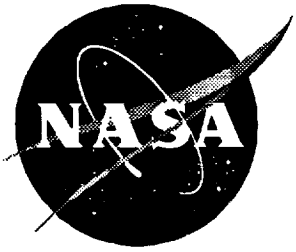
Conference Publication 3246
DOT/FAA/CT-94/03



FAA/NASA Joint University Program for Air Transportation Research 1992–1993

Frederick R. Morrell
Langley Research Center • Hampton, Virginia

Proceedings of a conference sponsored by the Federal Aviation Administration, Washington, D.C., and the National Aeronautics and Space Administration, Washington, D.C., and held in Athens, Ohio
June 17–18, 1993

National Aeronautics and Space Administration
Langley Research Center • Hampton, Virginia 23681-0001

February 1994

PREFACE

The Joint University Program for Air Transportation Research is a coordinated set of three grants sponsored by the Federal Aviation Administration and NASA Langley Research Center, one each with the Massachusetts Institute of Technology (NGL-22-009-640), Ohio University (NGR-36-009-017), and Princeton University (NGL-31-001-252). These research grants, which were instituted in 1971, build on the strengths of each institution. The goals of this program are consistent with the aeronautical interests of both NASA and the FAA in furthering the safety and efficiency of the National Airspace System. The continued development of the National Airspace System, however, requires advanced technology from a variety of disciplines, especially in the areas of computer science, guidance and control theory and practice, aircraft performance, flight dynamics, and applied experimental psychology. The Joint University Program was created to provide new methods for interdisciplinary education to develop research workers to solve these large scale problems. Each university submits a separate proposal yearly and is dealt with individually by FAA and NASA. At the completion of each research task, a comprehensive and detailed report is issued for distribution to the program participants. Typically, this is a thesis that fulfills the requirements for an advanced degree or a report describing an undergraduate research project. Also, papers are submitted to technical conferences and archival journals. These papers serve the Joint University Program as visibility to national and international audiences.

To promote technical interchange among the students, periodic reviews are held at the schools and at an FAA or NASA facility. The 1992-1993 year-end review was held at Ohio University, Athens, Ohio, June 17-18, 1993. At these reviews the program participants, both graduate and undergraduate, have an opportunity to present their research activities to their peers, to professors, and to invited guests from government and industry.

This conference publication represents the twelfth in a series of yearly summaries of the program. (The 1991-1992 summary appears in NASA CP-3193.) Most of the material is the effort of students supported by the research grants. Four types of contributions are included in this publication: a summary of ongoing research relevant to the Joint University Program is presented by each principal investigator, completed works are represented by full technical papers, research previously in the open literature (e.g., theses or journal articles) is presented in an annotated bibliography, and status reports of ongoing research are represented by copies of presentations with accompanying text.

Use of trade names of manufacturers in this report does not constitute an official endorsement of such products or manufacturers, either expressed or implied, by the National Aeronautics and Space Administration or the Federal Aviation Administration.

Frederick R. Morrell
NASA Langley Research Center

CONTENTS

PREFACE iii

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

AN INVESTIGATION OF AIR TRANSPORTATION TECHNOLOGY AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 1992-93..... 3 - 1
Robert W. Simpson

A GRAPHICAL WORKSTATION BASED PART-TASK FLIGHT SIMULATOR FOR PRELIMINARY RAPID EVALUATION OF ADVANCED DISPLAYS..... 9 - 2
C. Wanke, J. Kuchar, E. Hahn, A. Pritchett, and R.J. Hansman

A DATA FUSION ALGORITHM FOR MULTI-SENSOR MICROBURST HAZARD ASSESSMENT..... 17 - 3
Craig R. Wanke and R. John Hansman

OHIO UNIVERSITY

INVESTIGATION OF AIR TRANSPORTATION TECHNOLOGY AT OHIO UNIVERSITY, 1992-93..... 31 - 4
Robert W. Lilley

FAULT DETECTION AND ISOLATION..... 35 - 5
Greg Bernath

IMPROVED MODELING OF GPS SELECTIVE AVAILABILITY..... 45 - 6
Michael S. Braasch, AnnMarie Fink, and Keith Duffus

REALTIME MITIGATION OF GPS SA ERRORS USING LORAN-C..... 55 - 7
Soo Y. Braasch

A GPS COVERAGE MODEL..... 63 - 8
Trent A. Skidmore

GROUND STATION SITING CONSIDERATIONS FOR DGPS..... 73 - 9
James D. Waid

PRINCETON UNIVERSITY

INVESTIGATION OF AIR TRANSPORTATION TECHNOLOGY AT PRINCETON UNIVERSITY, 1992-93..... 87 - 10
Robert F. Stengel

DESIGNING ROBUST CONTROL LAWS USING GENETIC ALGORITHMS..... 99 - 11
Chris Marrison

PAGE 1 INTENTIONALLY BLANK

OPTIMAL NONLINEAR ESTIMATION FOR AIRCRAFT FLIGHT CONTROL IN WIND SHEAR 111
Sandeep S. Mulgund

AIR TRAFFIC MANAGEMENT AS PRINCIPLED NEGOTIATION BETWEEN INTELLIGENT AGENTS 131
J.P. Wangermann

OPTICAL COMMUNICATIONS FOR TRANSPORT AIRCRAFT 141
Robert F. Stengel

ROBUSTNESS OF SOLUTIONS TO A BENCHMARK CONTROL PROBLEM 147
Robert F. Stengel and Christopher I. Marrison

STOCHASTIC PREDICTION TECHNIQUES FOR WIND SHEAR HAZARD ASSESSMENT 155
D. Alexander Stratton and Robert F. Stengel

STOCHASTIC MEASURES OF PERFORMANCE ROBUSTNESS IN AIRCRAFT CONTROL SYSTEMS 161
Laura Ryan Ray and Robert F. Stengel

OMIT

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

AN INVESTIGATION OF

AIR TRANSPORTATION TECHNOLOGY

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

1992 - 1993

Robert W. Simpson
Flight Transportation Laboratory
M.I.T., Cambridge, MA 02139

SUMMARY OF RESEARCH ACTIVITIES

1. INTRODUCTION

One completed project and two continuing research activities are under the sponsorship of the FAA/NASA Joint University Program as the 1992-93 period ends. There were a number of publications during the year which are referenced in this report. A brief summary of the continuing research project is provided.

The completed project was:

1. Liu, Manly, "Tracking Aircraft around a Turn with Wind Effects", MIT Flight Transportation Laboratory Report, FTL 93-4, June 1993.

The active research projects are:

1. ASLOTS - An Interactive Adaptive System for Automated Approach Spacing of Aircraft.

2. Alerting in Automated and Datalink Capable Cockpits.

2. REVIEW OF CONTINUING RESEARCH ACTIVITIES

2.1 ASLOTS - Interactive, Adaptive Spacing on Final Approach

This research is aimed at providing ATC controllers concerned with approach spacing at busy airports with a decision support system which is:

1) Integrated across multiple simultaneous approaches
2) Interactive (so that they can direct its operation)
3) Adaptive (it adapts continuously to the real world situation).

The ASLOTS concept was described in last year's report. The effort during 1992-93 has been aimed at creating a high fidelity ATC simulation environment called ATCSIM. This simulation will provide realistic motion of aircraft under typical representation of errors from various navigation, guidance, surveillance, and ground tracking systems, as well as the time and spatial variation of winds. It has two components: an airborne simulation for arriving and departing aircraft, and a ground simulation of aircraft moving on the surface of the airport. A schematic of ATCSIM functionalities is shown in Figure 1.

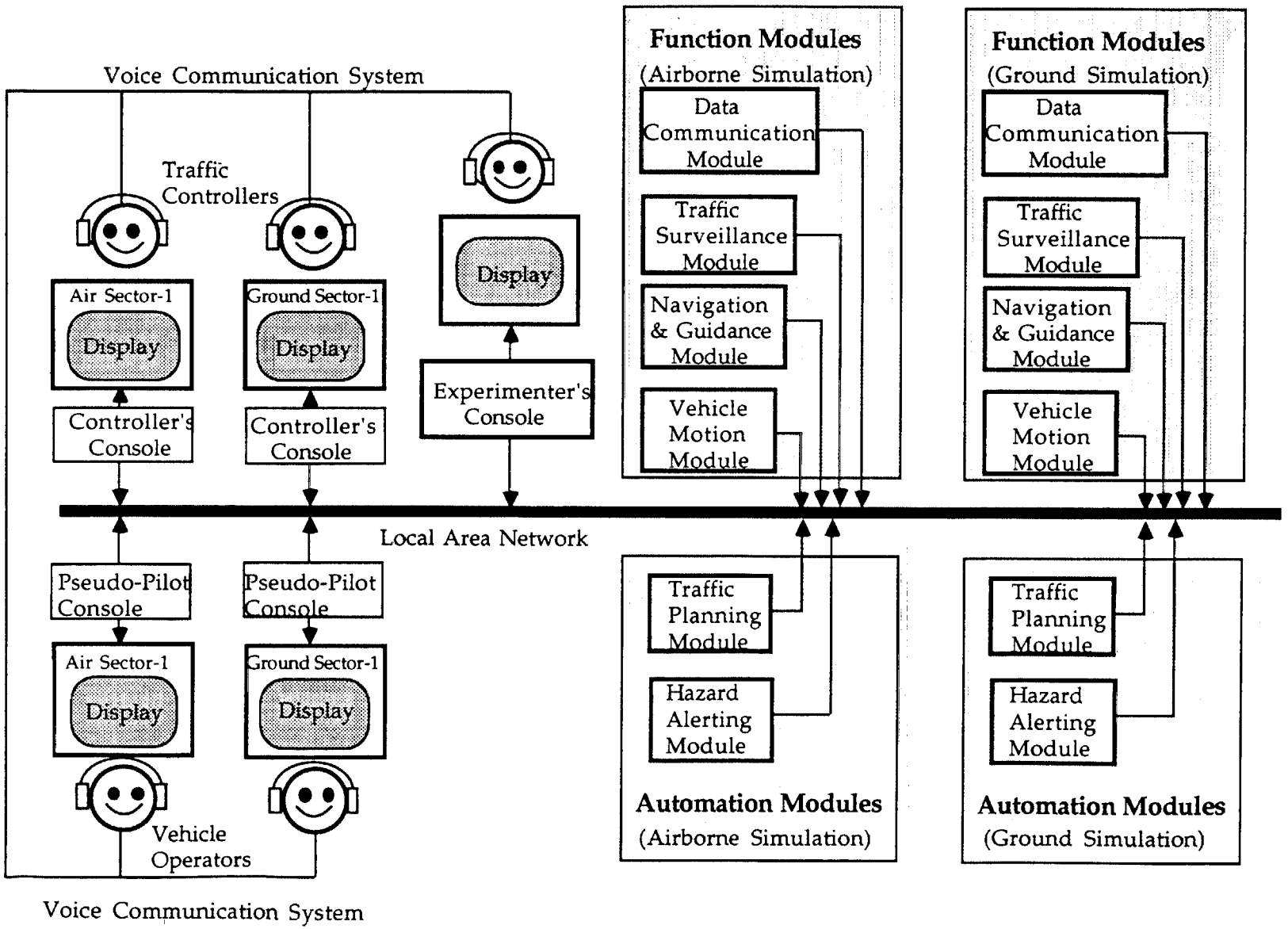
ATCSIM is designed using a generalized, modular software approach which can be easily adapted to new scenarios, and thereby provide a flexible, rapid-experimentation

tool for researchers interested in automation of ATC processes and Human Factors issues in ATC automation. It applies distributed processing using common workstations on a high speed local area network, and an object-oriented, modular approach to configuring the software which allows rapid reconfiguration of the traffic controller's console, its display formats, and its automation functions. It is written in standard ANSI C, uses X-windows for its graphics, Ethernet with TCP/IP protocol, and is currently in UNIX (AT&T System 5.3). This combination allows a variety of workstations to be used. ATCSIM will accommodate several ATC controller consoles (each with its pseudo-pilot station).

The modularity is indicated in Figure 1 where separate modules exist for communications, navigation and guidance, surveillance and tracking, and vehicle motion which provide realistic representation of the flight and ground paths followed by aircraft as they are controlled. Figure 1 also indicates that various modules for automating any or all of the various ATC processes (e.g., Conformance, Separation, Congestion Management, Hazard Alerting, etc.) can be developed separately. ATCSIM runs in real-time using a fixed script of arriving traffic, or can use Traffic Generators which construct a description of randomly arriving traffic with control over the longer term values for the mix of types, arrival rates, altitudes or gates, etc. Once a script is created it can be used by the experimenter for a series of tests. It is possible to "replay" any test run in fast-time, or "fast-forward" to any situation which is interesting. Such situations can then be the starting point for real-time experiments, and can be "doctored" to cause certain desired traffic situations to occur.

While the major effort in 1992-93 has been on creating ATCSIM, attention has now been returned to implementing ASLOTS. Current work is aimed at implementing its features (Feasible Slot Range, Auto Rearward Shift, Centerline Adaptation, Constrained Pattern Parameters, etc.) in an environment which will allow multiple runway approach and landings.

Figure 1 - ATCSIM FUNCTIONALITIES



2.2 Alerting in Automated and Datalink Capable Cockpits

Over the past few years, a variety of experiments have been performed in the MIT Advanced Cockpit Simulator in the area of weather hazard and terrain alerting. As a result of these experiments, it was noticed that there is a common generic thread in implementing advanced alerting systems. The focus of this research is to explore the idea of a general theory for Advanced Hazard Alerting in future situations where there may be a mix of airborne and ground sensors, and a reliable datalink to exchange information quickly. It is assumed both that pilots and controllers will be involved in detecting and resolving any deviation required by an unexpected hazard, and that their respective roles will be well defined.

While the different types of hazards (precipitation, wind shear, terrain, traffic) present different inputs, there are always five sequential processes in a Hazard Avoidance process:

1. Hazard Detection and Alerting
2. Communication / Display of Hazard Information
3. Generation and Decision on Hazard Resolution
4. Communication of Planned Resolution Path
5. Execution and Monitoring of the Resolution Path

The decision on the resolution path is assumed to be the responsibility of the captain of the aircraft. There will be "reaction" times necessary for the execution of each process, and the need to establish detection, intervention, and resolution criteria which are a function of hazard detection sensor performance, display capabilities, and aircraft state and performance capabilities. It is clear that the uncertainties in hazard detection vary with the "probe" or "lookahead" time. Various strategies for minimizing risk must be developed which are acceptable to both pilots and controllers. It is intended that pilot acceptance will be explored using the MIT Advanced Cockpit.

3. ANNOTATED REFERENCES OF 1992 - 93 PUBLICATIONS

3.1 Liu, Manly, Tracking Aircraft around a Turn with Wind Effects, SM Thesis, Department of Aeronautics and Astronautics , MIT, Cambridge, MA, 02139, June 1993

Currently, it is possible for ATC to use radar tracking to estimate an aircraft's current groundspeed and direction if it is flying a straight path, but large transient errors occur when the aircraft begins and ends a turn. The introduction of SSR Mode S datalink will make aircraft state information (heading, turn rate, groundspeed and direction, etc.) available for ground-based radar trackers, but it is desirable to minimize such transmissions. The minimal state information would be the declaration that the aircraft is no longer in a state of straight-line flight, but is currently turning. A "Turn Signal" indicating a left or right turn can be sent whenever the aircraft maintains a minimum bank angle for some period (e.g., 10° for more than 3 seconds).

In this research, two new "Turn Tracker" algorithms are devised to use the few radar position reports during a Turn Signal episode to estimate the initial position, groundspeed, and direction for the new straight line segment when normal radar tracking is resumed. The algorithms were implemented in a last-time simulation called TASIM, and compared with performance of an existing ATC tracker. The results show a significant reduction in average and maximum deviations of estimated values for groundspeed and direction during the turn, and a faster convergence on good estimates of the new groundspeed and direction along the new straight-line path after the turn.

4. REFERENCES TO PUBLICATIONS, 1992 - 1993

4.1 Liu, Manly, Tracking Aircraft around a Turn with Wind Effects, SM Thesis, and Flight Transportation Laboratory Report 93-4, Flight Transportation Laboratory, MIT, Cambridge, MA June 1993,

4.2 Hansman, R. John; Wanke, Craig; Kuchar, James; Mykitishyn, Mark; Hahn, Edward; Midkiff, Alan, Hazard Alerting and Situational Awareness in Advanced Air Transport Cockpits, Paper at 18th ICAS Congress, Beijing, China, September, 1992

4.3 Wanke, Craig; Hansman, R. John, A Data Fusion Algorithm for Multi-sensor Microburst Hazard Assessment, Preprint, AIAA Atmospheric Flight Mechanics Conference, Hilton Head, SC, August, 1992

4.4 Wanke, Craig; Hansman, R. John, Hazard Evaluation and Operational Cockpit Display of Grand Measured Windshear Data, AIAA Journal of Aircraft, Vol 29, No. 3, May-June, 1992

4.5 Wanke, Craig; Kuchar, Jim; Hahn, Edward; Pritchett, Amy; Hansman, R. John, A Graphical Workstation Based Part-Task Flight Simulator for Preliminary Rapid Evaluation of Advanced Displays, Preprint, SAE AEROTECH Conference and Exposition, Anaheim, CA, October, 1992

Preprints, SAE AEROTECH Conference and Exposition, Anaheim CA, October 5-8, 1992

A Graphical Workstation Based Part-Task Flight Simulator for Preliminary Rapid Evaluation of Advanced Displays

C. Wanke, J. Kuchar, E. Hahn, A. Pritchett, and R. J. Hansman
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA USA

ABSTRACT

Advances in avionics and display technology are significantly changing the cockpit environment in current transport aircraft. The MIT Aeronautical Systems Lab (ASL) has developed a part-task flight simulator specifically to study the effects of these new technologies on flight crew situational awareness and performance. The simulator is based on a commercially-available graphics workstation, and can be rapidly reconfigured to meet the varying demands of experimental studies. The simulator has been successfully used to evaluate graphical microburst alerting displays, electronic instrument approach plates, terrain awareness and alerting displays, and ATC routing amendment delivery through digital datalinks.

INTRODUCTION

The implementation of advanced technology has significantly changed the cockpit environment in current "glass cockpit" aircraft. Recent developments in display technology, on-board processing, data storage, and datalinked communications are likely to further alter the environment in second and third generation "glass cockpit" aircraft. It is important that these technologies be implemented in a manner which will enhance both the human and systems performances, in terms of both safety and efficiency. Because many of the changes in cockpit technology center around information management, proper design of advanced cockpit systems requires careful consideration of the human performance issues, particularly in the cognitive domain.

The MIT Aeronautical Systems Lab (ASL) has developed a part-task flight simulator specifically to study these issues. The simulator, based on a commercially-available graphics workstation, replicates the Electronic Flight Instrumentation System (EFIS), Flight Management Computer (FMC), and primary autoflight systems of a modern "glass-cockpit" aircraft such as the Boeing 757/767 or 747-400. Topics studied using this simulator include graphical displays for hazardous weather information, terrain awareness and alerting displays, datalink of ATC clearance amendments, and electronic approach plates.

The simulator provides high fidelity representations of electronic autoflight and instrumentation systems while remaining low-cost, rapidly reconfigurable, and portable enough to move to alternate sites if necessary. It allows new displays to be developed quickly and evaluated through flight simulations with active airline pilots of electronic cockpit aircraft. This paper discusses the design, advantages, and limitations of this approach.

DESIGN OBJECTIVES AND REQUIREMENTS

The design of the MIT Advanced Cockpit Simulator was motivated by the need for preliminary evaluation of new cockpit information systems. The primary area of interest is the effect of these new systems on human cognitive performance. This area includes such issues as information transfer efficiency, pilot decision-making performance, and flight crew situational awareness.

To evaluate cognitive performance issues, the autoflight systems and primary displays which affect decision-making needed to be simulated as exactly as possible. In addition, the need to test many different prototype displays demanded rapid reconfigurability. These requirements were achieved by simulating the graphical displays on a commercially-available graphics workstation. The simulation software was written by the researchers in modular fashion so that different displays could be implemented by recoding or replacing the appropriate modules.

A further requirement was simplicity. Since only cognitive-level issues were to be evaluated, it was assumed that all aircraft control would be performed using autoflight systems. Therefore, the autoflight and flight management systems needed to be simulated, but the direct flying controls (stick, rudder, throttles, etc.) could be omitted. For this reason no special hardware was required beyond general-purpose computers and some simple control panels, greatly reducing development time and simulator set-up time.

This simplicity also defines the limitations of this approach. Experiments involving flying performance, two- or three-man crews, or requiring a full cockpit workload situation are not practical with this simulator. However, this part-task approach can be useful for preliminary evaluation of candidate displays or procedures before a full-mission simulation is attempted.

Research supported by government grant.

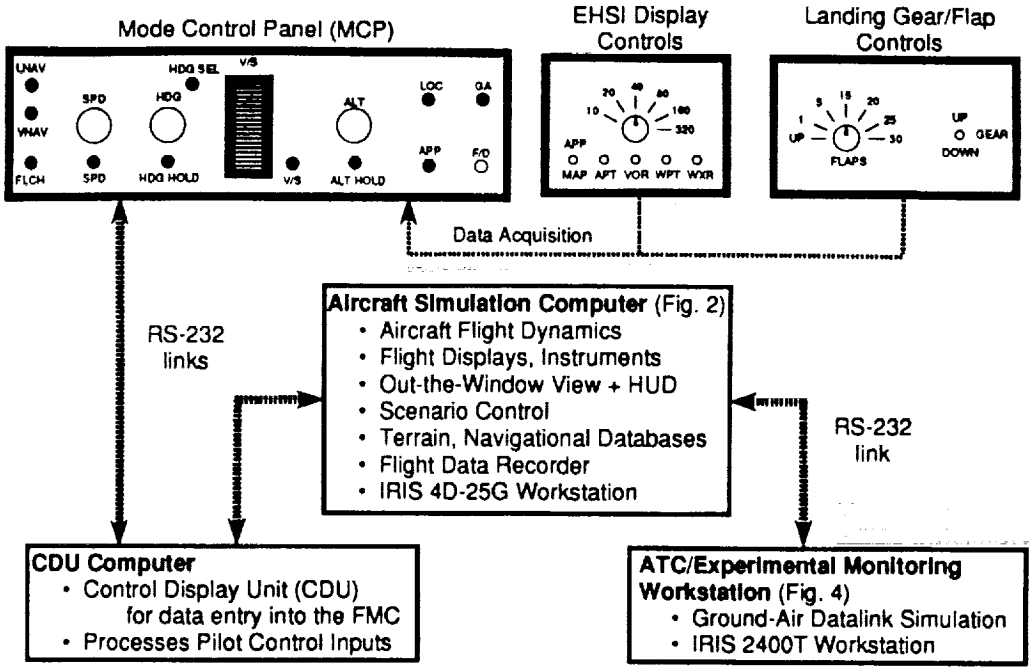


Figure 1: MIT ASL Advanced Cockpit Simulator. The simulator includes three computers and some auxiliary control panels, connected by standard RS-232 serial links.

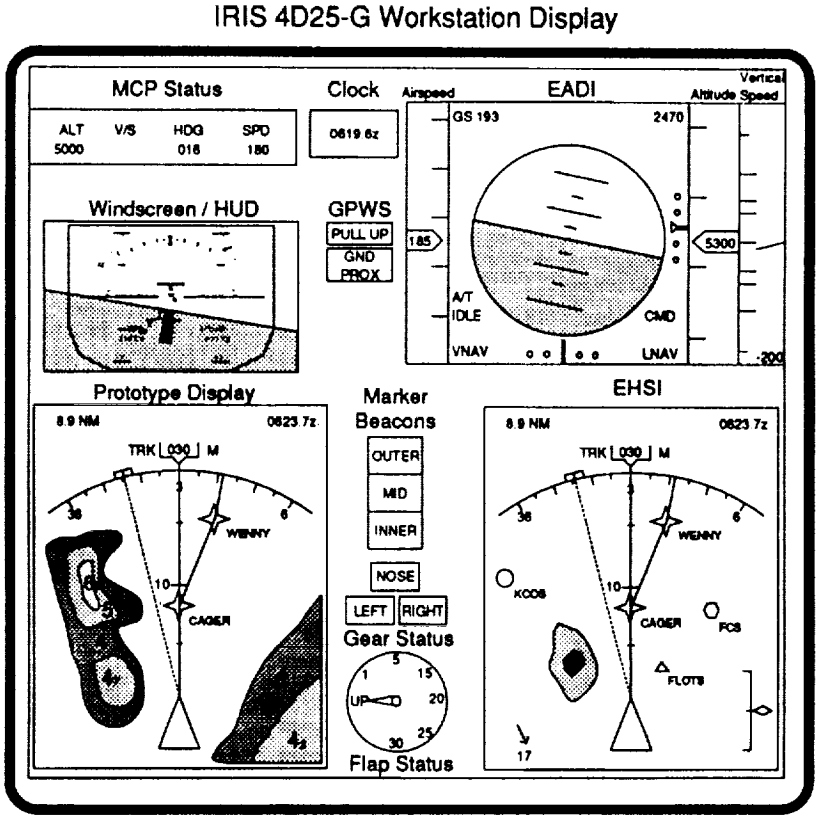


Figure 2: Primary Flight Instrumentation. This is a schematic view of the IRIS 4D-25G display in a typical configuration. Note that the electronic displays are actually in color on the simulator.

