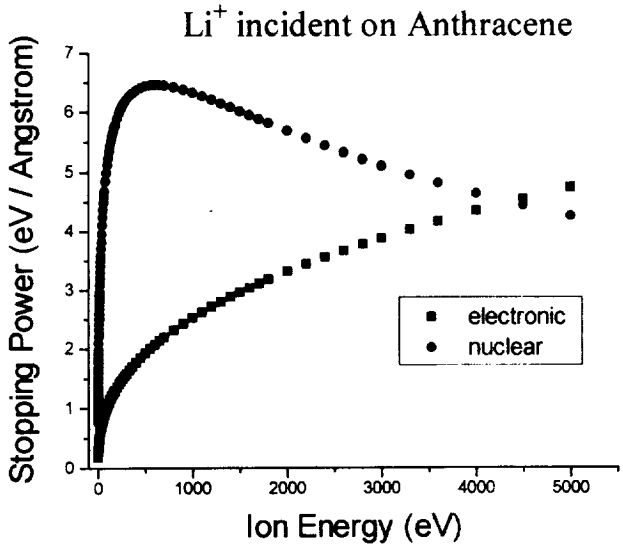
A primary mechanism for enhancement of nonlinear optical properties in organic material is the delocalization of electrons. The delocalization of electrons may be describe by the dynamics of electrons in combination with holes. Electrons in combination with holes are bound by an electrostatic attraction to the positive hole and this combination is known as an exciton. Exciton dynamics are complex that involves trapping, ionization, radiationless decay processes, exciton annihilation, and other diffusion mechanism. The exciton may migrate with recombining of the hole and the electron resulting in the emission of a photon (NLO properties). In organic material a primary mean for exciton transport is natural vibrational states or phonons.[7] The exciton-phonon interaction has been seen in reflection spectroscopy experiments of anthracene.[8] In organic material the mean for NLO properties is the mobilities of excitons in organic material.[9] The exact nature of exciton dynamics aren't fully understood. The migration of excitons are mainly a process of natural vibrational states or phonons. Theoretical analyses of the exciton-phonon have been made in optical cavity models in organic material.[10] However, experimentally the nature of exciton-phonon coupling is unclear. Exciton-phonon coupling in organic material can be studied by low energy ion scattering.

General Description of Simulation

In a low energy scattering event an ion can impinge into the outer most surface layer and be scattered from a target atom within outer most the layer. The ion losses its kinetic energy by elastic and inelastic processes. The inelastic process is the loss of energy by the ion to the electron cloud of the target atom. The inelastic process is the loss of energy by the target nucleus. The inelastic process is characterized by the electronic stopping power and the elastic process is characterized by a nuclear stopping power. The incident energy determines what

processes are dominant. Figure 1 shows the nuclear stopping power is dominate for an energy range of 1eV to 4 keV. This means that elastic process isdominate for Li^+ ions incident on anthracene at energies from 1eV to 4 keV. The interaction of the ion with the surface atom involves the transfer of energy by a recoil of

Figure 1

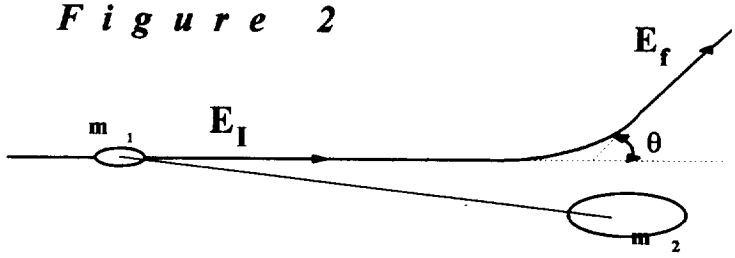


nuclei resulting in scattering. In this model all inelastic processes are neglible. In an elastic process of his type energy and momentum is conserved. Thus, a scattering event can be modeled by the use of conservation of energy and momemtum. Each scattering event is described as an ion scattering with one surface atom. Figure 2 is an illustration of the interaction of the ion and surface atom as a binary collision event. The use of conservation and momentum makes this scattering model simple such that the ratio of the scattering energy (E_f) and incidence energy (E_i)of the ion can be written as:

$$\frac{E_f}{E_i} = \left[\frac{\sqrt{1-\mu^2 \sin^2\theta} + \mu\cos\theta}{(1+\mu)} \right]^2$$

where $\mu = m_1 / m_2$ and θ - laboratory scattering angle.

Figure 2



In any real case atoms experience surface vibrations. These vibrations are produced at the surface and may be characterized as normal modes that are independent of the scattering event. The thermal contribution to the scattering event introduces the following conditions[11]

$E_i - E_f \rangle\rangle \frac{\hbar\omega_D}{2\pi}$ and $\lambda >> a$ and $K_B T >> \frac{\hbar\omega_D}{2\pi}$

where $\omega_D = \frac{6\pi v^3 N}{V}$ v- constant velocity of sound, V - volume

N - number of accoustic phonon modes

ω_D - Debye frequency of the crystal

λ - wavelength of the incoming ion

a - lattice constant of the crystal

The ion is assumed to have impulsive scattering. Whereas the collision time of the ion can be modeled as less than the vibrational period of the surface atom. Therefore we may write the following:

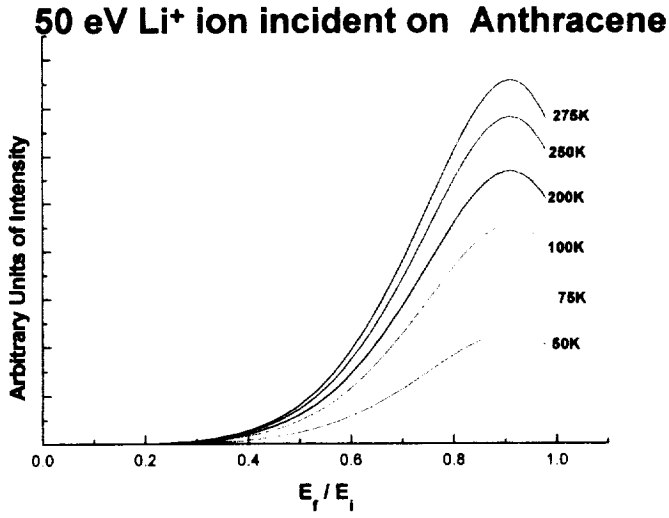
$$(E_i - E_f)^2 = h(\theta)\, E_r\, K_B T \quad \text{where } h(\theta,\mu) = \frac{4\mu}{(1+\mu)^2} \times \left(\frac{1+\mu\sin^2\theta}{1-\mu^2\sin^2\theta} - \frac{\cos\theta}{\sqrt{1-\mu^2\sin^2\theta}}\right)$$

The scattering model becomes simply a function of scattering geometry and momentum fluctuations can be measured by observing scattering trajectories by the ion.

Conclusion

Figure 3 is asimulation of single impulsive scattering of a 50 eV Li^+ from anthracene at surface temperatures 50 K to 275 K.

Figure 3



Peak broadening is observed with increasing temperature. Clearly a linear dependence of thermal fluctuation from surface atoms to peak broadening is

observed. Peak broadenings indicate a decrease within the vibrational modes with decreasing temperature. Damping towards intra to inter phonon modes maybe predicted with a changing temperature. However, the authors believe that peak brodening due to thermal fluctuations of anthracene surface atoms or any organic surface atoms would not be as distinct with a change temperature. This is due to the weak Vander Waal interactions that would make it difficult to distinquish peak broadening with surface temperature changes from 50 K to 275 K. An approximation must be made to the scattering model followed by experimental data comparison.

References

1. D. Ammerman, A. Bohler, C. Rompf, and W. Kowalsky, OSA / ACS conference on Organic thin films for photonic application in Portland Oregon (1995)
2. J. Belana, M. Mudarra, P. Colomer and M. Latour, Journal of Material Science, **30** (1995) 5241-5245
3. S. R. Forrest, L. Y. Leu, F. F. So, and W. Y. Yoon, Journal of Applied Physics, **66**,12(Dec. 1989) 5908-5914
4. M. L. Kaplan, S. R. Forrest, P. H. Schmidt, and T. Venkatesan, Journal of Applied Physics, **55**, 3 (Feb. 1984) 732-742
5. B. G. Penn, B. H. Cardelino, C. E. Moore, A. W. Shields, and D. O. Frazier, Prog. Crystal Groeth and Charact., **22** (1991) 19-51
6. M. D. Whitfield, S. Chan, and R. B. Jackman, Applied Physics Letters, **68**, 3 (Jan. 1996) 290-292
7. M.G. Sceats and G.C. Morris, Phys. Stat. Solidi, **14A** (1972) 643
8. M. G. Sceats and G. C. Morris, The Journal of Chemical Physics, **60**,2 (Jan. 1974) 375-392
9. M. Pope, and C. E. Swenberg, Electronic Process in Organic Crystals
10. X. Li, D. L. Lin, T. F. George, and X. Sun, Physics Review B, **42** (1990) 2977
11. D. M. Goodstein, C. A. Dirubio, and B. H. Cooper, Surface Review and Letters, **1**, 1 (1994) 175-185

observed. Peak broadenings indicate a decrease within the vibrational modes with decreasing temperature. Damping towards intra to inter phonon modes maybe predicted with a changing temperature. However, the authors believe that peak brodening due to thermal fluctuations of anthracene surface atoms or any organic surface atoms would not be as distinct with a change temperature. This is due to the weak Vander Waal interactions that would make it difficult to distinquish peak broadening with surface temperature changes from 50 K to 275 K. An approximation must be made to the scattering model followed by experimental data comparison.

References

1. D. Ammerman, A. Bohler, C. Rompf, and W. Kowalsky, OSA / ACS conference on Organic thin films for photonic application in Portland Oregon (1995)
2. J. Belana, M. Mudarra, P. Colomer and M. Latour, Journal of Material Science, **30** (1995) 5241-5245
3. S. R. Forrest, L. Y. Leu, F. F. So, and W. Y. Yoon, Journal of Applied Physics, **66**,12(Dec. 1989) 5908-5914
4. M. L. Kaplan, S. R. Forrest, P. H. Schmidt, and T. Venkatesan, Journal of Applied Physics, **55**, 3 (Feb. 1984) 732-742
5. B. G. Penn, B. H. Cardelino, C. E. Moore, A. W. Shields, and D. O. Frazier, Prog. Crystal Groeth and Charact., **22** (1991) 19-51
6. M. D. Whitfield, S. Chan, and R. B. Jackman, Applied Physics Letters, **68**, 3 (Jan. 1996) 290-292
7. M.G. Sceats and G.C. Morris, Phys. Stat. Solidi, **14A** (1972) 643
8. M. G. Sceats and G. C. Morris, The Journal of Chemical Physics, **60**,2 (Jan. 1974) 375-392
9. M. Pope, and C. E. Swenberg, Electronic Process in Organic Crystals
10. X. Li, D. L. Lin, T. F. George, and X. Sun, Physics Review B, **42** (1990) 2977
11. D. M. Goodstein, C. A. Dirubio, and B. H. Cooper, Surface Review and Letters, **1**, 1 (1994) 175-185