THE MIT ADVANCED COCKPIT SIMULATOR

OVERVIEW - As shown schematically in Figure 1, the full version of the MIT ASL Advanced Cockpit Simulator facility utilizes three computers and several control panels to emulate cockpit displays, autoflight systems, and Air Traffic Control (ATC). A Silicon Graphics IRIS 4D-25G graphics workstation is used to display the cockpit instruments (Figure 2) and compute flight dynamics. The Control Display Unit (CDU) is emulated by an IBM-XT computer, and a Silicon Graphics 2400T workstation is used as an Air Traffic Control workstation (Figure 4). The portable version of the simulator omits the ATC workstation. Pilot input through the control panels is detected by the IBM-XT through a data acquisition unit. All three computers exchange data through standard RS-232 serial communication links.

The simulator's cockpit displays are based on current "glass-cockpit" aircraft such as the Boeing 757/767 and 747-400. The IRIS screen depicts two major cockpit displays, the Primary Flight Display (PFD) and Electronic Horizontal Situation Indicator (EHSI), along with several secondary displays. Additional displays can be rapidly prototyped and added to the simulator for evaluation. The nominal flight displays may then be rearranged or modified to accommodate the new displays as needed.

Airspeed, altitude, and vertical speed are indicated on the PFD using moving tape displays similar to those found on the B747-400. An Electronic Attitude Director Indicator (EADI) displays the artificial horizon, ground speed, radio altitude, and Instrument Landing System (ILS) localizer and glideslope deviations.

The EHSI is located below the PFD, as in the B757 or 767. The EHSI is the primary navigational instrument, and the simulator version is based on the map mode used in the B757/767. It includes information such as aircraft heading, ground track, FMC-programmed route, nearby airports and navaids, and wind information. Weather radar returns can also be displayed. A control panel is provided for setting the EHSI display range (10 to 320 nm) and for suppressing weather radar returns or off-track intersections, navaids, or airports.

Flap, gear, and marker beacon light displays are provided to the left of the EHSI. Controls allow the pilot to set the flaps and lower or raise the landing gear during the approach. Additional controls such as a manual pressurization valve can be added to the simulation if a side task is necessary to increase the ambient crew workload.

A simple perspective out-the-window view is provided as a means by which to cue the pilot that the aircraft has descended below the cloud deck. While in instrument conditions, the display appears gray. Below the cloud deck, a perspective view of the airport appears.

A Head-Up Display (HUD) is also available, implemented over the out-the-window view (Figure 3). It uses symbology similar to that used on a commercially available HUD from Flight Dynamics, Inc. Roll, pitch, and heading scales as well as a flight path symbol are displayed in perspective format. Numeric information includes barometric altitude, airspeed, ground speed, vertical speed, and wind information.

AUTOFLIGHT AND FLIGHT MANAGEMENT SYSTEMS - The CDU for entry of flight path information into the Flight Management Computer (FMC) is simulated with an IBM-XT computer. Several screen displays, or "pages", can be selected: The Route page to select a destination, the Legs page to select waypoints and vertical path constraints, and the Direct-To page to change the immediate waypoint. The CDU is linked to the EFIS so that active and modified routes are displayed both textually, on the CDU, and graphically, on the EHSI. At first, the CDU interface used a standard computer keyboard and monochrome monitor. At this time, a replica of the 757/767 CDU display and keyboard is being integrated into the system to enhance realism.

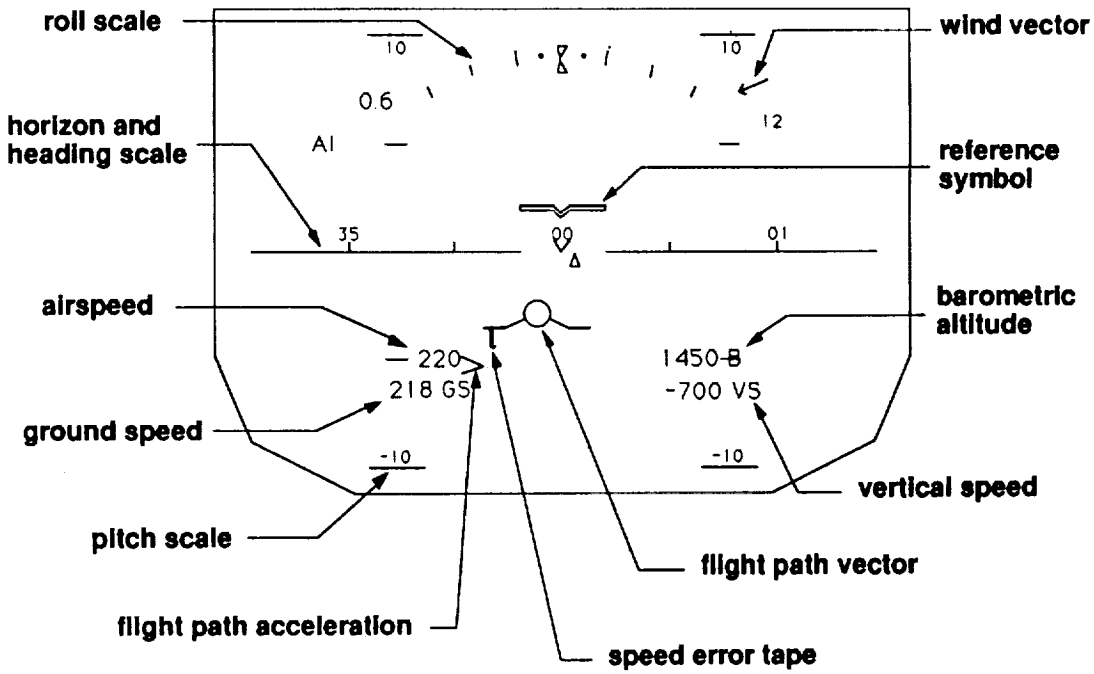


Figure 3: Head-Up Display (HUD). The HUD is overlaid on the windscreen, as seen in Figure 2.

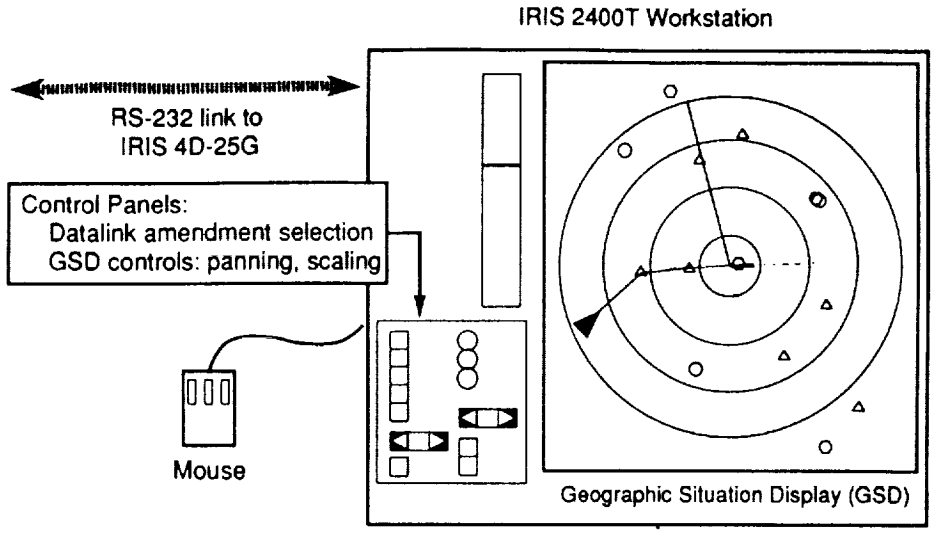


Figure 4: ATC/Experimental Control Workstation Display. The Geographic Situation display provides an airport-centered view of the scenario region, including the position of the simulator aircraft. It can be panned and rescaled by the simulation controller.

Non-FMC control of the aircraft is performed through an autopilot Mode Control Panel (MCP), similar to that used on the Boeing 757/767. A standard set of autothrottle and autoflight modes are available, including LNAV/VNAV flight (following FMC-programmed lateral and vertical flight paths) and the various capture ("select") and hold modes for airspeed, heading, vertical speed, and altitude. It is also possible to engage localizer and glideslope capture modes and a go-around mode for missed approaches.

AIRCRAFT DYNAMICS - The basic aircraft flight dynamics are based on longitudinal point-mass equations of motion in wind axes, and simple decoupled first-order roll angle dynamics. The aircraft data used (provided by NASA Langley Research Center, and used in [1]) is for a Boeing 737-100 aircraft, and includes non-linear curve fits for C_L and C_D as functions of angle-of-attack, flap position, and gear position. The multivariable inner-loop controller designed for this model took the form of a fully-coupled proportional-plus-integral cascade compensator and allows the aircraft to follow airspeed, flight path angle, and heading commands from the autoflight systems.

The autoflight systems provide outer loop control inputs and can operate in several different modes, ranging from simple altitude or heading holds up to full lateral and vertical path guidance based on the FMC programmed route. Localizer and glideslope tracking modes can be engaged for final approach. Because outer-loop controllers for the various autoflight modes are based on approximate frequencies and damping ratios for the Boeing 767 aircraft control system [2], the aircraft responds like a 767 when being controlled through the autoflight systems.

For the experiments including microburst wind shear events, a wind model is available including both constant wind components and simulated microburst winds from an analytical model [3].

ATC/EXPERIMENT CONTROL WORKSTATION - The ATC workstation (Figure 4) is used to monitor the progress of the aircraft's flight and to issue simulated datalink ATC clearance amendments. A mouse-based graphical user interface provides the ability to select and deselect navigational information, determine the aircraft location relative to a scenario reference point, and select and specify content and format of the scripted ATC messages. The controller is in contact with the pilot through a wireless headset, to simulate a standard VHF radio link. Simulated datalink messages are transmitted from the ATC workstation to the simulation computer via a serial communications link. It should be noted that this display was not intended to reproduce any actual or proposed advanced ATC workstation; it was designed only for simulation control.

RAPID PROTOTYPING CAPABILITIES - Commercially available display prototyping software was not used in order to reduce computational overhead. Instead, the flight displays were created using software written in the C programming language with IRIS Graphics Library primitives. This method of implementation allows flight displays to be rapidly reconfigured or redesigned to meet the varying demands of experimental studies. Typically, new displays may be created and added to the simulator in a matter of days.

Additional software was written to enable rapid creation of object-based charts for use with Electronic Instrument Approach Plate studies. Since a detailed object database was not available for use in the Advanced Cockpit Simulator, a software package was developed for the IRIS which facilitated the flexible, rapid creation of new chart display formats [4]. The program, called *Map*, allows the user to interactively create and modify electronic charts. Information may be grouped together in object-oriented layers which are then selectable by the pilot when flying the simulator. Also, a mouse-driven program called *WxrEdit* was developed to draw simulated weather radar reflectivity returns.

Scenarios can be set up and rapidly changed via English- language input files, which are read by the simulator software upon startup. These files define the starting aircraft position and state, pre-programmed FMC information, and scripted events to take place during a run. Scenario files also indicate *Map* and *WxrEdit* files to be loaded at start and during the runs.

EXPERIMENTAL PROCEDURES

In a typical experimental set up, an experimenter acting as air traffic controller is stationed at the ATC/Experimental Control Station and is in contact with the pilot through a simulated VHF link. The controller monitors the progress of the flight and issues vectors and approach clearance amendments according to a script for each scenario.

A second experimenter, acting as the Pilot Not Flying (PNF), is seated next to the subject pilot. In most experiments, the PNF experimenter handles ATC communications and is available to answer any questions about the simulator that occur during the experiment.

The cockpit is videotaped during the experiment to

record ATC and intra-cockpit communications and actions. In addition, the simulator software records all flight data and pilot control inputs for the entire experimental run.

In order to maximize the validity of the results and minimize simulator training requirements, the subject pool is normally limited to professional air carrier pilots currently qualified on autoflight aircraft.

A typical session begins with the pilot completing a brief background questionnaire. The experiment is described briefly, and the subject is introduced to the simulator displays and controls. Practice flights are flown until the pilot feels comfortable with the control of the simulator and its displays. Finally, the pilot is asked to fly the simulator as responsibly as he or she would on a normal flight, and to feel free to request different or additional vectors from ATC, should the need arise.

When the pilot is ready to begin, the appropriate Instrument Approach Plates and charts are issued. Airport information (ATIS) is also given to the pilot to describe weather conditions and other information usually received before an approach. Scenarios typically begin 50 to 150 nm from the destination airport with an initial route programmed into the aircraft's FMC and displayed on the EHSI.

After the pilot reviews the charts and feels comfortable with the situation, the simulation is started. Amendments to the programmed route are issued by the air traffic controller and the pilot may control the aircraft through the Mode Control Panel or by programming the FMC through the CDU. A typical test matrix would require that each pilot fly 9 to 12 descent-and-approach scenarios. Most experimental scenarios are set in the terminal area when the flight crew is busiest and handles the most information. The entire session takes three to four hours to complete.

When possible, the independent variables in each study are counterbalanced to reduce learning effects. Implicit measures of display efficacy are obtained by observing pilot reactions to scripted events that occurred during the flight, such as a vector into weather or a graphical microburst alert. In addition, subjective data is obtained through interviews with pilots both during and after the experiment.

SURVEY OF SIMULATOR EXPERIMENTS

Several studies involving cockpit information management have used the part-task simulator facility. The following list is a very brief description of several recent projects, which highlight the advantages of the simulator; the authors or references should be consulted for complete details on the experimental methods and results. Note that the figures in this section are schematic line drawings of color displays, and therefore lack some of the detail present on the actual displays.

Graphical microburst alerting displays. [5] Several different candidate displays for presenting microburst alerts were evaluated with the simulator (Figure 5). The rapid prototyping feature of the simulator was particularly useful in this study for design of several candidate displays. Also, the simulator was moved by van in order to do the experiment in a city with both an available facility and a large subject population, highlighting the advantage of portability.

Electronic instrument approach plates. [6] The advent of electronic library systems has made it possible to present charts electronically in the cockpit. This experiment compared several different possible formats and issues for electronic instrument approach plates (EIAP). The *Map* software

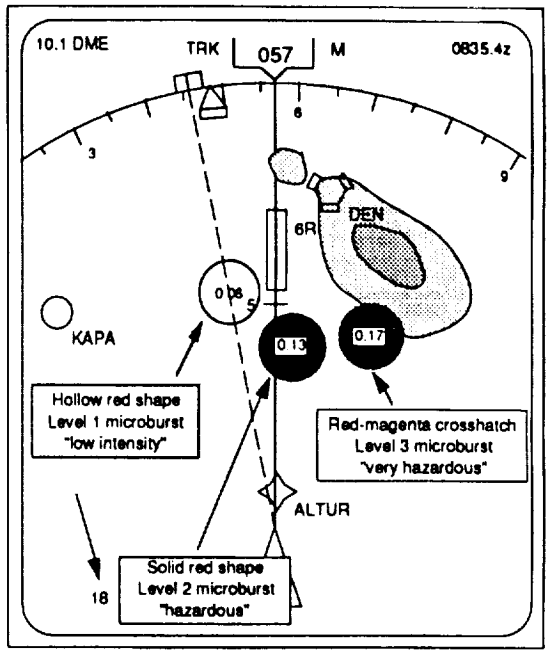


Figure 5: Graphical Microburst Alerting Display. Microburst alert icons are displayed directly on the EHSI display.

package was used to rapidly design these chart formats (Figure 6), which were then loaded and displayed by the simulator software. Chart information is grouped by type into layers, which can be selected or suppressed by the pilot with a switch panel similar to the EHSI control panel.

Terrain awareness displays. [7] Another application of electronic library systems is the presentation of terrain information, either as part of a ground proximity warning system or as a situational awareness display. One possible terrain awareness display (Figure 7) could present shaded contours. This display was also produced by the *Map* software, and was compared to more traditional spot elevation terrain representations in a piloted simulator study.

Graphical ATC datalink amendments. [8] The delivery of ATC clearance amendments through a digital air-ground datalink holds the potential to reduce voice congestion and information transfer errors associated with VHF radio communications. The ATC workstation (Figure 4) is linked directly with the FMC and the IRIS workstation to send datalink messages in either textual or graphical modes, and can directly reprogram new routings into the FMC if required. An experiment compared the effects of several types of datalink ATC amendment presentations on flight crew situational awareness. Figure 8 shows a datalink ATC amendment which has been delivered and displayed on the EHSI.

Topics for future experiments include HUDs, displays for airborne forward-looking wind shear sensors, continued study of terrain avoidance displays, and study of applications of digital ground-air datalinks.

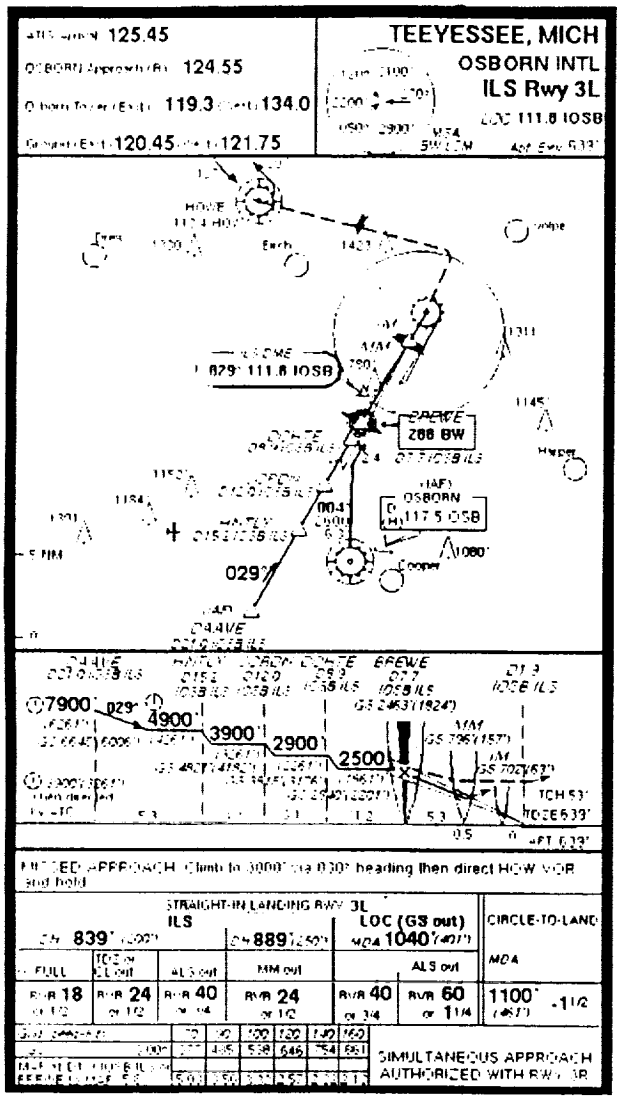


Figure 6: Sample Electronic Instrument Approach Plate (EIAP) format. On this display, information is grouped into color-coded layers and can be hidden or selected by the pilot.

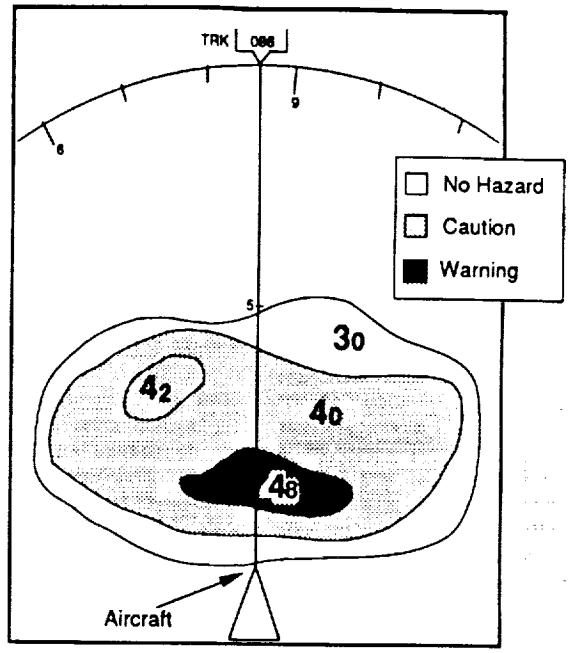


Figure 7: Contour Terrain Awareness Display with Graphical Ground Proximity Warning System (GGPWS). Terrain contours change to yellow or red colors when alerting criteria are satisfied.

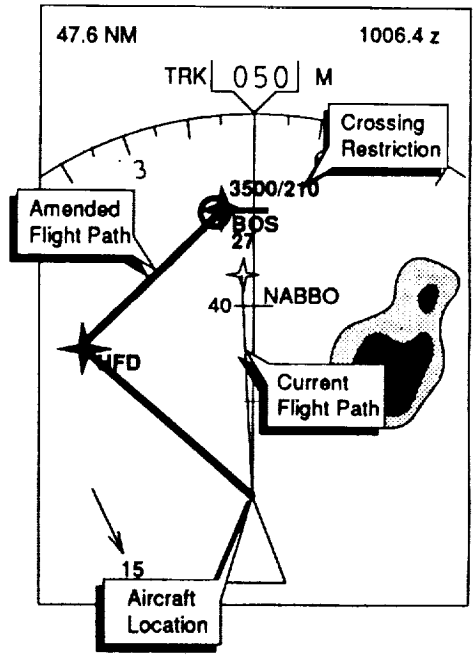


Figure 8: Graphical ATC Route Amendment. The bold line represents a new routing delivered by digital datalink; it flashes until accepted or rejected by the pilot.

CONCLUSIONS

A part-task advanced cockpit simulator has been developed to evaluate the effect of advanced cockpit information management systems on pilot cognitive performance. The utility of this part-task approach for rapid preliminary evaluation of new graphical displays and new datalink applications has been demonstrated through a series of successful experiments.

The MIT Advanced Cockpit Simulator replicates the major autoflight and electronic flight instrumentation systems of a modern "glass-cockpit" transport aircraft, but does not include manual flight controls or a cockpit mock-up. This simplicity reduces set-up time, cost, and allows the facility to be easily moved. Since the simulator is based on a commercially-available graphics workstation, it can be rapidly reconfigured and does not require special hardware. In order to maximize the validity of the results, the subject pool is limited to professional air carrier pilots currently qualified on "glass-cockpit" aircraft.

Concepts evaluated using this simulator include graphical microburst alerting displays, electronic instrument approach plates, terrain awareness and alerting displays, and ATC routing amendment delivery through digital datalinks. Topics for future experiments include HUDs, displays for airborne forward-looking wind shear sensors, and continued study of terrain avoidance displays and issues associated with digital ground-air datalinks.

ACKNOWLEDGEMENTS

Support for the research activities discussed above was received from the Federal Aviation Administration, National Aeronautics and Space Administration, DOT Volpe National Transportation Systems Center, Lincoln Laboratory, and the National Science Foundation Presidential Young Investigator Award Program under the following grants; MSS-8552702, NGL-22-009-640, BARR-10-119, NAG 2-12, NAG 2-716, DTRS 57-88-C-0078TD39, NAG-1-690. The simulation activities were aided by cooperation of the Air Line Pilots Association, the Allied Pilots Association, American Airlines, Delta Airlines, and United Airlines. The authors would especially like to thank the many subject pilots who participated in the experiments.

REFERENCES

[1] Hinton, David A., *Relative Merits of Reactive and Forward-Look Detection for Wind-Shear Encounters During Landing Approach for Various Microburst Escape Strategies*, NASA TM-4158, DOT/FAA/DS-89/35, February, 1990.

[2] Personal Communication with Arun Nadkarni, Boeing Commercial Airplane Company, August 1988.

[3] Oseguera, Rosa M. and Bowles, Roland L., *A Simple Analytic 3-Dimenstional Model Based on Boundary Layer Stagnation Flow*, NASA TM-100632, NASA Langley Research Center, Hampton, VA, July 1988.

[4] Kuchar, J.K., *'MAP' - Electronic Approach Plate Prototyping Software Documentation*, MIT Aeronautical Systems Laboratory Report ASL-90-6, October 1990.

[5] Wanke, C., and Hansman, R. J., "Experimental Evaluation of Candidate Graphical Microburst Alert Displays," AIAA Paper 92-0292, 30th Aerospace Sciences Meeting & Exhibit, Reno, NV, January 1992.

[6] Mykityshyn, M. and Hansman, R.J., *An Exploratory Survey of Information Requirements For Instrument Approach Charts*, MIT Aeronautical Systems Laboratory Report ASL-90-1-2, November 1990.

[7] Kuchar, J., and Hansman, R.J., *Advanced Terrain Displays for Transport Category Aircraft*, MIT Aeronautical Systems Laboratory Report, ASL-91-3, August 1991.

[8] Hahn, E. and Hansman, R.J., "Experimental Studies on the Effect of Automation on Pilot Situational Awareness in the Datalink ATC Environment," SAE Aerospace Technology Conference and Exposition '92, Anaheim, CA, (to be presented October 1992).

Preprint, AIAA Atmospheric Flight Mechanics Conference, Hilton Head, SC, USA, August 10-12 1992

A Data Fusion Algorithm for Multi-Sensor Microburst Hazard Assessment

Craig R. Wanke* and R. John Hansman**
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA USA

Abstract

A recursive model-based data fusion algorithm for multi-sensor microburst hazard assessment is described. An analytical microburst model is used to approximate the actual windfield, and a set of "best" model parameters are estimated from measured winds. The winds corresponding to the best parameter set can then be used to compute alerting factors such as microburst position, extent, and intensity. The estimation algorithm is based on an iterated extended Kalman filter which uses the microburst model parameters as state variables. Microburst state dynamic and process noise parameters are chosen based on measured microburst statistics. The estimation method is applied to data from a time-varying computational simulation of a historical microburst event to demonstrate its capabilities and limitations. Selection of filter parameters and initial conditions is discussed. Computational requirements and datalink bandwidth considerations are also addressed.

Introduction

Low altitude wind shear has been a major cause of fatal aviation accidents in the U.S.[1] The localized intense downdrafts known as microbursts are the most dangerous form of wind shear, and pose a serious hazard to aircraft during takeoff or approach. In a typical microburst encounter, an aircraft first encounters a performance-increasing headwind. This is followed by a downdraft and a rapid transition from headwind to tailwind, which produce sharp losses in altitude and/or airspeed.

Several systems for detection and measurement of microburst hazards are currently nearing the operational stage. Effective ground-based systems such as Terminal Doppler Weather Radar (TDWR) and the extended Low Level Wind Shear Alert System (LLWAS) are entering the deployment phase. TDWRs will be located at 47 major airports, and detect microbursts primarily by measuring the surface wind velocity component radial to the radar and identifying areas of radial shear.[2,3] LLWAS is a network of anemometers which measure horizontal windspeed and direction around the airport surface, and detect wind shear events from differences in wind speed and direction between sensors.[4] Airborne reactive wind shear alerting systems, currently in use, detect microburst penetration by comparing inertial and air data system measurements to compute the local winds. Several types of airborne forward-look sensor technologies are also under development, including infrared radiometry, Doppler radars and Doppler lidars.[5] Infrared systems measure the drop in temperature associated with the air in the center of a microburst, while Doppler radars and lidars measure wind velocities along the flight path ahead of the aircraft. In addition to new sensor developments, the development of digital air-ground datalink capabilities such as the Mode-S beacon system will allow microburst alert information to be exchanged between air and ground-based systems (Figure 1).[6]

As new detection systems become operational, it will become likely that more than one sensor system will be available in a given situation. Also, each of the aforementioned sensor systems has some geometrical observability problems. For example, both ground-based and airborne Doppler radars and lidars can only measure wind velocities radial to the sensor, not vertical winds. The aviation hazard posed by a microburst, however, is due to both horizontal wind shear and downdrafts in the microburst core. Therefore, a technique for combining data from different systems with different measurement characteristics could improve estimates of microburst hazards and aid alert generation.

The goal of this "data fusion" process is to provide a microburst detection and hazard assessment capability which is significantly better than that which can be achieved using a single sensor. The data fusion algorithm must provide *appropriate information* for alert generation, in a timely fashion, and be feasible with regard to the available air-ground datalink bandwidth and computational capabilities. Previous work at MIT has focused on

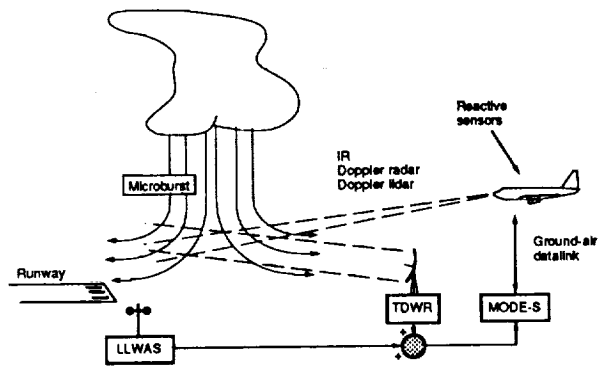


Figure 1: Advanced Microburst Detection and Alerting Systems

* Research Assistant

** Associate Professor, Associate Fellow AIAA

Research supported by government grant.

definition of appropriate information for microburst alerts. This work has included analytical studies to determine appropriate microburst hazard criteria,[7] and piloted part-task simulator studies to determine the information requirements of the flight crew when faced with a microburst alert situation.[8,9] The conclusions of these studies were that estimates of microburst location, extent, and approximate intensity are required for alerting purposes, and that microburst intensity can be quantified well using criteria which relate to the expected aircraft energy loss due to the microburst windfield. Therefore, a good data fusion algorithm should be able to compute the position and extent of a microburst, and to estimate intensity including contributions from both horizontal wind shear and downdrafts.

The data-fusion process can be done on a number of levels. One approach is to merge the final products of the sensor systems to produce improved alerts. For example, product-level techniques have been used to integrate TDWR and LLWAS information[10] and to determine the probability of hazardous wind shear given a wide range of evidence.[11] Another approach is to integrate sensors on the data level. The data-level approach is more complex, due to the large volume of data produced by several wind shear sensor systems. However, if correctly implemented, observability problems due to poor sensor geometry can be alleviated. Data from multiple sensors can be combined to form a "super sensor" which has improved sensing geometry. The technique proposed in this paper is a model-based data-level approach which attempts to gain this observability benefit without prohibitively large computational or data transmission requirements.

Model-Based Approach

Why Use a Model?

Representation of the actual microburst windfield with an analytical model has two major advantages. Firstly, once the model has been "fitted" to the windfield, the windfield can be approximated by the values of the model parameters. Thus, if the model represents the windfield well enough, the measured information (a large data set) can be encapsulated in a small set of "best-fit" model parameters. Since it is impractical (or at least undesirable) to transmit raw data between aircraft and ground-based systems, this is an important advantage.

Secondly, an analytical model can include additional information which can be used to infer quantities which cannot be directly measured, such as inferring vertical velocities from radar-measured radial velocities. Analytical models can be designed to satisfy basic fluid dynamic relationships such as mass continuity, and can be adjusted to reflect results of microburst field measurements.

Analytical Microburst Model

The analytical microburst model used in this work was developed at NASA Langley Research Center initially by Oseguera and Bowles,[12] and later improved by Vicroy.[13] The Oseguera-Bowles-Vicroy (OBV) model uses shaping functions to generate an axisymmetric flowfield which satisfies the mass continuity equation and is representative of the major characteristics of measured microbursts. Sample winds for a constant-altitude path through the model windfield are shown in Figure 2. The horizontal winds exhibit the classic microburst characteristic of a headwind increase, followed by rapid shearing to a tailwind. The vertical wind plot shows a downdraft in the microburst center and smaller updrafts at the edges.

The microburst winds are uniquely defined by a set of five parameters and three empirically-adjusted constants. For this study, a simple ambient wind (4 additional parameters) was added to the microburst windfield. The model parameters are summarized in Table 1. The total winds are given by non-linear, smooth, differentiable functions of the parameters and a given (x,y,h) position as follows:

$$U = U_{microburst} + U_0 + U_h h \quad (1)$$

$$V = V_{microburst} + V_0 + V_h h \quad (2)$$

$$W = W_{microburst} \quad (3)$$

U, V, and W are the Eastward, Northward, and vertical wind velocities; h is the altitude above ground level. "Microburst" quantities are functions of position and of the first five parameters in Table 1; these functions are summarized in Appendix A.

The OBV model is axisymmetric, but naturally occurring microbursts are often asymmetric.[14] In addition, multiple microbursts have been observed to occur close together and interact. To handle these cases, the model was extended to allow multiple interacting microbursts. For each microburst, another set of five microburst model parameters (the first five in Table 1) can be added. It is assumed that the ambient wind will be

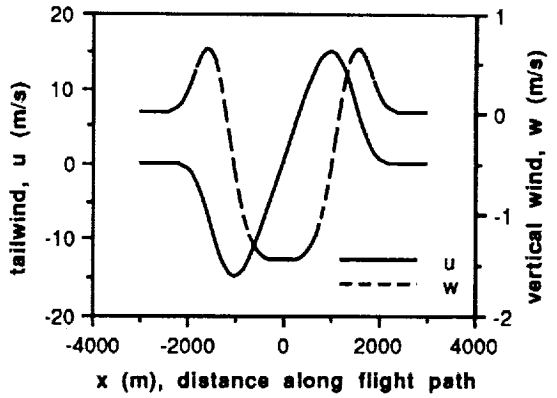


Figure 2: Microburst Model Windfield. Sample winds for a constant-altitude flight path through the center of a simulated microburst.

Table 1: Modified Oseguera-Bowles-Vicroy Microburst Model Parameters

Parameter	Description
x_0	X-coordinate (East) of microburst center (m)
y_0	Y-coordinate (North) of microburst center (m)
U_m	Maximum horizontal outflow speed (m/s)
R_p	Radius of maximum outflow (meters)
Z_m	Altitude AGL of maximum outflow (meters)
U_0	Eastward ambient wind constant component (m/s)
U_h	Eastward ambient wind altitude gradient (m/s/m)
V_0	Northward ambient wind const. component (m/s)
V_h	Northward ambient wind altitude gradient (m/s/m)

roughly constant throughout the x-y space of interest, i.e. near the airport, and so only one set of ambient wind parameters is used. The winds from each model microburst are summed to get the overall model windfield; this superposition does not violate mass continuity. In the simulation study below, when a "two-microburst" model is referred to it does not necessarily indicate that there are two microbursts being detected, but that two superimposed model microbursts are being used to simulate a complex microburst windfield with more than one area of high downdraft speed.

Model-Based Multi-Sensor Data Fusion

Given a suitable model, the fusion problem reduces to estimation of the "best" set of model parameters based on all available wind measurements. Once the best parameters have been estimated, alerting factors (intensity, extent etc.) can be derived from the analytical model windfield corresponding to the estimated model parameters.

This estimation procedure must satisfy several constraints to be practical. The estimation algorithm must be recursive, to handle new measurements as they become available. It must also account for time variation in the model parameters, since microbursts are dynamic phenomena with short lifetimes on the order of 15 minutes and sensor measurements will be taken at different times. It should also be probabilistic, to take advantage of microburst statistical characteristics from past field studies. A Kalman filter approach is proposed to satisfy these requirements.

Iterated Extended Kalman Filter Algorithm

Estimation Problem Structure

Kalman filtering techniques require a state-space dynamic model of the system and a relationship between system parameters and measured quantities. In this case, we would like to estimate analytical model parameters which best describe the microburst from measurements of the winds. The analytical microburst model parameters are therefore used as the filter state variables x(t). It was assumed that the time evolution of the microburst parameters can be adequately modeled by a linear, time-invariant, continuous-time system:

$$\dot{x}(t) = A\,x(t) + B\,u(t) + L\,w(t) \tag{4}$$

Deterministic inputs to the system are represented by u(t), and w(t) is a white Gaussian process noise input. The A, B, and L matrices define the dynamic model; they will be discussed below. Since the state variables x(t) are the analytical model parameters, they are related to the wind measurements through the analytical model wind equations. The resulting non-linear discrete-time measurement equation is:

$$z_k = h_k(x(t_k)) + v_k \tag{5}$$

where the measurement equations h_k are simply the wind equations from the analytical model, and v_k represents measurement noise. The state vector, x, and error covariance matrix, P, for a single downdraft case are defined as follows:

$$x = [x_0\ \ y_0\ \ U_m\ \ R_p\ \ Z_m\ \ U_0\ \ U_h\ \ V_0\ \ V_h]^T \tag{6}$$

$$P = E\left[(x - \hat{x})(x - \hat{x})^T\right] \tag{7}$$

where $\hat{x}$ is the current parameter estimate. The microburst eastward core location, x_0, should not be confused with x, the state vector. Process noise, w(t), and measurement noise, v_k, are white and gaussian with the following characteristics:

$$E\left[(L\,w(t))(L\,w(\tau))^T\right] = L\,Q(t)\,L^T\,\delta(t - \tau) \tag{8}$$

$$E\left[v_k\,v_k^T\right] = R_k \tag{9}$$

The aim of the filter is to produce the state estimate $\hat{x}$ which minimizes the error covariance P. Since the measurement equation is non-linear, this cannot be done with a standard Kalman filter algorithm. An extended Kalman filter (EKF) approach was chosen. The structure and principal equations for the EKF are briefly described below, based on the formulation given in Ref. 15. The filtering algorithm for discrete-time measurements is a two step process: (1) apply the system dynamic model to propagate the state estimate and state estimation error covariance between measurements, and (2) update the estimate when new measurements arrive.

Estimate Propagation: Microburst Dynamic Modeling

For linear, time-invariant, continuous-time system dynamics the propagation of the state estimate and estimation error covariance between measurements is governed by:

$$\dot{\hat{x}}(t) = A\,\hat{x}(t) + B\,u(t) \tag{10}$$

$$\dot{P}(t) = A\,P(t) + P(t)\,A^T + L\,Q(t)\,L^T \qquad (11)$$

The A, B, L, and Q matrices define the microburst time-evolution dynamics. Since the analytical model is time-invariant, these parameters must come from another source. Unfortunately, there is no simple time-varying analytical model available. However, measured microburst statistics can be used to approximate some dynamics. For example, microburst radial extent tends to increase steadily throughout the microburst lifecycle. Analysis of data from Colorado microburst measurements[16,17] indicates that the change in radial extent vs. time can be approximated by a constant bias (a) with additive white zero-mean gaussian noise (n):

$$\dot{R}_p = a + n(t) \qquad (12)$$

where a = 0.102 km/min and the noise term has a standard deviation of 0.15 km/min. The constant bias is treated as a deterministic input, and the noise term leads to a value for one element in the Q matrix. Similar modeling may be possible for some of the other state variables. For example, motion of the microburst core $(\dot{x}_o, \dot{y}_o)$ may be related to the ambient wind parameters, which would lead to non-zero entries in the A matrix. In the simulations discussed below, the A matrix was assumed to contain all zeros. The B and Q matrix elements were set based on statistical information where possible, and from engineering judgement when no statistical information was available. Further research on microburst dynamics is currently in progress.

Since the time behavior of the microburst parameters is not well modeled, significant process noise is required. The use of process noise to compensate for modeling deficiencies is similar to the well-known technique of applying a "forgetting factor" to older data in a batch least-squares formulation. In any case, these simple dynamics lead to sparse A, B, L, and Q matrices, and the propagation step in the filter requires little computation.

Incorporating Measurements

When new measurements are taken, the estimate is updated. The non-linear measurement equation, however, makes the update process difficult. The formulation presented here is based on the extended Kalman filter update with the addition of a local iteration procedure to reduce the effects of the measurement non-linearities.[15,18] At time t_k, a local iteration (over i) is performed. The i^{th} parameter estimate at time t_k, $\hat{x}_{k,i}$, is updated with the following expression:

$$\hat{x}^+_{k,i+1} = \hat{x}^-_k + K_{k,i}\left[z_k - h_k\left(\hat{x}^+_{k,i}\right) - H_k\left(\hat{x}^+_{k,i}\right)\left(\hat{x}^-_k - \hat{x}^+_{k,i}\right)\right] \qquad (13)$$

$$\hat{x}_{k,0} \equiv \hat{x}^-_k\,, \quad i = 0,1,... \qquad (14)$$

The Kalman gain, K, is ordinarily computed from:

$$K_{k,i} = P^-_k\,H^T_k\left(\hat{x}^+_{k,i}\right)\left[H_k\left(\hat{x}^+_{k,i}\right)P^-_k\,H^T_k\left(\hat{x}^+_{k,i}\right) + R_k\right]^{-1} \qquad (15)$$

and H_k is the locally linearized measurement matrix:

$$H_k(\hat{x}) = \left[\frac{\partial h_k(x)}{\partial x}\right]_{x=\hat{x}} \qquad (16)$$

In the above expressions, $\hat{x}^-_k$ and P^-_k indicate the propagated estimate and error covariance at time t_k (prior to updating), while $\hat{x}^+_k$ and P^+_k indicate the updated estimate and covariance based on the measurement z_k. The local iteration is repeated until the scaled norm of the parameter estimate does not change significantly. After the new estimate has been produced, the updated error covariance matrix is computed using values from the final iteration step:

$$P^+_k = \left[I - K_{k,i}\,H_k\left(\hat{x}^+_{k,i}\right)\right]P^-_k \qquad (17)$$

Some simple testing, in which winds generated directly from the OBV model were "identified" using this algorithm, indicated that the iterated filter results in significantly better estimates than the standard EKF; this has also been found by other investigators.[19] A probabilistic interpretation of this iteration based on Bayesian maximum likelihood estimation is given in Ref. 18.

One difficulty with the above updating algorithm is that there may be large numbers of measurements available at a single time step (as in TDWR data, for example), and the computation of the Kalman gain (Eqn. 15) requires inversion of an r-by-r matrix, where r is the number of measurements. The number of computations required to do this scales as r^3. In a linear filter, a large batch of measurements can be treated as a series of sequential scalar measurements (occurring at infinitesmal time spacing) without loss of information, thereby avoiding this problem. When the measurement equation is non-linear, the measurements cannot be incorporated sequentially without losing a significant amount of information. Therefore, an alternate form of the gain computation is required. When the number of measurements exceeds twice the number of states, and the measurement noises are independent (diagonal R_k) it is more efficient to use the "information form" of the gain computation:

$$\left(P^{-1}_k\right)^+ = \left(P^{-1}_k\right)^- + H^T_k\,R^{-1}_k\,H_k \qquad (18)$$

$$K_k = \left(P^{-1}_k\right)^+ H^T_k\,R^{-1}_k \qquad (19)$$

This form can be readily applied to the iterated EKF update described above. Although the covariance update must now be done inside the loop, the required matrix inversion is only n-by-n, where n is the number of states (model parameters). The computational requirement now scales linearly with r and cubically with n. In the

simulation cases below, where r > 100 and n = 9 or 14, this form was found to be much more efficient.

Multiple-Microburst Form

As discussed above, several model microbursts can be superimposed to simulate a more complex windfield. In this case, 5 new states are added for each additional model microburst. For i microbursts, the full state vector, x, is defined as:

$$x = \left[x_{mb,1}^T \; x_{mb,2}^T \cdots x_{mb,i}^T \; U_0 \; U_h \; V_0 \; V_h \right]^T \tag{20}$$

where:

$$x_{mb,i} = \left[x_{0,i} \; y_{0,i} \; U_{m,i} \; R_{p,i} \; Z_{m,i} \right]^T \tag{21}$$

In the simulations discussed below, one- and two-microburst forms are used.

Initialization

This algorithm can incorporate multi-sensor data, given that the microburst has previously been detected. The assumption is made that a single sensor has detected the event and has produced an initial parameter estimate and associated error covariance. The initialization algorithm therefore depends on the measurement characteristics of the initial sensor. The general process, however, is the same for all sensors. Quantities that can be directly measured are estimated from the initial data set, and quantities which are unobservable are initialized using statistics derived from microburst field studies.

For example, if a TDWR initially detects a microburst, estimates of maximum outflow speed (U_m), outflow radius (R_p), and core position (x_0,y_0) can be derived from the radial flowfield measured by TDWR. The outflow depth Z_m is an unobservable parameter, and must be initialized from statistics. Outflow depth statistics have been measured for 26 Colorado microbursts,[20] and the mean altitude of maximum outflow velocity was found to be 109 meters. This value was used to initialize the filter for the simulations discussed below, in which TDWR was always assumed to make the initial detection. The initial covariance matrix was diagonal, and values were chosen based on sensor resolution criteria or statistics where possible.

Simulation Results

Figures of Merit

As mentioned above, the important quantities for alerting purposes are position, extent, and approximate intensity. The "effectiveness" of the proposed algorithm can be defined in terms of its capability to produce these quantities. Therefore, two figures of merit were defined. The first concerns position and extent. Given a center point, an "extent polygon" can be drawn for a microburst windfield (example shown in Figure 3).[14] The vertices of the polygon correspond to the points of maximum radial outflow speed (measured radially outward from a center point). This polygon encloses the entire performance-decreasing portion of the microburst. The ability of the model-based algorithm to define this hazardous region can then be evaluated by comparing the extent polygon A of the truth windfield to the extent polygon B of the analytical model windfield corresponding to the estimated parameters:

$$M_{extent} = \frac{A \cap B}{A \cup B} \tag{22}$$

This quantity has a maximum value of 1 (for an exact match) and falls off for both underestimation and overestimation of the extent boundaries (Figure 4). Core position errors are also reflected, since the model extent polygon is then laterally translated with respect to the true extent polygon. This quantity is a function of altitude, but the dependence was found to be very weak and only results for a single altitude are presented in this paper.

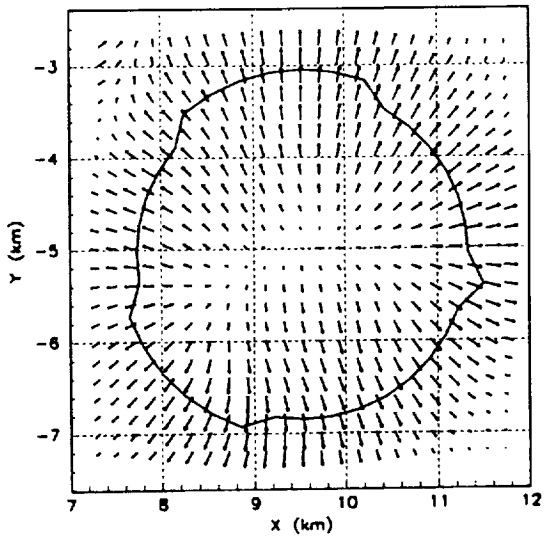


Figure 3: Microburst extent polygon. Horizontal windfield with outflow extent polygon superimposed.

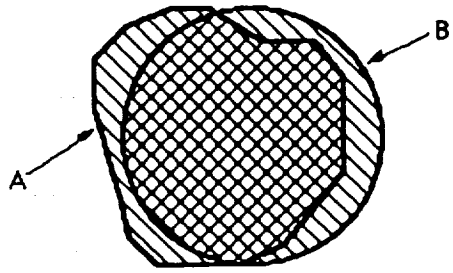


Figure 4: Extent Figure-of-Merit. Pictorial representation of Eqn. 22. The cross-hatched area is $A \cap B$ and the sum of the cross-hatched and striped areas is $A \cup B$. Note that the model "polygon" B is a circle for the single-microburst axisymmetric OBV model.

Microburst intensity was defined in terms of "F-factor", proposed by researchers at NASA Langley Research Center, which is based on the impact of a microburst windfield on the total energy (kinetic plus potential) of the aircraft.[5] It is a measure of the loss of potential rate-of-climb (or loss of effective thrust-to-weight ratio) due to the immediate windfield. It is dependent on the time rate of change in the aircraft frame of the tailwind velocity, the vertical wind velocity, and the aircraft airspeed. Positive values of F indicate a performance-decreasing situation, and negative values indicate a performance-increasing situation. As typical transport-category aircraft in landing or takeoff configuration have excess thrust-to-weight ratios between 0.1 and 0.15, an encounter with an F-factor in excess of that value would compel the aircraft to descend and is therefore hazardous.

$$F \equiv \frac{\dot{W}_x}{g} - \frac{W_h}{V} \qquad (23)$$

F, a point measurement, needs to be averaged over a distance to give a useful indication of aircraft hazard. Past work has determined that F averaged over 1 km of the aircraft flight path yields a good hazard estimate.[7] For evaluation purposes, however, it is desirable to assign a single hazard number to a microburst rather than one for each possible flight path through it. Therefore, for this work, the hazard number was defined as follows: (1) compute 1 km average F-factors for a large number of parallel constant-altitude paths through the microburst, (2) average the resulting values 500 m laterally across flight paths, and (3) pick the largest averaged F-factor as the hazard value.

This value depends strongly on the direction of the flight paths along which F is evaluated. In the simulation results presented below, averaged F-factors will be presented for either eastbound or northbound flight paths. In addition, F depends on altitude, and results will therefore be presented for several altitudes. For alerting purposes, however, it would be necessary to assign a single intensity value to a detected microburst, for example the largest value (over all directions) below a specified maximum altitude.

Simulated Microburst True Windfield

The windfield data used to evaluate the estimation algorithm was generated by the Terminal Area Simulation System (TASS).[21] It is a highly detailed computational simulation of a complex multiple microburst event which occurred at Denver-Stapleton airport on July 11, 1988. This event caused one near accident and a total of five aircraft to make missed approaches.[22] Windfield data from five times during this event was available, with a horizontal spacing of 200m and a vertical spacing of approximately 80m. For the following analyses the largest microburst in the event was selected (Figure 5). The horizontal windfield has a classic microburst outflow pattern. However, the vertical wind contours show some complex structure as indicated by two separate regions of high downdraft, neither of which correspond to the apparent horizontal windfield center (marked with an X). This rather complex event was chosen to test the estimation algorithm in a challenging but realistic situation.

OBV Model Best Fit

The first step for algorithm evaluation was to determine the ability of the analytical model to match the important characteristics of a microburst windfield, namely the figures of merit defined above. This was done with a deterministic non-linear batch least-squares optimization algorithm, similar to that used in Ref. 23 to model microburst winds with vortex rings. The "truth"

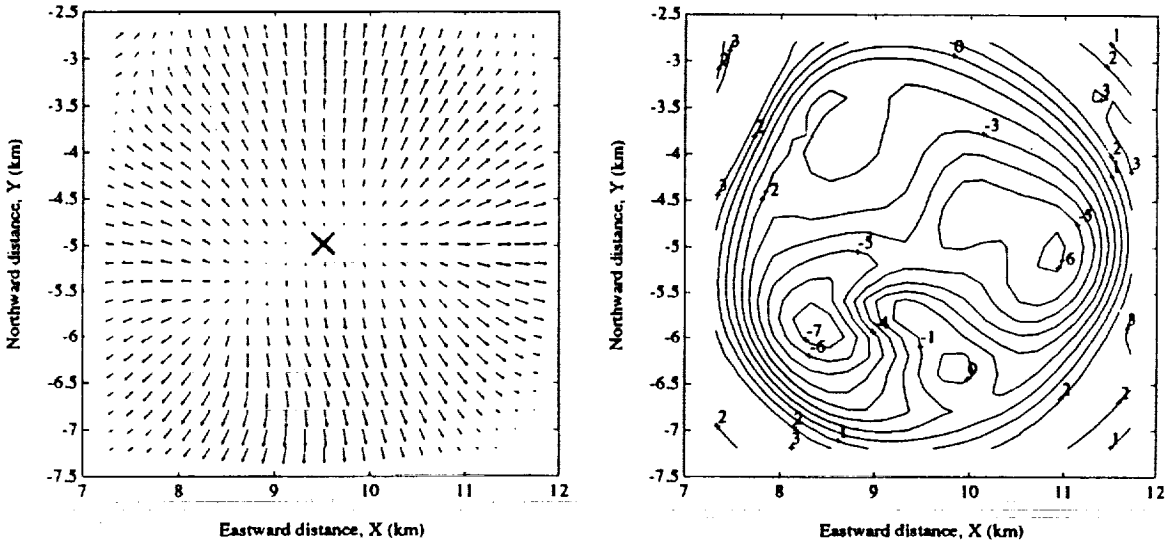


Figure 5. TASS-simulated windfield for 11 July 1988 microburst event at DEN. At left is a vector plot of horizontal winds: maximum velocity shown is 18.7 m/s. At right: vertical windspeed contours. Altitude shown is 177m AGL.

winds were taken from one time-step of the TASS model (shown in Figure 5). The data volume included three-axis winds at three altitudes with 200m lateral spacing for a total of approximately 5000 data points. The optimization procedure was a constrained version of the standard Gauss-Newton method,[24] and found the model parameters which minimized the mean square wind error, J:

$$J = \frac{1}{N}\sum_{j=1}^{N} e_j^T(p)\, e_j(p) \;:\; e(p) = V_{truth} - V_{model}(p) \quad (24)$$

where N is the number of total data points and V is the vector of all wind points, including East, North, and vertical components at all (x,y,z) locations. This procedure was done with a single-microburst model (9 parameters) and with a two-microburst model (14 parameters). The resulting parameters for the single-microburst case are given in Table 2. Note that the approximate radius of this microburst is 1700 meters, and the maximum outflow speed is approximately 18 m/s. The ambient wind magnitude is small in this case.

Table 2: Single-microburst least-squares parameter fit results

x_0 (m)	9528	U_0 (m/s)	0.9
y_0 (m)	-5047	U_h (m/s/m)	-0.001
U_m (m/s)	17.8	V_0 (m/s)	0.5
R_p (m)	1717	V_h (m/s/m)	-0.002
Z_m (m)	68.2		

The single-microburst fit produced an extent figure of merit of 0.92. The two-microburst fit result was slightly lower, at 0.85. As seen in Figure 5, this microburst was fairly axisymmetric in extent, so these good results are not surprising. However, plots of area-averaged F-factor looking Eastward and Northward for three altitudes (Figure 5) reveal that the microburst is not symmetric in intensity. As indicated by the "TASS windfield" points in Figure 6, the F-factors are larger when looking northward through the microburst than when looking eastward. This is due to the vertical wind distribution (Figure 5, righthand plot) which has multiple regions of high vertical windspeeds. For this reason, the single-microburst fit produces a single broad region of somewhat weak vertical winds in an attempt to globally match the windfield, and the result is that intensity is underestimated in both directions. The two-microburst fit, on the other hand, succeeds in matching the vertical windfield well and duplicates the intensity of the TASS windfield well in both directions.

For alerting purposes, both model windfields adequately represent the actual extent; however, the single-microburst model underestimates the intensity somewhat. The results of previous work, however, indicate that highly accurate intensity estimates are not critical for alert generation.[9] Based on these results, and similar results obtained using TASS windfields from another microburst

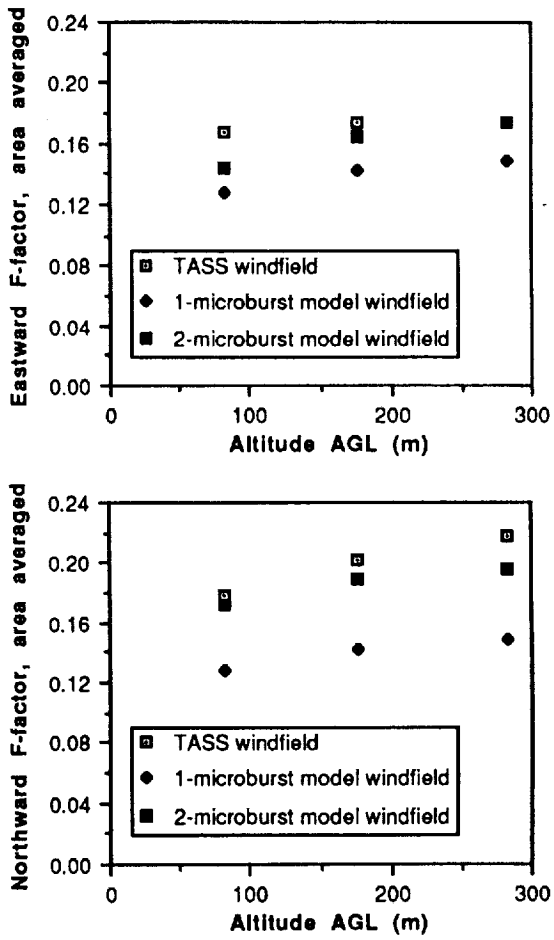


Figure 6: Microburst intensity matching performance for least-squares fit

event, the model was judged to be acceptable for estimation purposes.

Iterated EKF Update Simulation

The next step in algorithm evaluation was to determine if the iterated EKF update procedure was capable of taking sensor data (as modeled by small subsets of the entire windfield) and producing reasonable extent and intensity estimates. The TASS simulated winds were again considered to be the "truth" winds, and simulated sensor data subsets were taken from them. Assuming that the windfield was frozen in time (or alternatively, no time has lapsed between measurement sets), different combinations of sensor data were used sequentially to update the current estimate. Three sensors were considered in this way: (1) TDWR data, (2) winds measured from the aircraft, using inertial and air data measurements (referred to henceforth as INS data), and (3) airborne Doppler radar (ABDR) data.

For TDWR and ABDR data, it was assumed that the sensor was far enough from the microburst that radial

wind measurements could be considered parallel to each other, and that the antenna tilt angle was horizontal so that all data was taken at the same altitude. For example, for an eastward-looking radar, the U-components of the TASS windfield at a single altitude became the working data set. TDWR measurements were taken at an altitude of 82m AGL (the lowest TASS data altitude) for both eastward- and northward-looking cases, and ABDR measurements were taken from 177m and 283m AGL TASS data. In all cases, gaussian zero-mean white noise with a standard deviation of 1 m/s was added to the "truth" data to simulate measurement noise based on TDWR accuracy specifications.[25] All radar data sets were taken at 400m range and azimuth resolution; this is poorer than the resolution of operational radars, but reduced the computation time required to run the simulations.

Aircraft winds (INS data sets) consisted of 3-component winds along a straight flight path at constant altitude. Four 200m resolution INS data sets were defined, including eastward and northward flight paths at 177m and 283m AGL. All paths passed through the approximate center of the windfield, as marked in Figure 5. The measurement noise standard deviation used for aircraft wind measurements was 1.4 m/s.

For simulation purposes, it was assumed that TDWR made the initial microburst detection. Therefore, the first step was to initialize the filter as previously described, and then apply the iterated EKF update to incorporate the TDWR measurement. The resultant parameter estimate and error covariance were saved. Then the estimate was updated by incorporating either an INS data set or an ABDR data set, starting with the saved parameter estimate and covariance matrix. Twelve total sensor fusion cases were tested with both one-microburst and two-microburst versions of the filter.

Single-Microburst Filter

For all cases tested, the iteration procedure used in the update converged in 3 to 5 iterations. Results for four representative cases are presented here:

(1) Initialization only: Eastward-looking TDWR measurements alone (denoted TDWR-E)

(2) The results of (1) were updated using a sequence of eastbound aircraft-measured winds taken at an altitude of 177m AGL (denoted INS-E)

(3) The results of (1) were updated using a sequence of northbound aircraft-measured winds taken at an altitude of 177m AGL (denoted INS-N)

(4) The results of (1) were updated using northward-looking airborne Doppler radar data at 177m AGL (denoted ABDR-N)

The extent results are again good (Table 3), and illustrate the effect of fusing data from sensors with different measurement geometries. The extent figure-of-merit for TDWR-East is 0.85, and does not improve when an eastward path of INS data is incorporated. However, when northbound INS data or northward-looking airborne radar data is incorporated, the extent figure-of-merit increases to the .91 to .92 range. Since the microburst is not exactly symmetric in extent (it is slightly larger in the north-south direction), incorporation of northward-looking data increases the radius parameter in the OBV model to cover more area. This is equal to the performance achieved by the least-squares fit computation.

Table 3: Extent figures-of-merit: 1-microburst time-invariant data fusion.

TDWR-E alone	TDWR-E + INS-E	TDWR-E + INS-N	TDWR-E + ABDR-N
0.853	0.853	0.911	0.917

The effect of multi-directional data is also visible in the intensity results (Figure 7). As with the least-squares results, it is clear that the single-microburst model cannot match intensity with the complex windfield of this microburst. The TDWR-alone result is low, and incorporating an eastbound path of INS data actually lowers the estimate; this is because the path does not cross both regions of high vertical windspeed. Incorporating a northbound path of INS data or the ABDR data improves the estimate significantly at the higher altitudes, from which the INS and ABDR data are taken.

Two-Microburst Filter

The two-microburst version of the filter involved significantly more computation, since in general more iterations were required than for the single-microburst filter. Also, some cases did not converge consistently and required adjustment of the initial parameters. However, when the two-microburst filter did converge, the results were good. Extent figures-of-merit were between 0.85 and 0.90 for all cases. Figure 8 shows eastward intensity values for the algorithm applied to three cases:

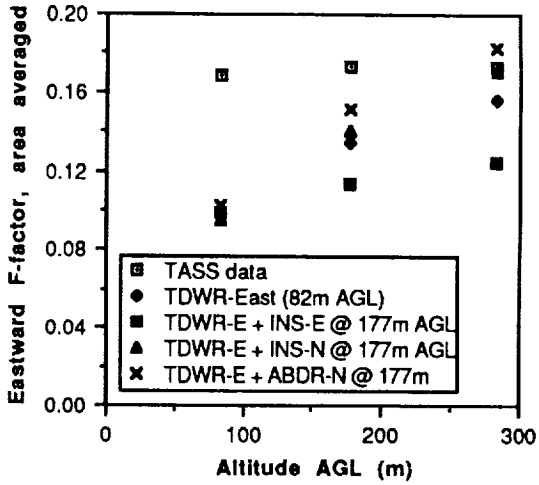


Figure 7: Intensity estimation - 1-microburst time-invariant data fusion

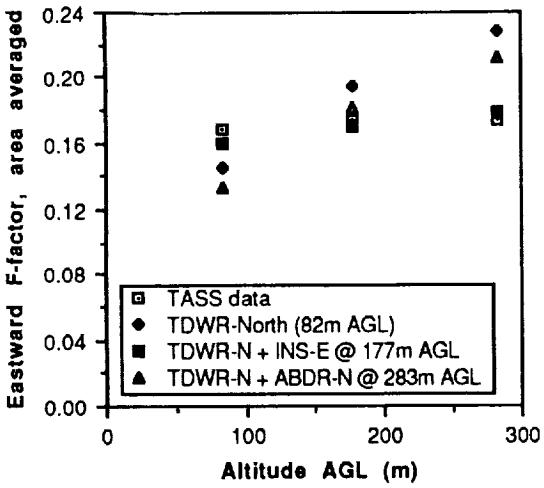


Figure 8: Intensity estimation - 2-microburst time-invariant data fusion

(1) northward-looking TDWR data alone (TDWR-N)

(2) TDWR-N updated with eastward aircraft-measured winds taken at 177m AGL (INS-E)

(3) TDWR-N fused northward-looking airborne radar data at 283m AGL (ABDR-N).

Although the TDWR is looking north and the intensity values shown are for eastbound paths, the results for TDWR data alone are fairly good. There is some overestimation at high altitudes. Inclusion of ABDR data with the same look angle as the TDWR (northward) improves the results slightly. As expected, inclusion of the eastbound INS data provides a second measurement direction and produces the best intensity estimates.

Full Iterated EKF Simulation

The third part of the algorithm evaluation was to include the microburst dynamic model (the propagation part of the filter) and apply the technique to time-varying data. For this analysis, data was taken from three different times in the evolution of the 7/11/88 microburst event. The three data sets were spaced two minutes apart, where the middle data set corresponds to the time-invariant data set used in the previous section and corresponds to the time at which the microburst was strongest.

The time spacing for this data was larger than desired, since TDWR data is updated at 1 min intervals and airborne radar data would be available even more frequently. However, it was still possible to construct illustrative examples. The following three sample cases assume that initial detection is made with northward-looking TDWR. Two minutes later, three different events are postulated:

Case 1. A sequence of eastbound aircraft-measured winds is downlinked to the ground and incorporated along with a second set of TDWR data

Case 2. An aircraft traveling northbound receives the previous TDWR estimate and updates using an airborne Doppler radar

Case 3. An aircraft traveling eastbound receives the previous TDWR estimate and updates using an airborne Doppler radar

At the third time step (+4 minutes), the parameter set is passed to the ground and an update is done using another set of TDWR data. The data sets were derived in the same way as for the time-invariant cases, and the estimate and error covariance were propagated between measurements as described above. In all cases, the single-microburst form of the filter was used.

The extent figures-of-merit (Figure 9) are fairly good (> 0.82) through the first two times, but are slightly lower (0.76) in the third time step. This is due to the distorting effect of an adjacent, weaker microburst on the shape of the primary microburst. The axisymmetric model used in the filter has difficulty representing this

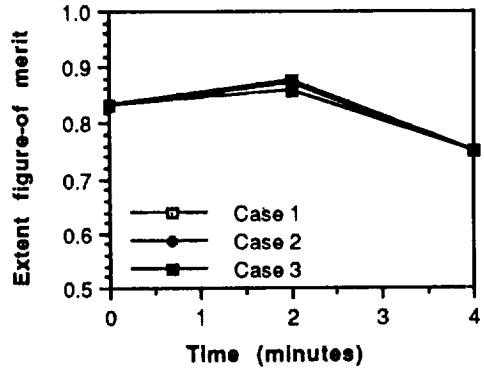


Figure 9: Extent figures-of-merit: time-varying cases

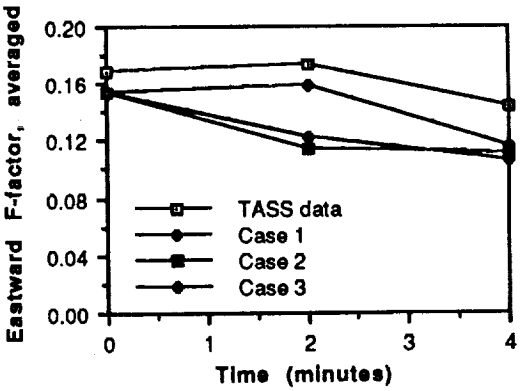


Figure 10: Eastbound intensity estimates time-varying cases

situation. There is little difference between the three sample cases.

The intensity results (Figure 10) are similar to those from the single-microburst time-invariant runs, in that all of the estimates are low. As evident from the "TASS data" curve, the actual microburst increases in strength in the first two minute span and then decreases in the last two minutes. Only in scenario 3, in which northbound TDWR measurements were combined with eastbound ABDR measurements, was the filter able to follow this trend. The low estimates are most likely due to difficulties matching this complex microburst with the single-microburst filter. However, the intensity results were somewhat sensitive to the choice of process noise strength, which indicates a need for further study of microburst time dynamics.

Discussion

The simulations demonstrate the potential usefulness of this technique, particularly for estimating the size and position of the microburst hazard region. Several other characteristics of the algorithm were also observed during the simulation runs, although it should be noted that the use of computational data for a single historical microburst event limits the scope of the conclusions that can be drawn. Further simulation work is planned, using data from actual field measurements.

The single-microburst algorithm appeared to be numerically robust. Errors in initial conditions and reasonable variations in choice of filter parameters did not produce filter instability in either the time-invariant or the time-varying simulations. The two-microburst form, however, was numerically sensitive. In several cases the filter diverged during the update iteration, and choice of parameters such as the initial covariance matrix appeared to have a large impact on the convergence properties of the filter. In cases where convergence was reached the results tended to be dependent on the actual windfield shape. When the windfield had two clear downdraft centers, the convergence was steady and the results for both intensity and extent were good. In cases where there was only one region of high downdraft (such as the first time-step of TASS model data) then the two sets of microburst parameters either became coincident, or one microburst became very weak. This mismodeling problem was also apparent in the diagonal elements of the covariance matrix; the covariance elements corresponding to the unnecessary microburst parameters grew very large. Possible solutions to this problem include more intelligent initialization based on recognized windfield features, or running multiple filters of different types in parallel. In any case, the improved estimation possible from the two-microburst filter must be weighed against the associated numerical difficulties.

Aside from numerical robustness and algorithm tuning issues, there are other implementation issues to be considered. The computational requirements of the filter need to be assessed with respect to available computational resources. Computational load can be decreased by thinning large data sets, at the expense of estimation accuracy. Also, datalink bandwidth needs to be considered. A model parameter list of 9 elements, for example, has an associated 81 element covariance matrix (of which only 45 are unique). It is likely that the entire covariance matrix is not necessary to initialize the next update step, and that some elements could be omitted without loss of performance.

Although the algorithm has been presented in the context of multi-sensor data fusion, it does not require multiple sensors. Benefits would still be gained if it were used with a single sensor due to the additional information contained in the analytical model (correct fluid dynamics, empirical data). Also, the algorithm could be adapted to other fluid dynamic phenomena which can be represented by simple analytical models.

Conclusion

A recursive model-based data fusion algorithm for multi-sensor microburst hazard assessment was presented. A simple analytical microburst model is used to approximate the actual windfield, and a "best" set of model parameters are estimated from measured winds using an extended Kalman filtering technique. The resulting parameter estimate and associated error covariance encapsulate the current state of knowledge about the actual windfield, and can be used to compute estimates of microburst position, extent, and intensity for alert generation. Microburst state dynamics and process noise parameters for the filter were chosen based on statistical data from microburst field studies.

Simulated measurements for three types of sensors were derived from a time-varying computational model of a historical microburst event. Two forms of the algorithm were then tested, one using a single-microburst model and one using a two-microburst model. It was found through both time-invariant and time-varying simulations that both forms of the algorithm were able to estimate the position and extent of the simulated microburst well. The two-microburst model produced better intensity estimates, but suffered from numerical robustness problems. These preliminary results are promising, and further work is planned including simulations using field measurements and study of feasibility issues such as computational requirements.

Acknowledgements

The authors would like to thank Steven Campbell of MIT Lincoln Laboratory and David Hinton and Fred Proctor of the National Aeronautics and Space Administration (NASA), Flight Management Division, Langley Research Center, for their input, advice, and support. This work was supported by NASA and MIT Lincoln Laboratory under Air Force Contract No. F19628-90-C-0002, and the Federal Aviation Administration and NASA under grant NGL-22-009-640.

References

1. National Research Council, *Low Altitude Wind Shear and Its Hazard to Aviation*, National Academy Press, 1983.

2. Evans, James E., "Status of the Terminal Doppler Weather Radar One Year Before Deployment," *4th International Conference on Aviation Weather Systems*, Paris, France, June 24-26, 1991, American Meteorological Society, Boston, MA.

3. Campbell, Steven D., Merritt, Mark W., and DiStefano, John T., "Microburst Recognition Performance of TDWR Operational Testbed," *4th International Conference on Aviation Weather Systems*, Anaheim, CA, January 30 - February 3, 1989, American Meteorological Society, Boston, MA.

4. Wilson, F. Wesley, Jr., and Gramzow, Richard H., "The Redesigned Low Level Wind Shear Alert System," *4th International Conference on Aviation Weather Systems*, Paris, France, June 24-26, 1991, American Meteorological Society, Boston, MA.

5. Bowles, Roland L., "Reducing Windshear Risk Through Airborne Systems Technology," 17th Congress of the International Council of the Aeronautical Sciences, Stockholm, Sweden, September 9-14, 1990.

6. Orlando, V. A., and Drouilhet, P. R., Mode S Beacon System: Functional Description, DOT/FAA/PM-86/19, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA, August 1986.

7. Wanke, C., and Hansman, R. J., "Alert Generation and Cockpit Presentation for an Integrated Microburst Alerting System," AIAA Paper 91-0260, 29th Aerospace Sciences Meeting, Reno, NV, USA, January 7-10, 1991.

8. Wanke, C., Chandra, D., Hansman, R. J., and Bussolari, S. R., "A Comparison of Voice and Datalink for ATC Amendments and Hazardous Wind Shear Alerts," 4th International Symposium on Aviation and Space Safety, Toulouse, France, 20-22 November 1990.

9. Wanke, C., and Hansman, R. J., "Experimental Evaluation of Candidate Graphical Microburst Alert Displays," AIAA Paper 92-0292, 30th Aerospace Sciences Meeting, Reno, NV, USA, January 6-9, 1992.

10. Corman, Larry B., and Mahoney, William P. III, "Integration of the TDWR and LLWAS Wind Shear Detection Systems," *4th International Conference on Aviation Weather Systems*, Paris, France, June 24-26, 1991, American Meteorological Society, Boston, MA.

11. Stratton, D. Alexander, and Stengel, Robert F., "Probabilistic Reasoning for Intelligent Wind Shear Avoidance," *Journal of Guidance, Control, and Dynamics*, Vol. 15, No. 1, Jan.-Feb. 1992.

12. Oseguera, Rosa M., and Bowles, Roland L., "A Simple Analytic 3-Dimensional Downburst Model Based on Boundary Layer Stagnation Flow," NASA Technical Memorandum 100632, July 1988.

13. Vicroy, Dan D., "A Simple, Analytical, Axisymmetric Microburst Model for Downdraft Estimation," NASA Technical Memorandum 104053, February 1991.

14. Hallowell, Robert G., "Aspect Angle Dependence of Outflow Strength in Denver Microbursts: Spatial and Temporal Variations," *16th Conference on Severe Local Storms*, Kananaskis Park, Alberta, Canada, October 22-26, 1990, American Meteorological Society, Boston, MA.

15. Gelb, A., ed., *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.

16. Wilson, J. W., Roberts, R. D., Kessinger, C., and McCarthy, J., "Microburst Wind Structure and Evaluation of Doppler Radar for Airport Wind Shear Detection," *Journal of Climate and Applied Meteorology*, Vol. 29, June 1984.

17. Hjelmfelt, Mark R., "Structure and Life Cycle of Microburst Outflows Observed in Colorado," *Journal of Applied Meteorology*, Vol. 27, August 1988.

18. Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

19. Wishner, R. P., Tabaczynski, J. A., and Athans, M., "A Comparison of Three Non-Linear Filters," *Automatica*, Vol. 5, 1969.

20. Biron, P. J., and Isaminger, M. A., *High Resolution Microburst Outflow Vertical Profile Data from Huntsville, Alabama, and Denver, Colorado*, Project Report ATC-163, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA., April 1991.

21. Proctor, Fred H., *The Terminal Area Simulation System - Volume I: Theoretical Formulation; Volume II: Verification Cases*, NASA CR-4046 and CR-4047, DOT/FAA/PM-86/50, 1987.

22. Schlickenmaier, Herbert W., *Windshear Case Study: Denver, Colorado, July 11, 1988*, Federal Aviation Administration, DOT/FAA/DS-89/19, 1989.

23. Schultz, Thomas A., "Multiple Vortex Ring Model of the DFW Microburst," *Journal of Aircraft*, Vol. 17, No. 2, Feb. 1990.

24. Scales, L. E., *Introduction to Non-Linear Optimization*, Springer-Verlag, New York, 1985.

25. Personal Communication with Steven Campbell, Lincoln Laboratory, June 1992.

Appendix A: Oseguera-Bowles-Vicroy Microburst Model Equations

The U, V, and W wind components are functions of position (x,y,z) and the model parameters:

$$u = \frac{\lambda \tilde{x}}{2}\left[e^{C_1 (z/z_m)} - e^{C_2 (z/z_m)}\right] e^{\left[\frac{2 - \left(\tilde{x}^2 + \tilde{y}^2\right)^{\alpha} / r_p^{2\alpha}}{2\alpha}\right]} \quad \text{(A1)}$$

$$v = \frac{\lambda \tilde{y}}{2}\left[e^{C_1 (z/z_m)} - e^{C_2 (z/z_m)}\right] e^{\left[\frac{2 - \left(\tilde{x}^2 + \tilde{y}^2\right)^{\alpha} / r_p^{2\alpha}}{2\alpha}\right]} \quad \text{(A2)}$$

$$w = -\lambda \left\{\frac{z_m}{C_1}\left[e^{C_1 (z/z_m)} - 1\right] - \frac{z_m}{C_2}\left[e^{C_2 (z/z_m)} - 1\right]\right\} \times \left[1 - \frac{\left(\tilde{x}^2 + \tilde{y}^2\right)^{\alpha}}{2\, r_p^{2\alpha}}\right] e^{\left[\frac{2 - \left(\tilde{x}^2 + \tilde{y}^2\right)^{\alpha} / r_p^{2\alpha}}{2\alpha}\right]} \quad \text{(A3)}$$

where the position offsets are given by

$$\tilde{x} = x - x_0 \quad \text{(A4)}$$

$$\tilde{y} = y - y_0 \quad \text{(A5)}$$

and the radial scale factor λ is:

$$\lambda = \frac{2\, u_m}{r_p \left(e^{C_1} - e^{C_2}\right) e^{(1/2\alpha)}} \quad \text{(A6)}$$

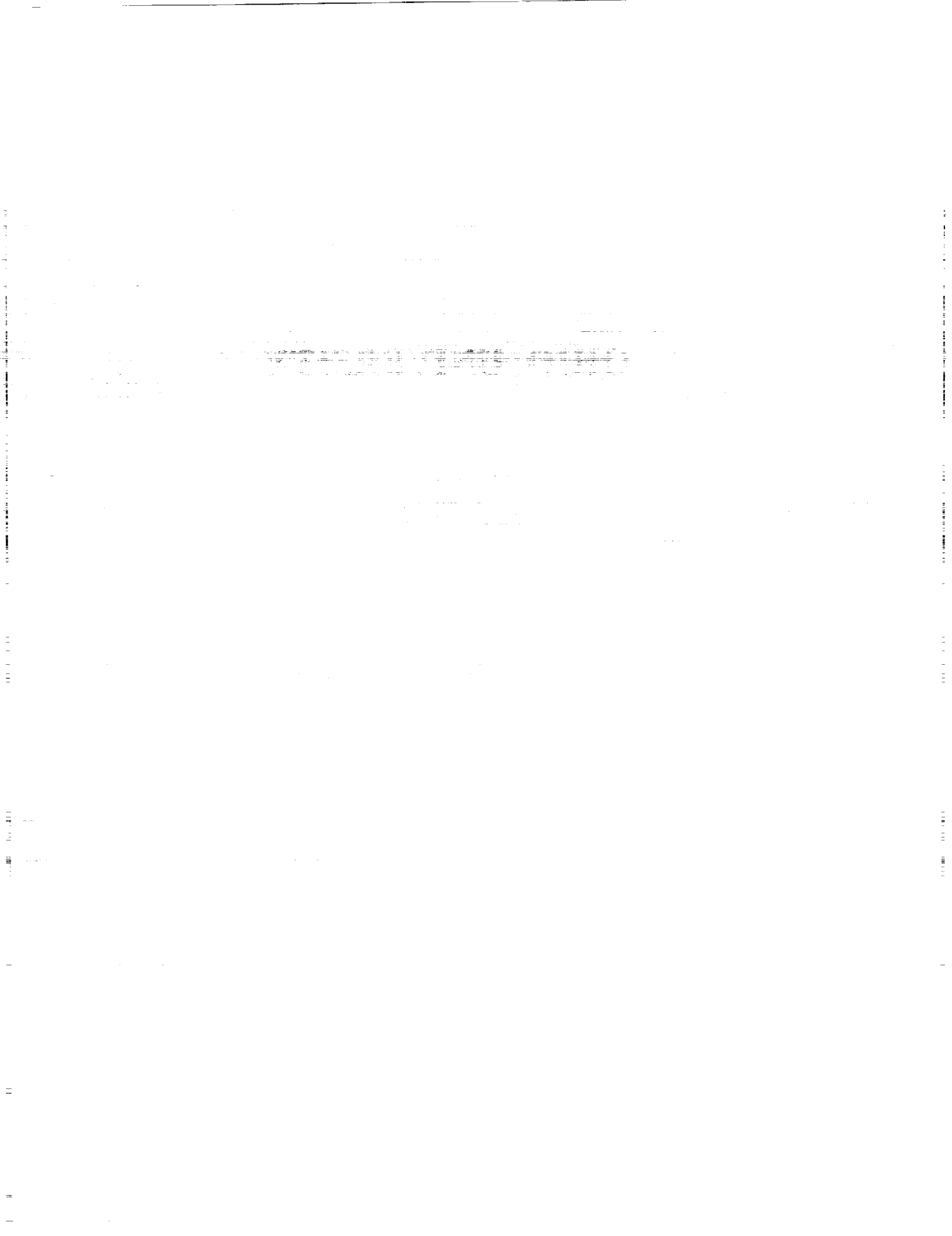
C_1 and C_2 are empirically adjusted constants with the following values:

$$C_1 = -0.15 \quad \text{(A7)}$$

$$C_2 = -3.2175 \quad \text{(A8)}$$

and α is a shaping parameter which was set to 2.0 for the work presented here.

OHIO UNIVERSITY



INVESTIGATION OF AIR TRANSPORTATION TECHNOLOGY AT OHIO UNIVERSITY 1992-1993

Robert W. Lilley
Avionics Engineering Center
Department of Electrical and Computer Engineering
Ohio University
Athens, Ohio

SUMMARY OF RESEARCH

The Joint University Program in Air Transportation Systems provides opportunities for progress by students, staff and faculty at the Avionics Engineering Center, Ohio University. During the 1992-93 year, four conference papers and two M. S. theses were produced; these are summarized in the bibliography below. The conference papers are included in their entirety, for reference.

- Interest in the satellite-based Global Positioning System (GPS) in the interferometric mode implies the need for highly accurate position and velocity estimates in real time, from multiple antennas. Such advanced applications require also an excellent knowledge of the transmitted signal's characteristics (studies of Selective Availability and methods for mitigation).

  Differential mode operations are also implicit when interferometric GPS is applied to aircraft approach operations (studies of ground station siting and performance).

  GPS hybridization with other systems is a key element in eventual sole-means navigational use of the system. Studies of combined GPS/Loran-C and GPS/IRS are supporting this future priority.

  GPS system availability is a pervasive concern, and is a complicated quantity related to required user accuracy, position and time. A comprehensive coverage model is under development.

- Although specific papers were not generated in the weather-uplink research area, this work did support a spin-off effort. Knowledge gained in the weather-uplink work is now being applied in differential GPS uplink studies supported by FAA.

- Fault detection and isolation (FDI) work continues, in direct support of GPS integrity assurance standards being developed for FAA by RTCA. Much of the past FDI work generated in the Joint University Program has been adopted as part of these national/international standards.

ANNOTATED BIBLIOGRAPHY OF 1992-93 PUBLICATIONS

1. Braasch, M. S.; Fink, A. B.; Duffus, K.: Improved Modeling of GPS Selective Availability. Proceedings of the ION National Technical Meeting, San Francisco, CA, January 20-22, 1993.

Selective Availability (SA) represents the dominant error source for stand-alone users of GPS. Even for DGPS, SA mandates the update rate required for a desired level of accuracy in realtime applications. As has been witnessed in the recent literature, the ability to model this error source is crucial to the proper evaluation of GPS-based systems. A variety of SA models have been proposed to date; however, each has its own shortcomings. Most of these models have been corrupted by additional error sources. This paper presents a comprehensive treatment of the problem. The phenomenon of SA is discussed and technique is presented whereby both clock and orbit components of SA are identifiable. Extensive SA data sets collected from Block II satellites are presented. System Identification theory then is used to derive a robust model of SA from the data. This theory also allows for the statistical analysis of SA. The stationarity of SA over time and across different satellites is analyzed and its impact on the modeling problem is discussed.

2. Braasch, S.: Realtime Mitigation of GPS SA Errors Using Loran-C. Wild Goose Association, Annual Convention and Technical Symposium, August 24-27, 1992, Birmingham, England.

3. Braasch, S.: Realtime Mitigation of GPS Selective Availability Using Loran-C. M.S. Thesis, Ohio University, Department of Electrical and Computer Engineering, Athens, OH, June 1993.

The hybrid use of Loran-C with the Global Positioning System (GPS) has been shown capable of providing a sole-means of enroute air radionavigation. By allowing pilots to fly direct to their destinations, use of this system is resulting in significant time savings and therefore fuel savings as well. However, a major error source limiting the accuracy of GPS is the intentional degradation of the GPS signal known as Selective Availability (SA). SA-induced position errors are highly correlated and far exceed all other error sources (horizontal position error: 100 meters, 95 %). Realtime mitigation of SA errors from the position solution is highly desirable. This paper discusses how that can be achieved. The stability of Loran-C signals is exploited to reduce SA errors. The theory behind this technique will be discussed and results using bench and flight data will be given.

4. Skidmore, T. A.: A GPS Coverage Model. Proceeding of the ION National Technical Meeting, Washington, DC, June 29 - July 1, 1992.

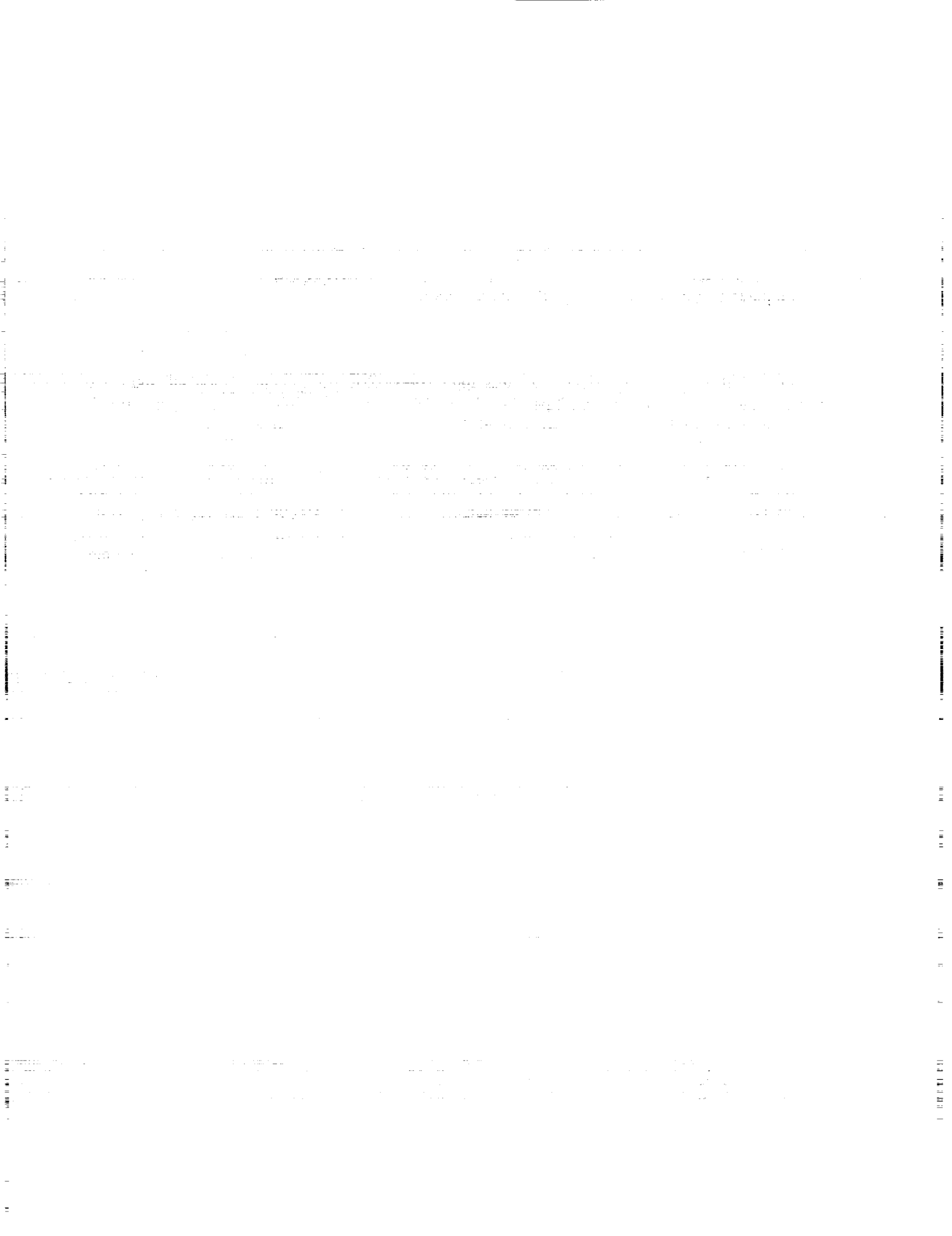
This paper summarizes the results of several case studies using the Global Positioning System coverage model developed by Ohio University. Presented are results pertaining to outage area, outage dynamics, and availability. Input parameters to the model include the satellite orbit data, service area of interest, geometry requirements, and horizon and antenna mask angles. It is shown for precision-landing Category I requirements that the planned GPS

21 Primary Satellite Constellation produces significant outage area and unavailability. It is also shown that a decrease in the user equivalent range error dramatically decreases outage area and improves the service availability.

5. Waid, J. D.: Ground Station Siting Consideration for DGPS. Proceedings of the ION National Technical Meeting, San Francisco, CA, January 20-22, 1993.

6. Waid, J. D.: Development of an Interferometric Differential Global Positioning System Ground Reference Station. M.S. Thesis, Ohio University, Department of Electrical and Computer Engineering, Athens, OH, March 1992.

Aircraft guidance and positioning in the final approach and landing phases of flight requires a high degree of accuracy. The Global Positioning System operating in differential mode (DGPS) is being considered for this application. Prior to implementation, all sources of error must be considered. Multipath has been shown to be the dominant source of error for DGPS and theoretical studies have verified that multipath is particularly severe within the final approach and landing regions. Because of aircraft dynamics, the ground station segment of DGPS is the part of the system where multipath can most effectively be reduced. Ground station siting will be a key element in reducing multipath errors for a DGPS system. This situation can also be improved by using P-code or narrow correlator C/A-code receivers along with a multipath rejecting antenna. This paper presents a study of GPS multipath errors for a stationary DGPS ground station. A discussion of GPS multipath error characteristics will be presented along with some actual multipath data. The data was collected for different ground station siting configurations using P-code, standard C/A-code and narrow correlator C/A-code receiver architectures and two separate antenna constructions.



FAULT DETECTION AND ISOLATION

Greg Bernath
Ohio University
Athens, Ohio

SUMMARY

In order for a current satellite-based navigation system (such as the Global Positioning System, GPS) to meet integrity requirements, there must be a way of detecting erroneous measurements, without help from outside the system. This process is called Fault Detection and Isolation (FDI). Fault detection requires at least one redundant measurement, and can be done with a parity space algorithm. The best way around the fault isolation problem is not necessarily isolating the bad measurement, but finding a new combination of measurements which excludes it.

BACKGROUND

The objective of fault detection and isolation is to use inconsistencies in redundant sensor measurement data to detect and isolate sensor malfunctions. If a given single measurement is in error, it will cause the navigation solution to be in error, possibly greater than the allowable error threshold. Outside sources may not be able to broadcast in a timely manner that a signal is in error; for instance, if a single GPS satellite malfunctions, it could be from 15 minutes to several hours before the information is made public in the satellite broadcast data. Therefore, it is imperative for FDI algorithms to be able to detect and isolate instrument errors using only data from the instruments themselves.

FDI can be implemented in any multisensor navigation system with redundant measurements. Current work is focusing on satellite navigation using GPS, along with hybrid systems such as GPS/Loran-C (Long Range Navigation - C) or GPS/IRS (Inertial Reference System) [3]. FDI used specifically with GPS is also known as RAIM, or Receiver Autonomous Integrity Monitoring [4].

To detect step errors or fast growing ramp errors, a Kalman filter will work well. However, it will not detect a slow growing ramp error, such as might be caused by a GPS satellite clock drift. To detect slow growing errors, the Kalman filter algorithm should be used in parallel with a parity space algorithm.

PARITY SPACE AND ESTIMATION SPACE

Estimation space contains the actual horizontal measurement error and the alarm threshold for a given error. However, actual positions and actual errors are not known, given that all of the measurement data is coming from imperfect sensors. Therefore, the work of detecting and isolating errors is done in parity space. Parity space is a mathematical tool where measurement noise and biases are used to create a parity vector. The parity vector determines the detection statistic, d_k, which is compared to a detection threshold, T_D, in order to determine whether an alarm condition exists.

Errors and biases in parity space and estimation space are related, but it is not a one to one correspondence. The exact correspondence will be determined by measurement geometries. For instance, with a good geometry, a large measurement error (parity space) will result in only a small position error (estimation space). The reverse can also be true. Figure 1 illustrates two different slow growing ramp errors plotted in parity space versus estimation space. In case I, the detection threshold is crossed before the alarm threshold, yielding a false alarm. As the error continues to grow, the alarm threshold is crossed, turning it into a correct fault detection. In case II, the alarm threshold is crossed before the detection threshold, resulting in a missed detection. As the error continues to grow, the detection threshold is crossed, turning it into a correct fault detection. An ideal algorithm would minimize both the number of false alarms and missed detections.

LEAST SQUARES ESTIMATOR ALGORITHM

In a least-squares approach to fault detection, the relationship between the measurements and the user state (position) is given by:

$$\underline{y} = \mathbf{H}\underline{\beta} \qquad (1)$$

where: $\underline{y}$ = measurement vector (n-by-1)
H = data matrix (n-by-m)
$\underline{\beta}$ = user state vector (m-by-1)

$\underline{y}$ is a vector of n measurements, one from each instrument. In the case of using only GPS satellites, it would consist of the pseudoranges. $\underline{\beta}$ is the m-element user state vector, consisting of the user position coordinates and other navigation state elements such as clock offset with respect to GPS time. H is an n-by-m matrix which relates the measurements to the user states.

There are three possible cases:

1) $n < m$: Underdetermined system
2) $n = m$: Exactly determined system

3) $n > m$: Overdetermined system

In the underdetermined case, a navigation solution is not possible. In the exactly determined case, a navigation solution is possible, but fault detection is not.

Algorithms for managing the redundant measurements in an overdetermined system form the basis of fault detection. A parity equation can be derived from equation 1, starting with a mathematical manipulation called the QR factorization on the data matrix H (ref. 2):

$$H = QR \tag{2}$$

H is factored into an n-by-n orthonormal matrix Q ($Q^TQ = I$) and an n-by-m upper triangular matrix R. R contains (n-m) rows of zeros along the bottom, due to the n-m redundant measurements in H. Substituting QR for H in equation (1) gives:

$$\begin{aligned} \underline{y} &= QR\underline{\beta} \\ Q^T\underline{y} &= Q^TQR\underline{\beta} \\ Q^T\underline{y} &= R\underline{\beta} \end{aligned} \tag{3}$$

Now partition R into an m-by-m upper triangular matrix U and (n-m) rows of zeros, denoted by 0. Similarly, partition Q^T into Q_1 (m-by-n) and Q_2 ((n-m)-by-n rows).

$$\begin{pmatrix} Q_1 \\ \hline Q_2 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} U \\ \hline 0 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \tag{4}$$

The least squares navigation state solution is:

$$\underline{\beta} = U^{-1}Q_1\underline{y} \tag{5}$$

U is an upper triangular matrix. Due to the nature of the QR factorization, all matrix elements on the diagonal must be non-zero. Therefore, U is always non-singular and this equation always has a solution.

The parity equation is:

$$Q_2\underline{y} = 0 \tag{6}$$

The measurement vector $\underline{y}$ contains noise ($\underline{e}$) and measurement biases ($\underline{b}$). If $\underline{y}$ is replaced by

($\underline{y}$ - $\underline{e}$ - $\underline{b}$), the 0 in equation (6) can be replaced by the parity vector $\underline{p}$.

$$\underline{p} = Q_2\underline{y} - Q_2\underline{e} - Q_2\underline{b}$$
$$\underline{p} = -Q_2\underline{e} - Q_2\underline{b} \quad (7)$$

Thus, a parity vector will be determined by the noise and bias errors. From the parity vector, it can be determined whether an instrument is in error and an alarm should be raised.

PARITY SPACE AND DETECTION PROBABILITIES

Consider a situation with one redundant measurement. In this case, the parity vector will be reduced to a scalar, and the detection statistic reduces to the absolute value of the scalar. In the case where no measurement bias exists, figure 2 shows the distribution of the parity scalar. Since there is no bias error, the position error is definitely under the alarm threshold and the system is either in the normal operation condition or the false alarm condition. The probability of a false alarm (P_{FA}) is obtained by integrating the areas outside of T_D. For noise having a normal distribution (generally a good assumption), this integral is a standard Gaussian function.

Figure 3 illustrates the case where a large measurement bias exists, making the position error larger than the alarm threshold. In this case the system is either in the correct fault detection condition or in the missed detection condition. The probability of a missed detection (P_{MD}) is the integral of the area inside T_D. Again, if Gaussian noise is assumed, this is a standard Gaussian function.

PROTECTION RADIUS

The above example uses detection threshold, measurement noise, and measurement bias error as parameters to find P_{FA} and P_{MD}. Accuracy requirements are stated in a form like "the probability of exceeding 100 meters accuracy is no greater than 0.05". In order to compare FDI results with such specifications, it helps to rearrange the procedure. This means using the parameters alarm threshold, measurement noise, P_{FA}, and P_{MD} to determine the protection radius, which is the smallest horizontal position error that is guaranteed to be detected with the given probabilities. If all parameters are kept constant, the protection radius will vary only as a function of satellite geometry.

The method resulting in the best protection radius uses all satellites in view. However, many receivers are limited to six channels and are incapable of using more than six measurements. A way around this is to search all possibilities of combinations of 5 or 6 satellites for the set with the best geometry, and use that set to find the protection radius.

Figure 4 shows a comparison of each method for a given location over the span of one day. The parameters used to generate these plots are: σ = 32 meters, $P_{FA} = 6.67 \times 10^{-5}$, and $P_{MD} = 3.3 \times 10^{-9}$.

Since all aircraft carry a baroaltimeter, this can be used as another instrument to improve the algorithm. The altimeter adds another measurement without requiring more channels. The altimeter measurement is weighted according to its accuracy and the phase of flight. Figure 5 shows the effect of altimeter aiding, using the same parameters as before and an altimeter with statistics identical to the GPS satellites.

FAULT ISOLATION

The fault isolation problem is very difficult. Previous work explored fault isolation using both a snapshot method and a time history method. Since the objective is to ensure that the aircraft is flying with a set of good measurements, it is not necessary to isolate the bad measurement. It is only required that the bad measurement is not used in the navigation solution. With this in mind, Fault Detection and Exclusion (FDE) was devised.

In FDE, once an alarm is raised, the algorithm discards the present combination of satellites and looks for the combination with the next-best geometry. If this set also raises the alarm, the algorithm goes on to the next best set. Once a set is found that doesn't raise the alarm, that set is used from then on for navigation. In this manner, the bad satellite is not necessarily isolated, but it is excluded.

CONCLUSIONS

A fault detection algorithm for a multisensor navigation system has been presented. A protection radius has been calculated using several different algorithms, with the best-of-six plus altimeter aiding method being chosen as the best method that will work with all receivers. The fault isolation problem has been bypassed by using fault exclusion. The only remaining work for the algorithm is to program it into a receiver and flight test it.

ACKNOWLEDGEMENTS

This work was supported by the Federal Aviation Administration and the National Aeronautics and Space Administration through the Joint University Program for Air Transportation Research (Grant NGR 36-009-17).

REFERENCES

1. Kline, P. A. and van Graas, F.: Fault Detection and Isolation for Multisensor Navigation Systems. NASA Conference Publication 3131, Joint University Program for Air Transportation Research 1990-1991.

2. Golub, G. H., and Van Loan, C. F.: Matrix Computations, Second Edition. The Johns Hopkins University Press, Baltimore, MD, 1989.

3. Brenner, M.: Implementation of a RAIM Monitor in a GPS Receiver and an Integrated GPS/IRS. Proceedings of the Third International Satellite Division Meeting of the Institute of Navigation, September 1990.

4. Van Graas, F. and Farrell, J. L.: Receiver Autonomous Integrity Monitoring (RAIM): Techniques, Performance and Potential. Proceedings of the 47th Annual Meeting of the Institute of Navigation, Williamsburg, VA, June 10-12, 1991.

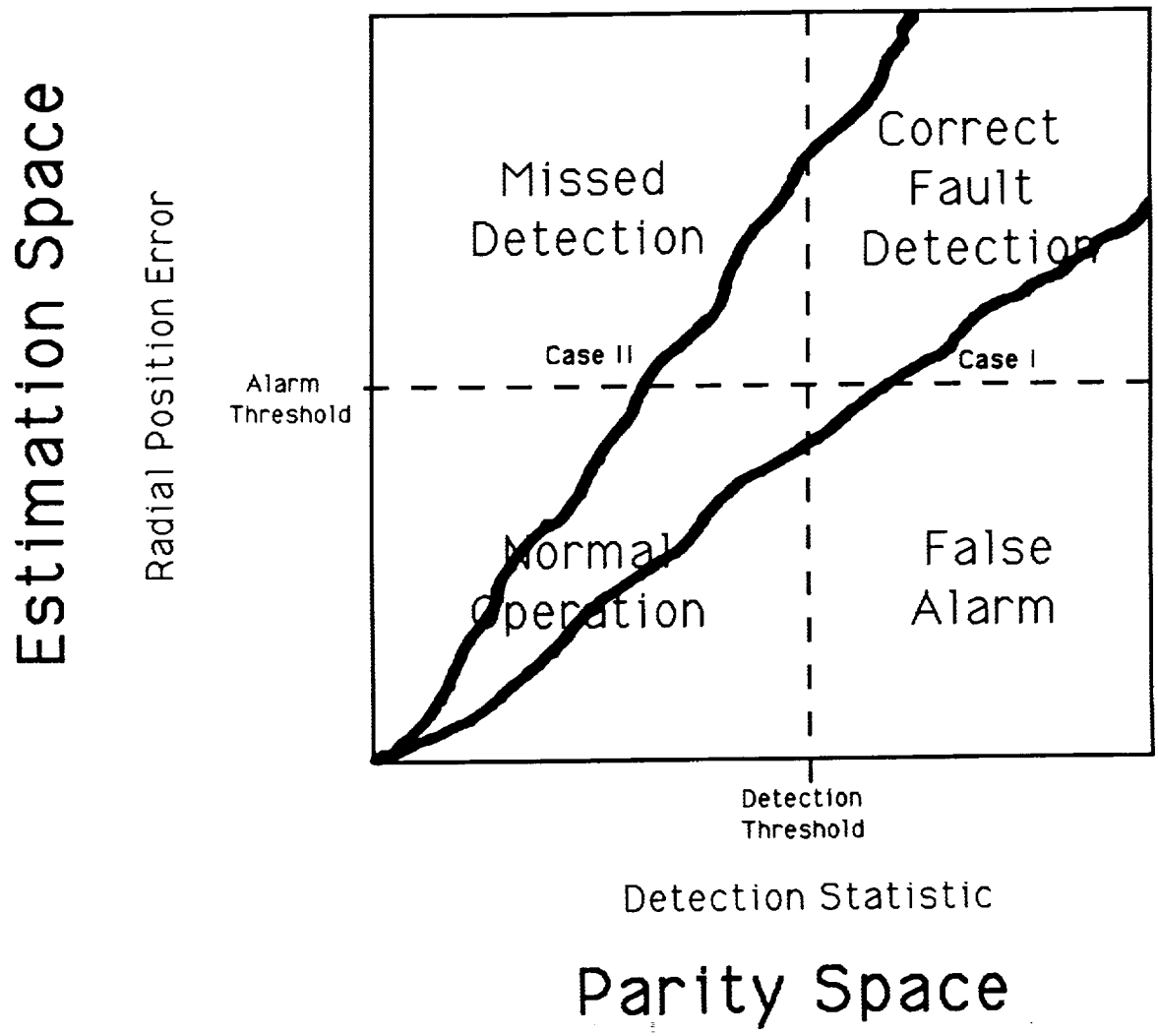


Figure 1. Two slowly growing measurement errors plotted in parity space versus estimation space.

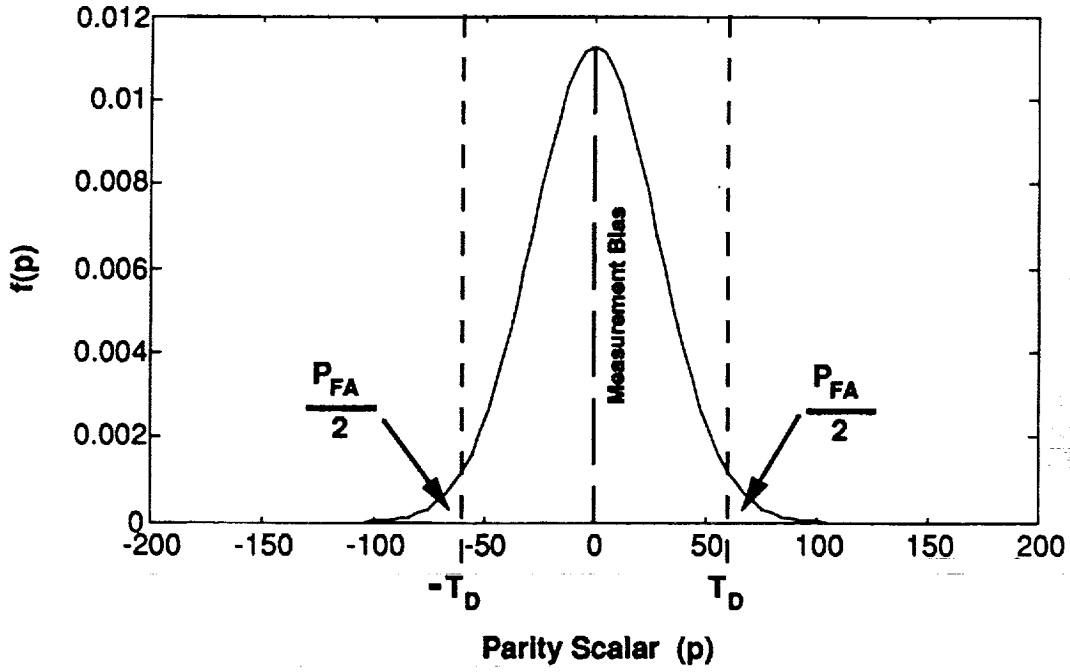


Figure 2. Probability density function for the parity scalar in the absence of a measurement bias error.

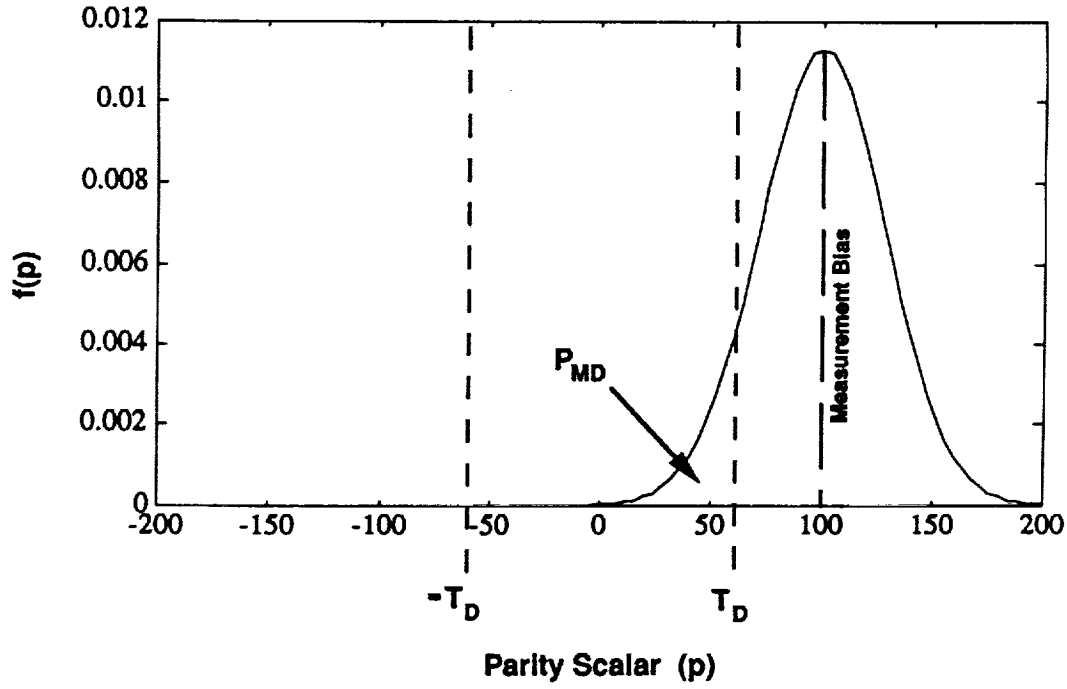


Figure 3. Probability density function for the parity scalar in the presence of a measurement bias error.

Figure 4. Worst Case Protection Radius (36N 140E)

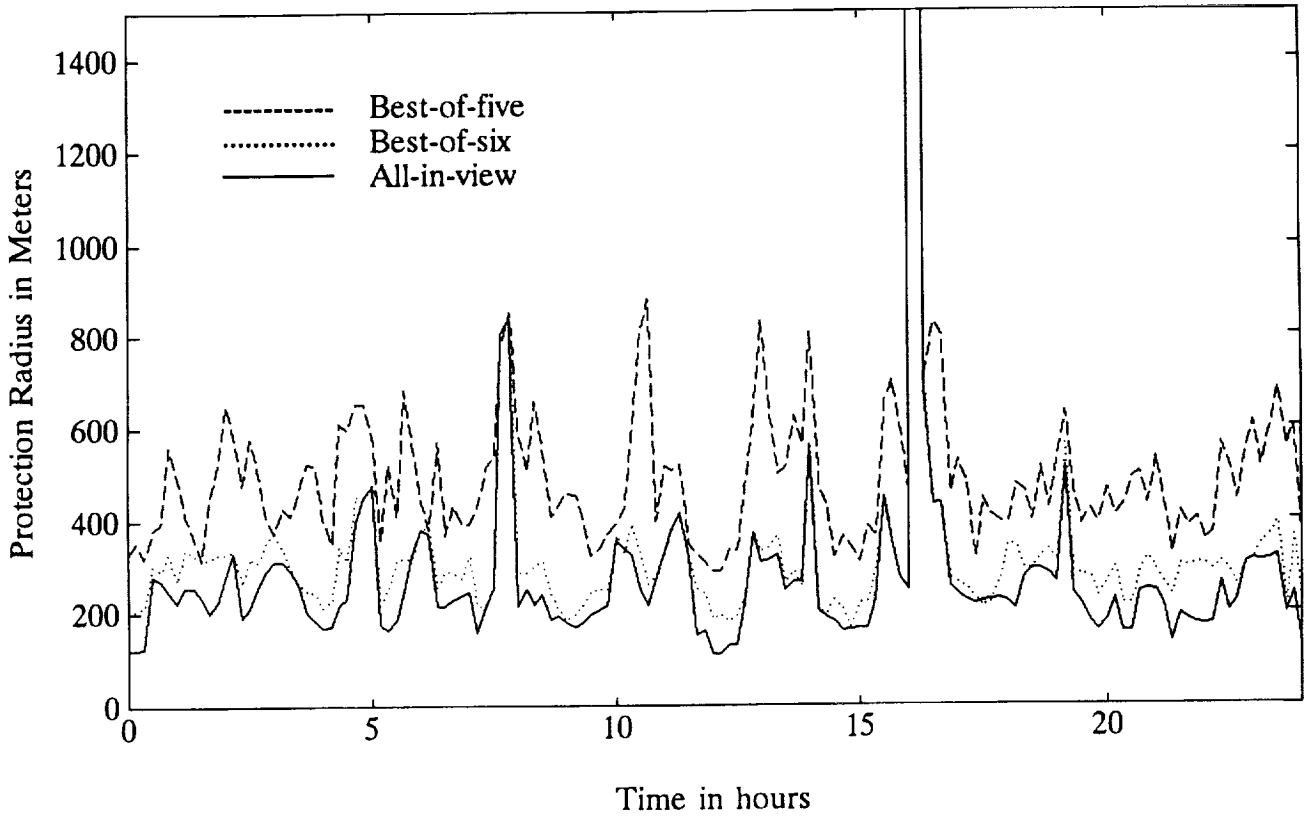
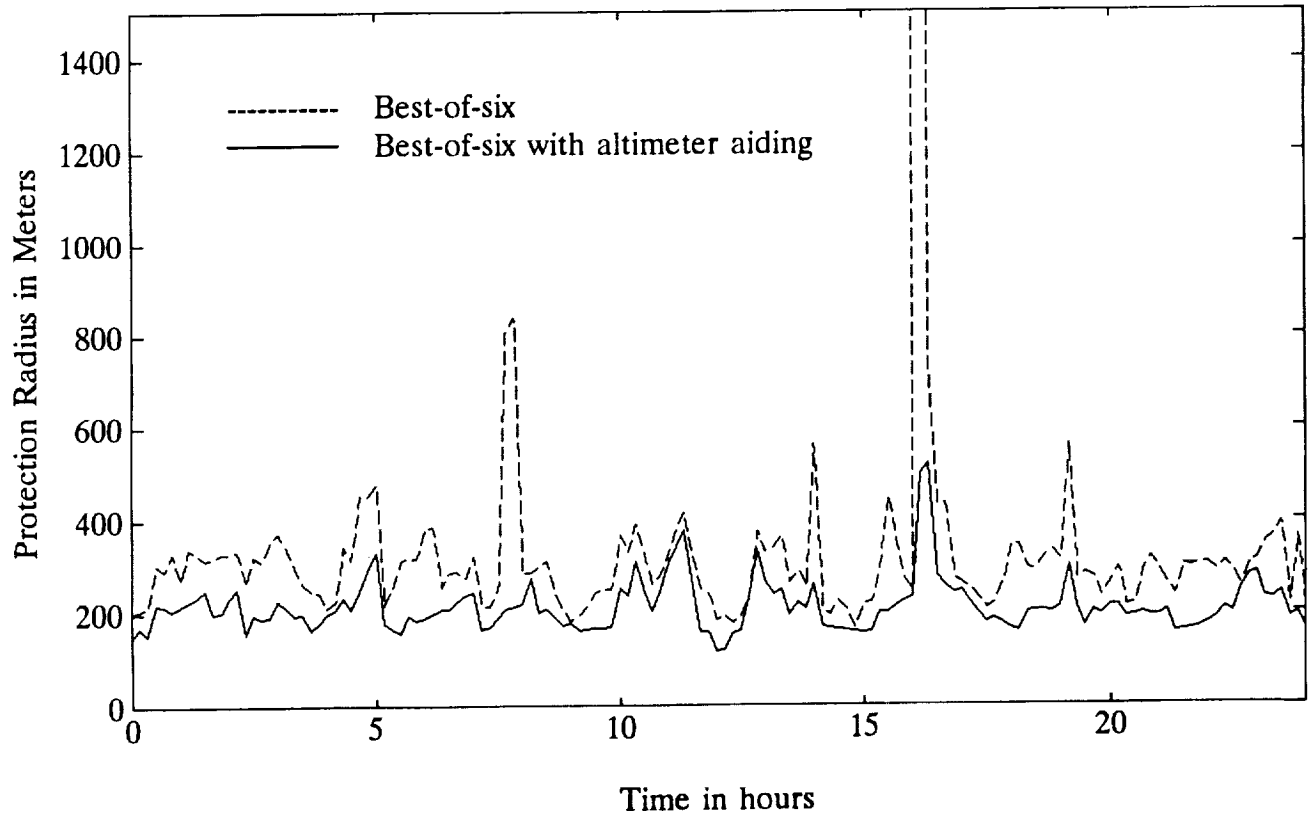


Figure 5. Worst Case Protection Radius (36N 140E)



IMPROVED MODELING OF GPS SELECTIVE AVAILABILITY

Michael S. Braasch
Avionics Engineering Center, Ohio University
& Telecommunications and Traffic-Control Systems Group, Delft University of Technology

AnnMarie Fink
Keith Duffus
Avionics Engineering Center, Ohio University

ABSTRACT

Selective Availability (SA) represents the dominant error source for stand-alone users of GPS. Even for DGPS, SA mandates the update rate required for a desired level of accuracy in realtime applications. As has been witnessed in the recent literature, the ability to model this error source is crucial to the proper evaluation of GPS-based systems. A variety of SA models have been proposed to date; however, each has its own shortcomings. Most of these models have been based on limited data sets or data which have been corrupted by additional error sources. This paper presents a comprehensive treatment of the problem. The phenomenon of SA is discussed and a technique is presented whereby both clock and orbit components of SA are identifiable. Extensive SA data sets collected from Block II satellites are presented. System Identification theory then is used to derive a robust model of SA from the data. This theory also allows for the statistical analysis of SA. The stationarity of SA over time and across different satellites is analyzed and its impact on the modeling problem is discussed.

INTRODUCTION

The intentional degradation of the GPS signal known as Selective Availability (SA) is the single largest error source for open loop (non-differential) users of GPS. This degradation is accomplished through manipulation of the broadcast ephemeris data and through dithering of the satellite clock (carrier frequency). Manipulation of the satellite ephemeris data results in erroneous computation of satellite position. This is a long term, non-periodic error trend over the duration of the satellite pass. Dithering of the satellite clock results in erroneous code-phase and carrier-phase measurements. This error trend consists of random oscillations with periods on the order of 5 to 10 minutes.

As the recent literature has shown, a software-centered GPS signal model is essential for the bench testing and evaluation of a variety of GPS-based systems [Bar-Sever, et al, 1990; Braasch, 1990-91; Feit, 1992; Lear, et al, 1992]. A key element in this model is the module for SA. Several SA models have been presented over the past few years; however, each has been derived based on limited data sets or data which have been corrupted by other error sources. An accurate SA-only model is needed. Ideally, this model should be able to generate the typical kinds of SA error traces observed on any satellite at any time. Furthermore, since the two error sources behave quite differently, independent characterization of the orbit and clock components of SA is required. This paper presents work performed to address these issues.

SA DISCUSSION

SA was formally implemented by the Department of Defense on March 25, 1990 [Anon., 1990]. At that time, however, SA had been on experimentally for nearly one year. Various groups reported observing SA-like errors soon after the launch of the first Block II satellite, SVN 14, in February of 1989 [Braasch, 1990-91; Kremer, et al, 1990].

These observations led to the development of the first model of SA based on actual data [Braasch, 1990-91]. In subsequent years, other researchers developed additional SA models [Chou, 1990; Lear, et al, 1992]. None of the investigations, however, were able to answer some fundamental questions: 1) Is SA the same on all satellites? 2) For a given satellite, is SA a stationary random process? That is, do the statistical properties of the SA vary as a function of time? 3) Quantitatively speaking, what is orbital SA?

Presented at the Institute of Navigation 1993 National Technical Meeting, San Francisco, CA, January 20-22, 1993

ORBIT ERROR ANALYSIS

Accurate modeling of SA requires consideration of both the orbital and clock error components. Previous SA investigations have focussed on the clock component only without consideration of the orbital component.

The ability to observe the orbital error component relies on the data provided by various public and private GPS tracking networks. These networks employ a variety of GPS tracking stations which make range measurements to the satellites. Since the locations of the tracking stations are known, this information can be coupled with the range measurements to calculate the position of the satellites.

The result is the so called precise ephemeris or orbit data. Since the precise orbits are calculated according to where the satellites currently are located, they are more accurate than the broadcast ephemeris data (even without SA) which represents a prediction of where the satellites will be in the future. This precise orbit data is used in a variety of non-realtime GPS applications which require the utmost of accuracy.

The precise orbit data are made available to the public in a variety of formats and media. The data used in this study were obtained from the National Geodetic Survey (NGS) through the Navstar GPS Information Center Bulletin Board and from the Scripps Institution of Oceanography (University of California at San Diego) through their own bulletin board service. The various computer programs required to read the data formats and perform the required interpolations were provided by the NGS [Remondi, 1985; Remondi, 1989; Remondi, 1991]. For verification, precise data were obtained both from NGS and Scripps and compared.

During April of 1992 (days 104, 112,113), broadcast ephemeris data were collected from 4 Block I satellites and 11 Block II satellites. Some months later, after the precise ephemeris data had been posted, the precise orbits were compared with the orbits calculated using the broadcast ephemeris. Along-track, cross-track and radial errors were calculated and plotted. Since orbit predictions are never perfect, errors on the order of a few meters were expected even in the absence of SA [Ananda, et al, 1984; Bowen, et al, 1985]. Surprisingly, the error plots for all satellites (Block I and Block II) were on the order of a few meters. Figures 1 through 3 show an example of orbital errors computed for satellite 19.

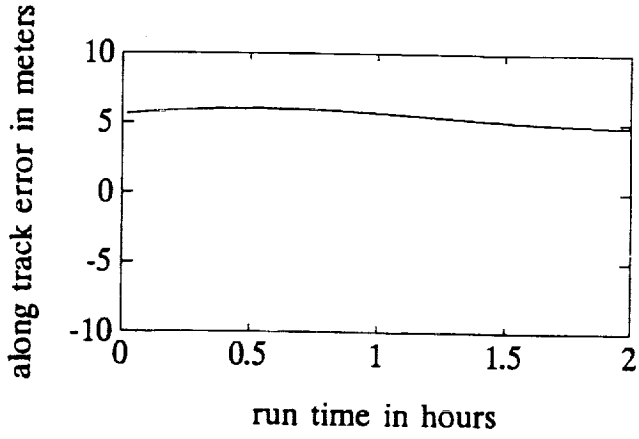


Figure 1. SV 19 ATK Error

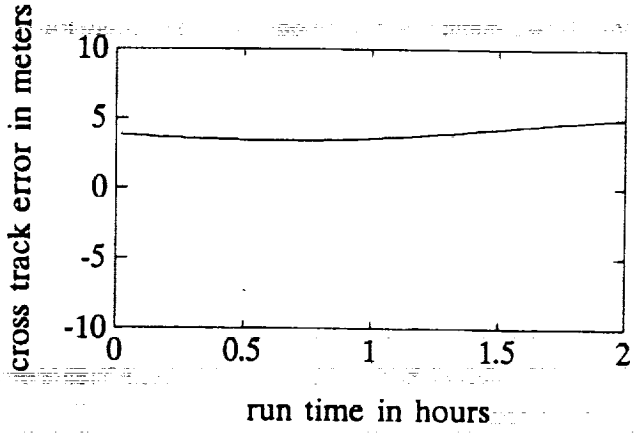


Figure 2. SV 19 XTK Error

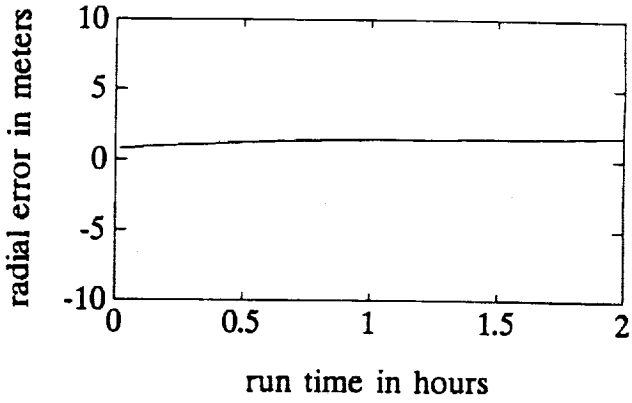


Figure 3. SV 19 Radial Error

Based on these limited data sets, it would seem that the orbital component of SA has not been implemented. It is possible that SA was turned off at this time. However, at the very least, a method now exists whereby the orbital component of SA can be observed. Further data collection efforts are planned to determine if this lack of orbital SA is a regular phenomenon or not.

SA (CLOCK COMPONENT) DATA COLLECTION AND REDUCTION

Having performed the orbital error analysis, the next phase in the study was to collect data for analysis of the clock component of SA. As was noted by Lear, et al (1992), the clock component of SA is a smooth error trace over time and therefore carrier-phase (integrated doppler) data must be collected for the data reduction. This was one of the greatest drawbacks of the models presented in Braasch (1990-91). Since only pseudorange data were available for that study, the data reduction process left a combination of SA and receiver noise. Since filtering could not be performed without imposing assumptions on the underlying SA waveform, it was decided that a model would be derived for the combination of SA and receiver noise [Braasch, 1990-91]. An additional problem with that study was the fact that the data were collected (and hence the model operated) at a data rate of 1/6 Hz. The need for an SA-only model operating at the standard 1 Hz rate served as the original motivation for this study.

During the first week of December (November 30 - December 4), 1992, integrated doppler data were collected at a known location from 10 Block II satellites. The data were collected at Ohio University using a Stanford Telecommunications, Inc. modified Time Transfer System model TTS-502B under the control of a personal computer. The term "modified" refers to the fast-sequencing version of the receiver produced by Stanford Telecommunications, Inc. For the purposes of this study, the important aspect of the modified receiver is its ability to make continuous carrier-phase (integrated doppler) measurements with fine resolution and low noise. The data rate was 1 Hz.

In order to extract the SA waveform, the following steps were taken. First, the true ranges from the satellite to the known antenna location were calculated for the duration of the satellite pass. These were subtracted from the integrated doppler measurements. What remains are referred to as measurement residuals and are a combination of SA, receiver clock drift, atmospheric delay, multipath and a bias due to the ambiguity in the integrated doppler measurements. For environments in which the strength of the multipath is less than the direct signal, the carrier-phase multipath error is guaranteed to be less than 5 cm [Braasch, 1992]. Although it will not be proven here, suffice it to say that the antenna environment used in this study satisfies this criterion. Since a rubidium standard was used as the time base for the receiver, the receiver clock drift is extremely stable and is typically modeled as a first order polynomial [Kremer, et al, 1990]. However, since dual-frequency measurements were not available, ionospheric delay could not be removed. In addition, tropospheric delay is also present. It should be recognized though, that the delays due to the atmosphere are typically long term trends. The result then, is the combination of bias, clock drift and atmospheric delay can be removed by fitting a second-order polynomial to the measurement residuals and subtracting it out. If any bias or long term drift component is present in SA, it will be removed also [Braasch, 1990-91; Lear, et al, 1992]. If an extremely long term error component does exist in the clock SA, it can only be observed if the user clock is synchronized to GPS time [Braasch, 1990-91]. It should also be noted that since the precise ephemerides for the satellites were not available at the time of this writing, broadcast ephemeris was used in the computation of the true ranges. However, under the assumption that the broadcast ephemeris is as accurate as in our previous analysis, this error component is virtually negligible. Even if an orbital SA component is present, it will tend to be removed through the subtraction of the best-fitting second-order polynomial.

The results of the data collection and reduction are shown in figures 4 through 13. The SA error amplitude varies from 40 to 70 meters and the oscillations have periods on the order of 5 to 10 minutes. The variations in the data record length are due to several factors including satellite availability, truncation of records due to receiver glitches and more importantly, truncation of records in order to achieve stationarity. More detail on this last point will be given in a later section.

SA MODEL IDENTIFICATION

Over the past few years, various models have been used to simulate SA. The first SA model was not based on actual SA data but was deduced from a sample probability distribution curve [Matchett, 1985]. The GPS Joint Program Office (JPO) generated SA samples and then computed the curve from these samples. A second-order Gauss-Markov process was postulated and the coefficients were adjusted until its distribution curve matched the one provided by the

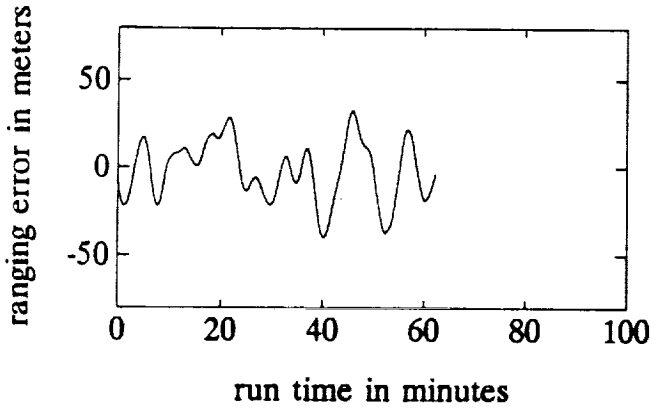


Figure 4. SA on SV 14

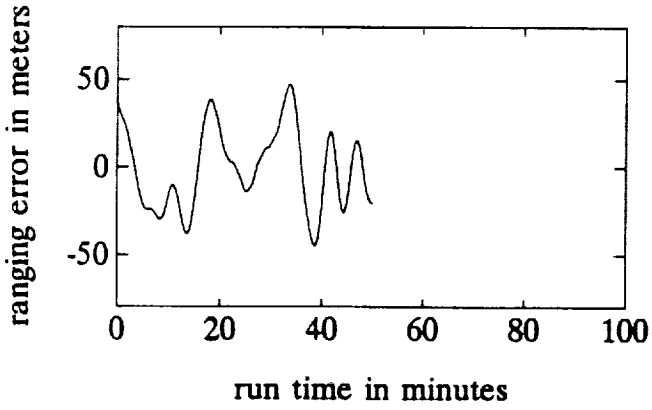


Figure 5. SA on SV 15

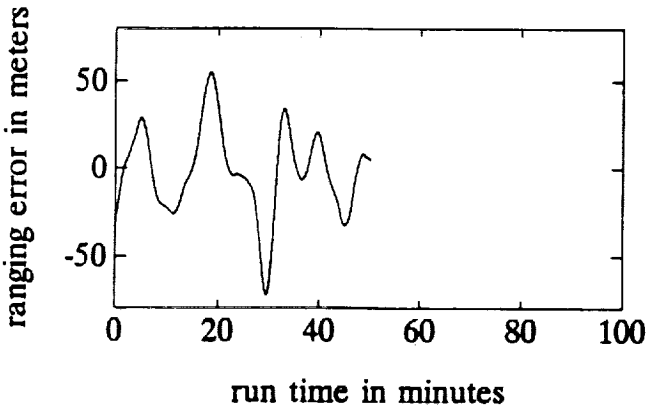


Figure 6. SA on SV 16

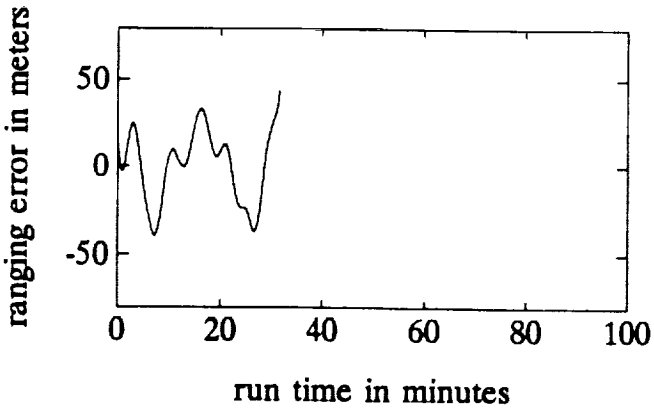


Figure 7. SA on SV 17

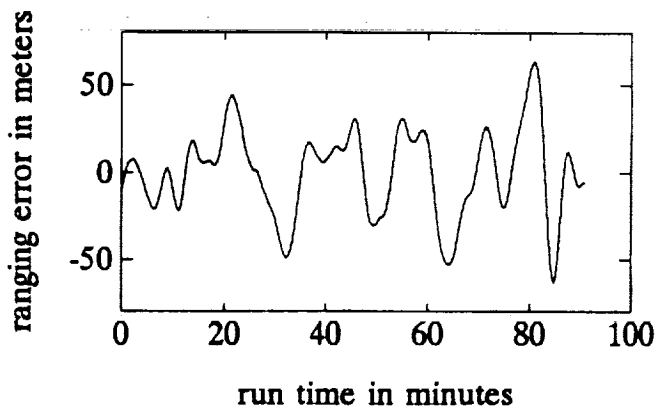


Figure 8. SA on SV 19

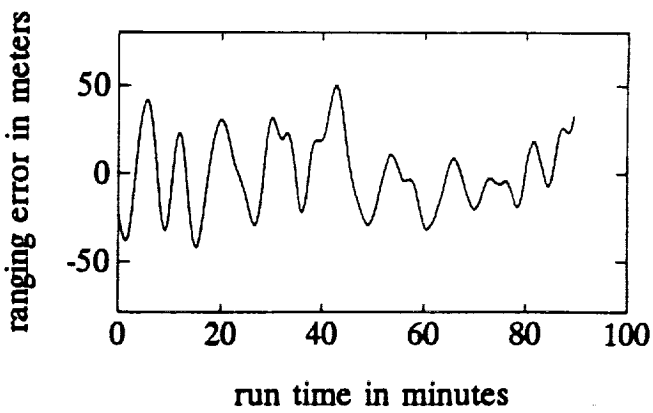


Figure 9. SA on SV 21

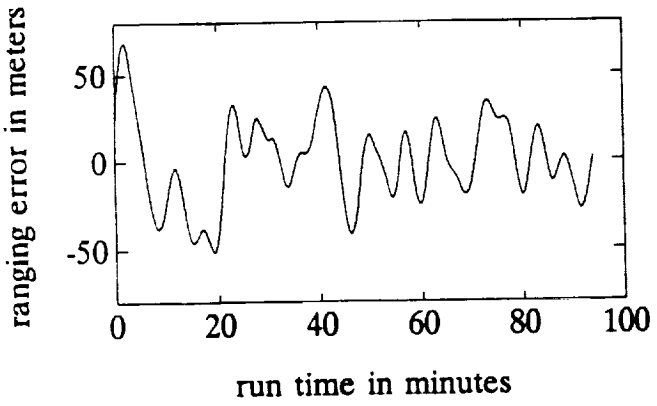


Figure 10. SA on SV 23

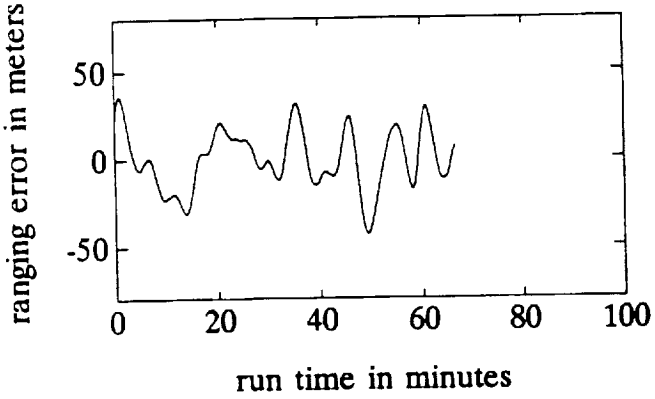


Figure 11. SA on SV 24

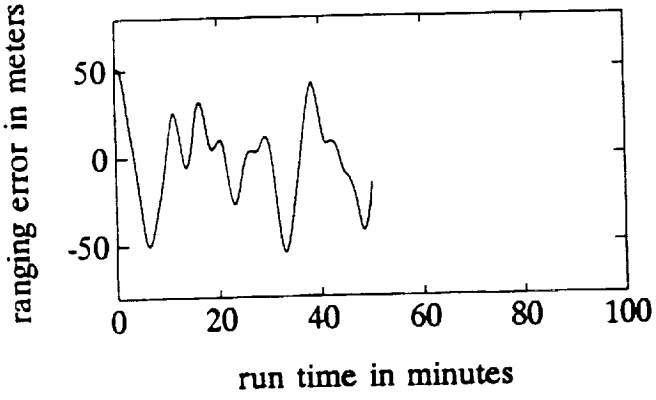


Figure 12. SA on SV 25

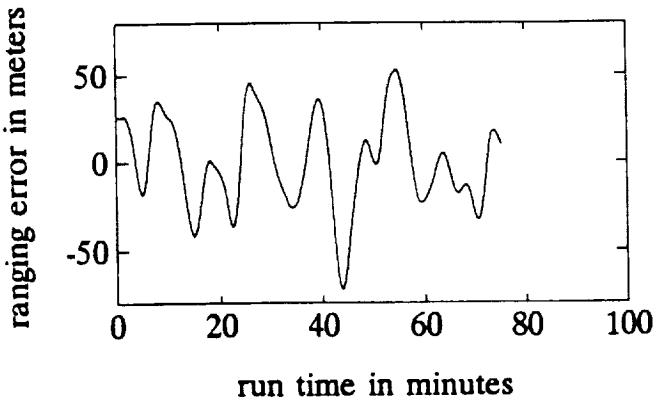


Figure 13. SA on SV 28

JPO. The first models obtained from actual SA data were time series models derived using System Identification theory [Braasch, 1990-91]. Later, Chou also implemented a second order Gauss-Markov process but his was based upon actual SA data [Chou, 1990]. In their recent paper, Lear, et al (1992) present several time series and analytical models also based upon actual SA data.

For this study, System Identification theory was employed to derive time series models in a manner similar to that used in Braasch (1990-91). In general, time series models are based upon the assumption that the data of interest (SA in this case) can be modeled as the output of a linear system (pole-zero filter) driven by Gaussian white noise. Conceptually, derivation of time series SA models can be thought of as a two-step process. The first step is to send the SA data through a filter and adjust the poles and zeros (or equivalently, filter coefficients) such that the output is Gaussian white noise with minimum variance (the output is referred to as residuals). The second step is then to compute the inverse of the filter determined in the first step. Model identification is now complete. Statistically equivalent SA data can then be generated by driving the inverse filter with Gaussian white noise (whose variance is equivalent to that of the residuals in the first step). Kelly (1992) provides an excellent overview of time series model identification and its application to the problem of microwave landing system (MLS) signal modeling.

Three decisions are inherent in the above procedure. The first is the choice of model (filter) type. Three are possible: 1) a pole-zero filter (giving rise to what is known as an Autoregressive Moving Average or ARMA model); 2) an all-pole filter (yielding an Autoregressive or AR model); 3) an all-zero filter (yielding a Moving Average or MA model). The

second decision is the choice of model order. That is, if an AR model is chosen, how many poles will be used? The third decision is related to the first two and involves determining if a given residual sequence is white.

Since the primary goal in this study was to derive an accurate SA-only model, an AR model type was chosen. This stems from the fact that ARMA and MA models tend to be noisy. In fact, Braasch (1990-91) concluded that an ARMA model was the best model type for the combination of SA and receiver noise. An autoregressive model of order p (referred to as an AR(p)) is defined as follows [Marple, 1987]:

$$y(n) = -\sum_{k=1}^{p} a(k)y(n-k) + e(n) \qquad (1)$$

where y is the model output, n is the time index, a(k) is the kth filter coefficient, and e is the input Gaussian white noise. Note that the SA models derived from the data will operate at 1 Hz since they are tied to the data collection rate.

Having made the decision to use an AR model type, the rest of the process involved finding the optimum model order and coefficients (pole locations). For a given model order, many methods exist for optimizing the coefficients [Kay, 1987; Ljung, 1987; Marple, 1987]. The one chosen in this study was the Modified Covariance or Forward-Backward method. The second name stems from the fact that the optimization criterion is the minimization of forward and backward prediction errors. As will be shown later, this method performs quite well with SA data.

Several methods exist for model order selection. The majority of these methods have been developed for extremely short data records. The main issue is that one wants to derive a model for the underlying statistical process which gave rise to the data. When model orders are selected which are too high (i.e. approaching the number of data points in the sample), the result is a "fit" of the sample data record rather than the underlying statistical process. The model order selection method used in this study is known as the Principle of Parsimony. The simplest acceptable model is the one chosen. An acceptable model is the inverse of the filter which outputs white noise when driven with SA. Note that if the model order is too low, the residuals will not be white even though the coefficients have been optimized.

The model identification, therefore, proceeds as follows. For a given sample of SA data, the coefficient is optimized for a first-order filter and the residuals are examined. If they are not white, then the coefficients for a second-order filter are optimized and the residuals are examined again. The process is repeated until the model order and optimum coefficients are found for which the residuals are white. This process was performed for each of the SA data sets shown earlier. Depending upon the data set, models of either 9 or 11 coefficients were derived.

The method for determining whiteness involved examination of the autocorrelation function. An example is given in figure 14 where the autocorrelation function is plotted for the residuals from the SA data of satellite 28. Ideally, the autocorrelation function of white noise has a spike at lag 0 and is zero everywhere else. However, that can be obtained only for infinite length sequences. As a result, some minor "sidelobes" will occur at lags other than zero for white noise sequences which are finite. The dotted lines in the figure represent the 99% confidence intervals for the sidelobes. As can be seen in the plot, the sidelobes lie inside the confidence intervals for the most part and thus the model is acceptable.

Further validation of the model can be performed by generating some waveforms and comparing the power spectral densities (PSD's) of the generated and collected data. An example is shown in figures 15 and 16. Figure 15 shows the waveform generated by the SA model which was derived from the SV 28 data. Note that if one compares the waveform to that of the collected data (figure 13), they are not the same. However, they are statistically equivalent. That is, the periods and amplitudes of the generated data are the same as for the collected data. This is better illustrated in figure 16 where the PSD's of the two waveforms are plotted. Although it is difficult to see, there are actually two PSD's plotted. The solid line represents the collected data and the dashed line represents the generated waveform. PSD comparisons were performed on all of the models derived from the data. In each case the result was similar to that shown here.

A final step in model validation concerns the power in the residuals. Recall that in step one of the model derivation process, the goal was to find a filter which output white noise (residuals) with minimum variance when driven with SA. The need for minimum variance is important from both a theoretical and practical viewpoint. Theoretically, having residuals with minimum variance means that the filter has been optimized and embodies the structure (i.e. correlation or information) of the SA. Kelly (1992) refers to this as the filter "explaining" the data. However, from a practical viewpoint, minimum variance is also required. This is particularly true when trying to model random, yet smooth, waveforms such as SA.

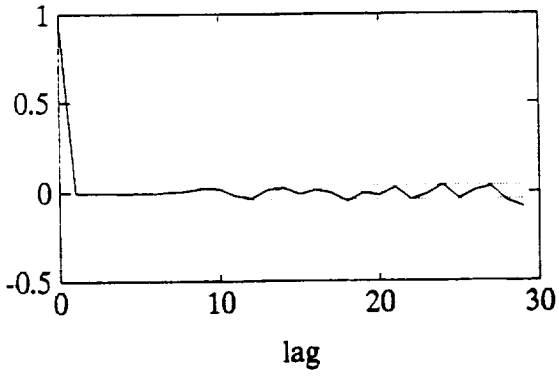


Figure 14. Autocorrelation of SV 28 residuals

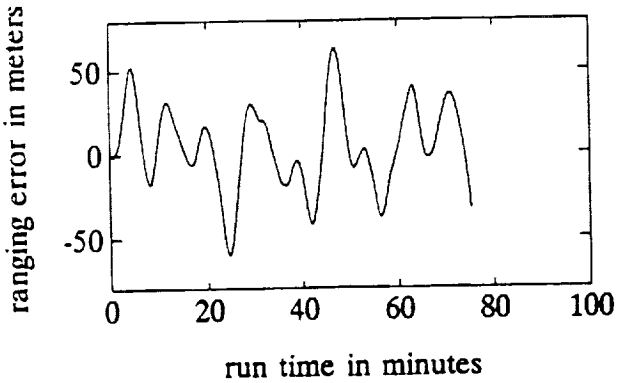


Figure 15. SA model output

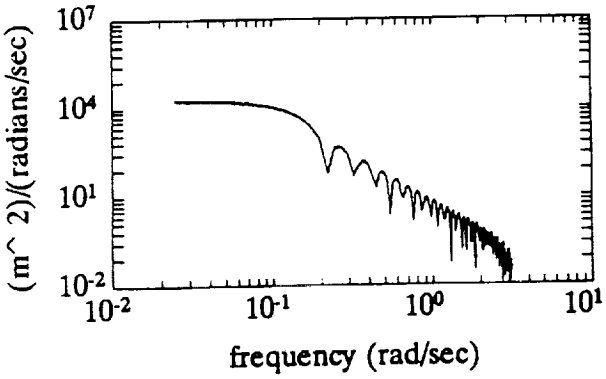


Figure 16. Modeled and measured SA PSD's

Figures 17 and 18 illustrate the success of the AR model type in this respect. The residuals plotted in figure 17 have a standard deviation of 4.12 mm (4.12×10^{-3} m). Since this represents the amplitude of the noise driving the model (see equation 1), it follows that any noise-like behavior in the generated SA waveforms will be negligible. This is verified in figure 18 which shows the smooth waveform of the generated SA over a short time interval.

MODEL IDENTIFICATION RESULTS

Having derived ten models for SA, the question which poses itself is: Which one do I use? Ideally, one would like to use a single model to generate the SA from all satellites. Multiple SA waveforms corresponding to different satellites could then be generated simply by driving the model with multiple Gaussian white noise sequences. It is therefore necessary to compare the models and the collected data to determine if any equivalence exists. If the collected data share similar PSD's and their corresponding models are similar, then a single SA model is feasible.

As mentioned in the previous section, models with either 9 or 11 coefficients were derived from the collected SA data. For the purposes of comparison, 11th order models were derived for those data sets initially giving rise to 9th order models. Although, strictly speaking, this violates the Principle of Parsimony, the additional complexity of having two more coefficients is negligible.

Although they will not be listed here in their entirety, a comparison of the coefficients for the ten models would seem to indicate little similarity. However, examination of their corresponding pole plots provides more insight. An example is given in figure 19 where the poles of two models are plotted. The models were derived from the data sets of SV 28 and SV 25. The ellipses around the poles indicate the two-sigma confidence regions. Notice that for all of the poles the confidence regions of the two models either overlap or are in close proximity to each other. Admittedly, this is not a strict statistical proof of model equivalence (for that, a multivariate analysis of variance hypothesis test is required; see Kelly (1992)). However, it is at least an indication of model similarity.

Pole-plot comparisons were performed with all of the models. Five were found to be similar. These five were the models derived from SV's 28, 25, 19, 16 and 15. The similarity was verified through comparison of the PSD's of collected and generated SA waveforms. Since the five models are similar, any one of them can be chosen and used as the SA model. The coefficients

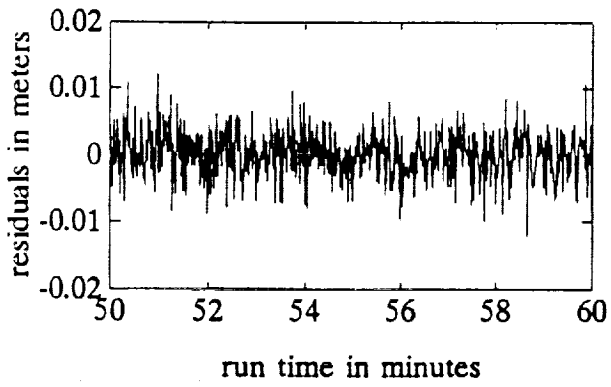


Figure 17. SV 28 residuals

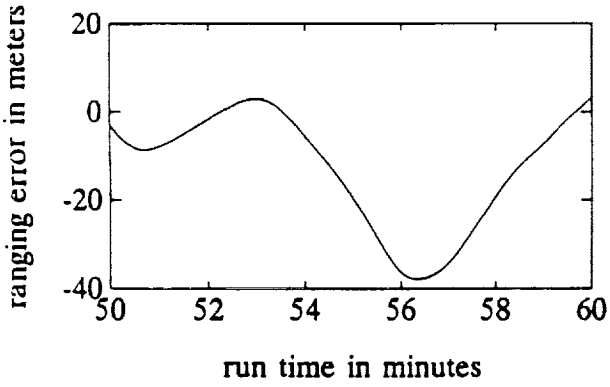


Figure 18. SA model output - expanded scale

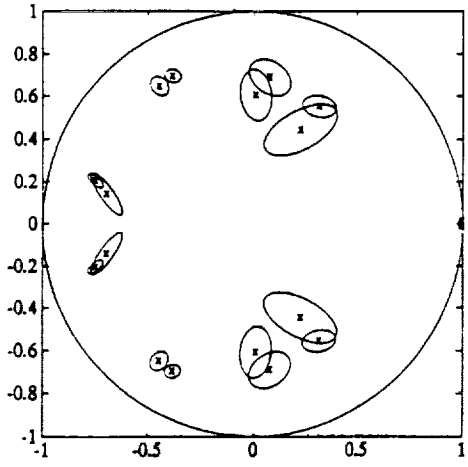


Figure 19. Pole comparison

for the model derived from the SV 28 data will be listed here:

a(1) = -1.36192741558063
a(2) = -0.15866710938728
a(3) = +0.13545921610672
a(4) = +0.21501267664869
a(5) = +0.30061078095966
a(6) = -0.12390183286070
a(7) = +0.10063573000351
a(8) = +0.02694677520401
a(9) = -0.12898590228866
a(10) = +0.05083106570666
a(11) = -0.05600186282898

$\sigma_e^2 = 1.6993 \times 10^{-5}$ (meters^2)

σ_e^2 is the variance of the Gaussian white noise input. The seemingly excessive amount of significant figures is required to ensure filter stability. Note in figure 19 that three out of the eleven poles are extremely close to the unit circle. Truncation of the coefficients can cause these poles to move outside the unit circle yielding instability. It is thus very important that the significant figures be maintained. Towards this end, it is suggested that double-precision arithmetic be employed in the generation of SA waveforms using this model.

The distribution of these poles makes sense from the point of view of filter theory. The three poles grouped near the unit circle and the real axis represent a type of low-pass filter with an extremely narrow bandwidth. This is necessary since the input to the filter is wide-band noise and the output is extremely narrow-band SA. Although the low frequency components dominate the SA waveform, higher frequencies are present also and the other poles of the model contribute to these components.

Stationarity

As was mentioned earlier, some of the collected data records had to be truncated in order to achieve stationarity. A random process is said to be stationary if its statistics do not change with time. Unfortunately, some of the original collected SA records did exhibit non-stationary behavior. Another way of viewing this is to assume that SA is truly generated by a time series model but that the coefficients change as a function of time. Powerful as they are, the vast majority of model identification techniques assume a stationary data record. Non-stationary records are typically examined by segmenting the data into stationary sections and identifying a model for each one separately. The non-stationary behavior of the data then can be determined by examining the change in the models from segment

to segment [Marple, 1987].

Having SA with this kind of behavior makes sense, at least from a security point of view. A non-stationary random process is much harder to "crack" than a stationary one. It should be pointed out, however, that the collected data did exhibit stationarity for periods of up to one and a half hours. Since this was the maximum data collection period, no conclusions can be made for longer periods. Future data collection efforts are being planned to examine this phenomenon more closely. In the mean time, the SA models derived from the data are good approximations to the truth.

CONCLUSIONS AND RECOMMENDATIONS

Simulations are often necessary in the process of development and testing of GPS-based systems. For those users of GPS not having the benefits of DGPS corrections, SA represents the dominant source of error. For would-be developers of DGPS systems, SA dictates the trade-off between the update rate (of the differential corrections) and system accuracy. Simulations therefore must account for SA. In this paper, the issue of SA analysis and modeling has been revisited. Using post-processed, precise ephemeris data, a technique has been described whereby the clock and orbital components of SA can be identified separately. For the data collected for this paper, the orbital component of SA seems not to have been implemented.

SA data (clock component) has been collected from over half of the current Block II satellites and a robust model has been derived. The model has been demonstrated to be accurate and robust. It is suggested that this model be implemented in GPS receiver test equipment and in GPS-based system simulations. Since the model is capable of generating virtually unlimited amounts of data, the design and test engineers need not be constrained to a few collected waveforms.

ACKNOWLEDGEMENTS

The work presented in this paper was funded in part by the National Aeronautics and Space Administration under the Joint University Program for Air Transportation Research Grant NGR 36-009-017. Ashtech, Inc. is thanked for their provision of the GPS receiver which was used in the collection of the broadcast ephemeris data. Dr. Frank van Graas of the Avionics Engineering Center is thanked for his support.

REFERENCES

Ananda, M., H. Bernstein, R. Bruce, K. Cunningham, W. Feess, P. Jorgensen, M. Menn and C. Price (1984) "Autonomous Navigation of the Global Positioning System Satellites," AIAA Guidance and Control Conference, Seattle, WA, August.

Anonymous (1990), "Newsfront" GPS World, May/June.

Bar-Sever, Y., T. Yunck and S. Wu (1990), "GPS-Based Orbit Determination and Point Positioning Under Selective Availability," ION GPS-90, Colorado Springs, CO, September 19-21.

Bowen, R., P. Swanson, F. Winn, N. Rhodus and W. Feess (1985), "Global Positioning System Operational Control System Accuracies," NAVIGATION: Journal of the Institute of Navigation, Vol. 32, No. 2, Summer.

Braasch, M. (1990-91), "A Signal Model for GPS," NAVIGATION: Journal of the Institute of Navigation, Vol. 37, No. 4, Winter.

Braasch, M. (1992), "On The Characterization Of Multipath Errors In Satellite-Based Precision Approach and Landing Systems," Ph.D. dissertation, Department of Electrical and Computer Engineering, Ohio University, Athens, Ohio, June.

Chou, H. (1990), "An Anti-SA Filter for Non-differential GPS Users," ION GPS-90, Colorado Springs, CO, September 19-21.

Feit, C. (1992), "GPS Range Updates in an Automatic Flight Inspection System: Simulation, Static and Flight Test Results," ION GPS-92, Albuquerque, NM, September 16-18.

Kay, S. (1987), Modern Spectral Estimation: theory and application, Prentice-Hall, Englewood Cliffs, NJ.

Kelly, R. (1992), "MLS System Error Model Identification and Synthesis," IEEE Trans. on Aerospace and Electronics Systems, Vol. 28, No. 1, January.

Kremer, G., R. Kalafus, P. Loomis and J. Reynolds (1990), "The Effect of Selective Availability on Differential GPS Corrections," NAVIGATION: Journal of the Institute of Navigation, Vol. 37, No. 1, Spring.

Lear, W., M. Montez, L. Rater and L. Zyla (1992), "The Effect of Selective Availability on Orbit Space Vehicles Equipped with SPS GPS Receivers," ION GPS-92, Albuquerque, NM, September 16-18.

Ljung, L. (1987), System Identification - Theory for the User, Prentice-Hall, Englewood Cliffs, NJ.

Marple, S. L. (Jr.) (1987), Digital Spectral Analysis with Applications, Prentice-Hall, Englewood Cliffs, NJ.

Matchett, G. (1985), "Stochastic Simulation of GPS Selective Availability Errors," Technical Memorandum, FAA Contract DTRS-57-83-C-00077, June.

Remondi, B. (1985), "Distribution of Global Positioning System Ephemerides by the National Geodetic Survey," First Conference on Civil Applications of GPS - ION, Greater Philadelphia Section, Warminster, PA, September 12.

Remondi, B. (1989), "Extending the National Geodetic Survey Standard GPS Orbit Formats," NOAA Technical Report NOS 133 NGS 46, National Geodetic Information Branch, National Oceanic and Atmospheric Administration, Rockville, MD, November.

Remondi, B. (1991), "NGS Second Generation ASCII and Binary Orbit Formats and Associated Interpolation Studies," Twentieth General Assembly of the International Union of Geodesy and Geophysics, Vienna, Austria, August 11-24.

REALTIME MITIGATION OF GPS SA ERRORS USING LORAN-C

Soo Y. Braasch

Avionics Engineering Center
Ohio University
Athens, Ohio, USA

Abstract

The hybrid use of Loran-C with the Global Positioning System (GPS) has been shown capable of providing a sole-means of enroute air radionavigation. By allowing pilots to fly direct to their destinations, use of this system is resulting in significant time savings and therefore fuel savings as well. However, a major error source limiting the accuracy of GPS is the intentional degradation of the GPS signal known as Selective Availability (SA). SA-induced position errors are highly correlated and far exceed all other error sources (horizontal position error: 100 meters, 95%). Realtime mitigation of SA errors from the position solution is highly desirable. This paper discusses how that can be achieved. The stability of Loran-C signals is exploited to reduce SA errors. The theory behind this technique will be discussed and results using bench and flight data will be given.

Introduction

The hybrid use of Loran-C with the Global Positioning System (GPS) has been shown to be capable of providing a sole-means of enroute air radionavigation [1]. Standardization committees such as the RTCA are currently working on developing minimum operational performance standards for this system. By allowing pilots to fly direct to their destinations, use of this system will result in significant time savings and therefore fuel savings as well. By not confining all aircraft to a small portion of the airspace (which results when using the Victor airways), the risk of collision undoubtedly will be reduced as well.

However, a major error source limiting the accuracy of GPS is the intentional degradation of the GPS signal known as Selective Availability (SA). SA manifests itself in the form of erroneous orbital data broadcast by the satellites and in dithering of the satellite clock. The result is position determination which, according to the Department of Defense (DoD), will be in error by one hundred meters 95% of the time in the horizontal plane. Previous work performed at Ohio University showed that SA-induced position errors are highly correlated [2]. Since the correlation time is on the order of minutes, it follows that the error falls well within the passband of the aircraft's dynamic response. The result is that the aircraft will follow the deviations induced by SA.

Realtime mitigation of SA errors from the position solution is highly desirable. This paper discusses how that can be achieved. The stability of Loran-C signals is exploited to reduce SA errors. In the typical hybrid use of Loran-C and GPS, the Loran-C signal stability is not exploited. This stems from the relatively poor absolute accuracy of Loran-C (relative to GPS). However, it is possible to use the stability of Loran-C positioning to reduce SA-induced GPS positioning errors. The theory behind this technique will be discussed and results will be given. First, the phenomenon of SA will be described.

Selective Availability

As mentioned in the introduction, SA is an intentional corruption of the GPS signal by the DoD to limit the accuracy available to the public. The degradation is achieved in two ways. First, false satellite orbit parameters are broadcast to the users. This results in incorrect positioning of the satellites in the navigation solution. Secondly, code and carrier tracking errors are induced through dithering the satellite clock (carrier frequency). The erroneous orbit parameters lead to position errors which vary slowly throughout the satellite pass. Code-phase and carrier-phase errors due to the dithering of the satellite clock are random but also are highly correlated. Correlation times of several minutes are typical. As a result, simple filtering schemes are not effective and aircraft will follow the deviations. Virtually all of the information available to date about SA has been gathered through data collection efforts by civilian organizations. The DoD, however, has stated that SA shall be instituted in such a way as to yield horizontal position errors at a 95% level of 100 meters [3].

Mitigation Methodology

The heart of the mitigation scheme lies in the differences between Loran-C and SA-induced GPS position errors. Loran-C position errors in general are biased and noisy. The level of noise depends upon the receiver architecture

Wild Goose Association, Annual Convention and Technical Symposium, August 24-27, 1992, Birmingham, England

and specifically upon the tracking loop bandwidth. However, noise levels on the order of 5 to 10 meters can be achieved for airborne applications. The Loran-C position bias is primarily composed of unmodeled additional secondary phase factors (ASF). In general the bias does not remain constant over any given flight path but the variation is usually quite slow in comparison to the clock component of GPS SA error. This phenomenon is what makes Loran-based SA mitigation possible. The long-term stability of the Loran-C measurements is exploited to smooth the SA-induced variations in the GPS measurements.

Conceptually, the mitigation scheme works as follows. The Loran-C sensor computes the horizontal position of the aircraft. A vertical input is needed and is supplied by the barometric altimeter (again, a sensor which is biased but stable). The combination provides a three-dimensional position of the aircraft. Range values are computed from the GPS satellites to the Loran-C/altimeter position. These range values are then the reference against which the measured GPS pseudoranges are filtered.

Note that the technique depends upon the assumption that SA error is composed only of high frequency components relative to the Loran-C bias error variations. Strictly speaking, this assumption is not valid since the orbital component of SA error is slowly varying. However, as was shown in [2], the clock component of SA error has periods on the order of five to ten minutes. As such it is a high frequency error source relative to the non-noise component of Loran-C error. Although this has not been rigorously proven, flight data (to be shown later) supports the conclusion. Thus, the technique is able to reduce the clock component (or roughly speaking, the variance) of SA error.

The filtering is accomplished by complementary Kalman filters which are applied to each pseudorange measurement [4,5]. The inputs to each filter are the given GPS pseudorange measurement and the corresponding range computed from the satellite to the Loran-C/altimeter position. At each measurement epoch (current time given by the index k), the complementary Kalman filter is executed as follows:

$$d_k^- = d_{k-1}^+ + (L_k - L_{k-1}) \qquad (1)$$

$$p_k^- = p_{k-1}^+ + q \qquad (2)$$

$$k_k = \frac{p_k^-}{p_k^- + r} \qquad (3)$$

$$d_k^+ = d_k^- + k_k(z_k - d_k^-) \qquad (4)$$

$$p_k^+ = (1 - k_k)p_k^- \qquad (5)$$

where the subscript represents the time index. The superscripts '-'and '+'represent predicted and estimated quantities respectively. 'd$^+$' represents the estimated pseudorange with variance q. 'z'represents the measured pseudorange with error variance r. Note that r is due primarily to SA. 'L'represents the range computed from the satellite to the Loran-C/altimeter position. 'p' represents the prediction or estimation error variance. 'k' is the Kalman gain. In equation (1), the current pseudorange prediction is computed by updating the previous pseudorange estimate with the difference between the current and previous Loran-C/altimeter ranges. The prediction error variance is computed in equation (2) and is used to compute the Kalman gain in equation (3). The difference between the measured and predicted pseudoranges is weighted by the Kalman gain in the computation of the current estimate (equation 4). Finally, the current estimation error variance is computed (equation 5).

Position Solution

Given at least four GPS pseudoranges, position may be computed. As will be shown in the next two sections, significant reduction in SA-error may be achieved when using the mitigation technique just described.

For both the simulation and flight test results (to be shown later), the ordinary least-squares (OLS) estimator is used to determine position and clock bias from the pseudoranges. In the absence of measurement errors, the relationship between satellite location, receiver location, clock bias and pseudorange is given by:

$$R_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} + b \qquad (6)$$

where R_i is the pseudorange to the i^{th} satellite, $(x_i,y_i,z_i)$are the coordinates of the satellite, (x,y,z)are the coordinates of the receiver and b is the receiver clock bias (converted to units of distance through multiplication by the speed of light). Since the receiver coordinates and clock bias must be solved for simultaneously, at least four measurements are required.

However, instead of attempting simultaneous solution of non-linear equations, the standard technique is to solve iteratively a set of equations which have been linearized about an initial estimated position and clock bias

(x_o, y_o, z_o, b_o). This is achieved by forming a Taylor series expansion and retaining the zeroth and first order terms:

$$R_i = R_{io} + (\delta x)\frac{\partial R_i}{\partial x}\Big|_{x_o y_o z_o} + (\delta y)\frac{\partial R_i}{\partial y}\Big|_{x_o y_o z_o} + (\delta z)\frac{\partial R_i}{\partial y}\Big|_{x_o y_o z_o} + (\delta b)\frac{\partial R_i}{\partial b} \quad (7)$$

where R_{io} is the range from the satellite to the initial position estimate. δx, δy, δz and δb represent the corrections to the initial estimates. If the initial estimate is close to the truth, no iterations are required. However, if the initial estimate is not close, the corrections are used to update the initial estimate and the process is repeated. Convergence is declared if the magnitudes of the corrections are below a desired threshold.

The partial derivatives are evaluated as follows:

$$\frac{\partial R_i}{\partial x} = \frac{x_o - x}{R_i - b_o} = \alpha_{i1} \quad (8)$$

$$\frac{\partial R_i}{\partial y} = \frac{y_o - y}{R_i - b_o} = \alpha_{i2} \quad (9)$$

$$\frac{\partial R_i}{\partial z} = \frac{z_o - z}{R_i - b_o} = \alpha_{i3} \quad (10)$$

$$\frac{\partial R_i}{\partial b} = 1 = \alpha_{i4} \quad (11)$$

Substitution of equations (8) through (11) into (7) yields:

$$\delta R_i = (\delta x)\alpha_{i1} + (\delta y)\alpha_{i2} + (\delta z)\alpha_{i3} + (\delta b)\alpha_{i4} \quad (12)$$

where:

$$\delta R_i = R_i - R_{io} \quad (13)$$

Four pseudorange measurements allow for the following simultaneous set of equations:

$$\begin{pmatrix} \delta R_1 \\ \delta R_2 \\ \delta R_3 \\ \delta R_4 \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{pmatrix} \begin{pmatrix} \delta x \\ \delta y \\ \delta z \\ \delta b \end{pmatrix} \quad (14)$$

which may be rewritten more succinctly:

$$\underline{y} = H\underline{\beta} \quad (15)$$

The presence of measurement errors may be accounted for by the addition of an error vector:

$$\underline{y} = H\underline{\beta} + \underline{\varepsilon} \quad (16)$$

The ordinary least-squares solution is then given by:

$$\underline{\hat{\beta}}_{OLS} = (H^T H)^{-1} H^T \underline{y} \quad (17)$$

After one iteration then, the position and clock bias estimate is given by:

$$\begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} x_o \\ y_o \\ z_o \\ b_o \end{pmatrix} + \begin{pmatrix} \delta x \\ \delta y \\ \delta z \\ \delta b \end{pmatrix} \quad (18)$$

Simulation

To determine the feasibility of the technique, a simulation was performed. A simple flight-path was modeled with the aircraft traveling to the east for 900 seconds at 100 meters/second, followed by a 2g turn and then returning to the west (figure 1). For the sake of simplicity in the calculations, a static satellite constellation was modeled. In order to focus on the effects of SA, all other GPS error sources were assumed to be zero. The Loran-C/altimeter errors were modeled in the position domain by a constant 200 meter bias on each axis.

The SA model was obtained from collected data using the System Identification procedure described in [2]. In order to model SA rather than the combination of SA and receiver noise, integrated Doppler data (rather than pseudorange data) were used. The System Identification procedure yielded a 16th order autoregressive (AR) filter. When Gaussian white noise (of proper variance) is input to this filter model, the output is statistically equivalent to the collected SA data. An example of the output is given in figure 2.

The positioning errors resulting from the SA corruption are given in figures 3 and 4. Both the east and north components of the position error exhibit similar characteristics to the SA error on the pseudorange measurements. As discussed earlier, the errors are highly correlated and reach up to 100 meters. However, use of the Loran-C/altimeter data in the complementary Kalman filter yields significant reduction of SA error (figures 5 and 6).

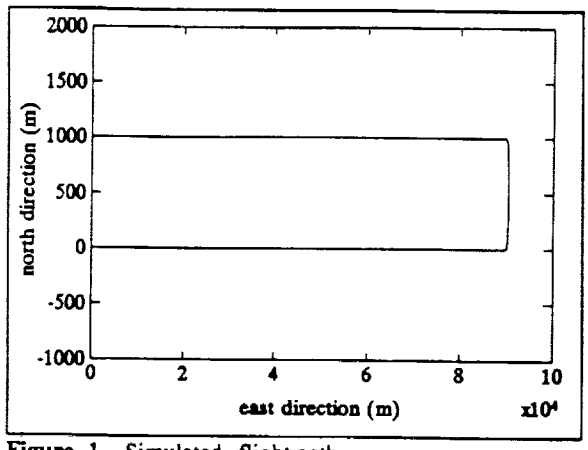


Figure 1. Simulated flight path.

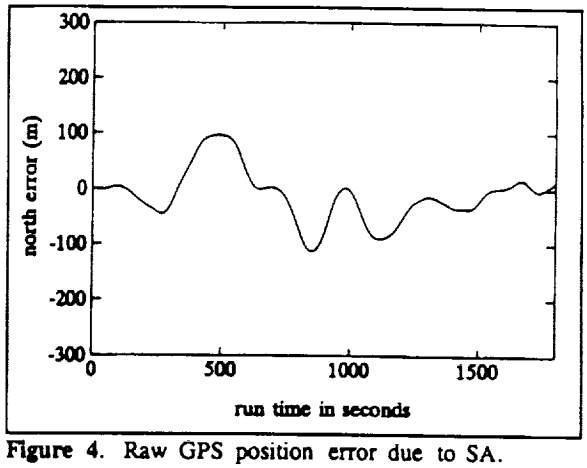


Figure 4. Raw GPS position error due to SA.

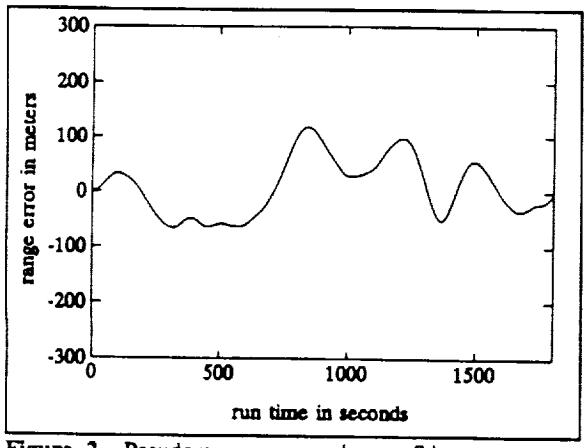


Figure 2. Pseudorange error due to SA.

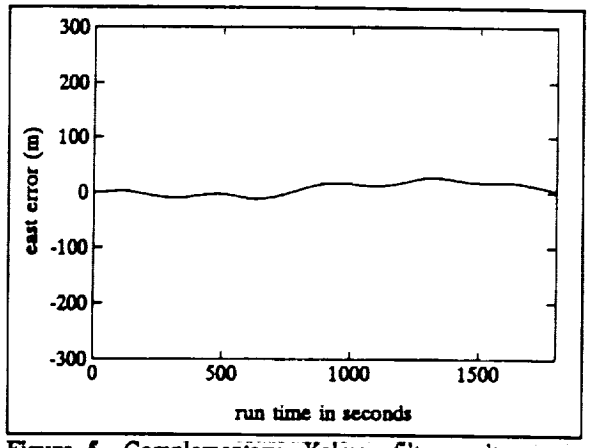


Figure 5. Complementary Kalman filter results.

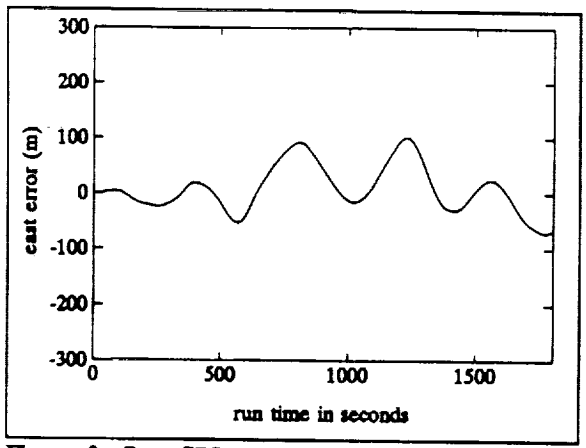


Figure 3. Raw GPS position error due to SA.

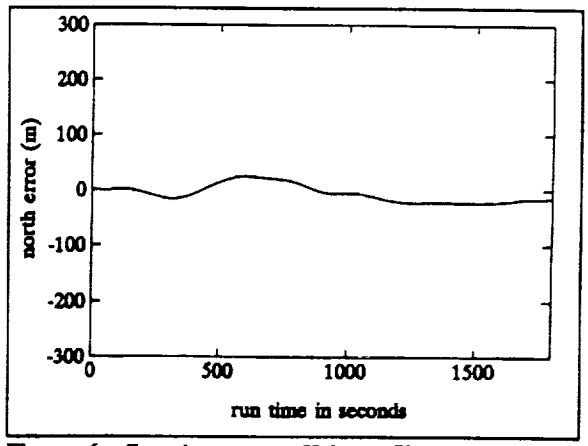


Figure 6. Complementary Kalman filter results.

Flight Test

Although extremely encouraging, the simulation results were obtained using a simple model for Loran-C position errors. In order to verify the robustness of the technique, actual flight data was used. This is necessary since Loran-C position error bias is spatially dependent.

The flight data employed here were collected during a trip from Cleveland to Athens, Ohio in Fall of 1990 (figure 7). It may be recalled that SA was temporarily turned off at that time because of military use of civilian GPS receivers during Operation Desert Shield [6]. As a result, the GPS horizontal positioning accuracy is on the order of 10-20 meters [1]. For this flight, the GPS-derived position was therefore used as a rough truth reference.

SA was generated by the model described earlier and added to the raw GPS pseudorange measurements (figure 8). As expected, the Loran-C position error is biased but the bias is not constant with position (figures 9 and 10). As was done earlier, altimeter error was modeled as a constant 200 meter bias. Raw SA-induced position errors are as expected with large excursions and high correlation (figures 11 and 12). Again, position errors after smoothing are significantly reduced (figures 13 and 14). It is important to note that even in the face of spatially varying Loran-C position errors, the mitigation scheme continues to perform well.

Conclusions

A technique has been described whereby the stability of Loran-C signals are exploited to reduce SA-induced GPS position errors. The viability of the technique has been confirmed using simulations as well as actual flight data. Future work will consider the possibility of realtime SA model identification.

Acknowledgements

This work was supported by the Federal Aviation Administration and the National Aeronautics and Space Administration through the Joint University Program for Air Transportation Research (Grant NGR 36-009-17). The author would like to express her gratitude to the entire staff of the Avionics Engineering Center at Ohio University for their encouragement and support. Special thanks are also given to Professor Robert W. Lilley and Professor Durk van Willigen for their careful review of this paper.

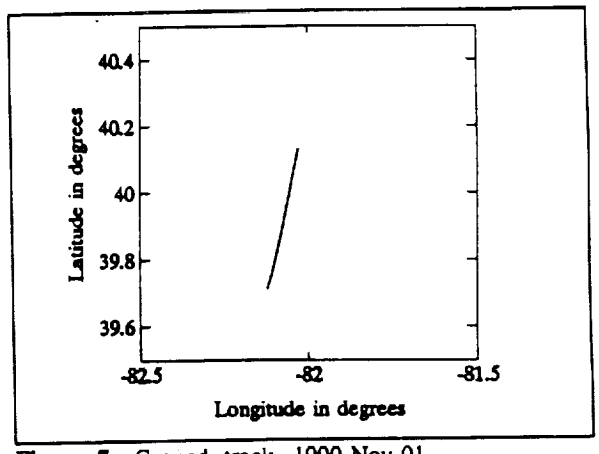


Figure 7. Ground track, 1990-Nov-01.

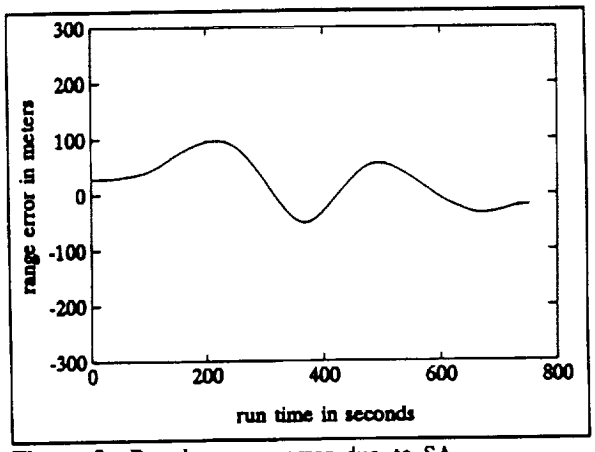


Figure 8. Pseudorange error due to SA.

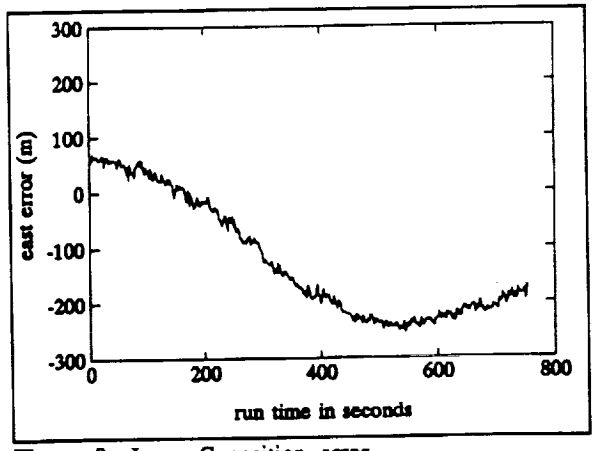


Figure 9. Loran-C position error.

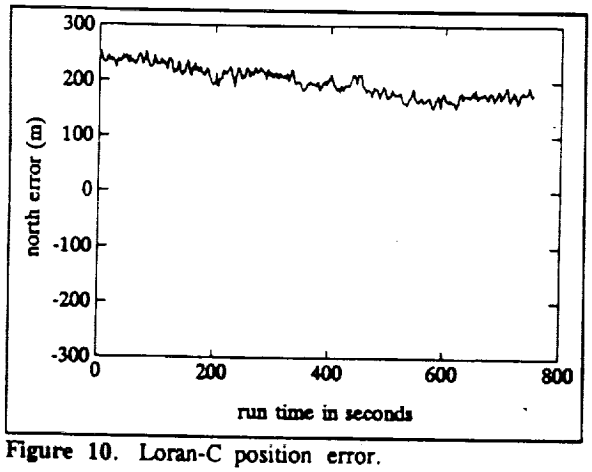


Figure 10. Loran-C position error.

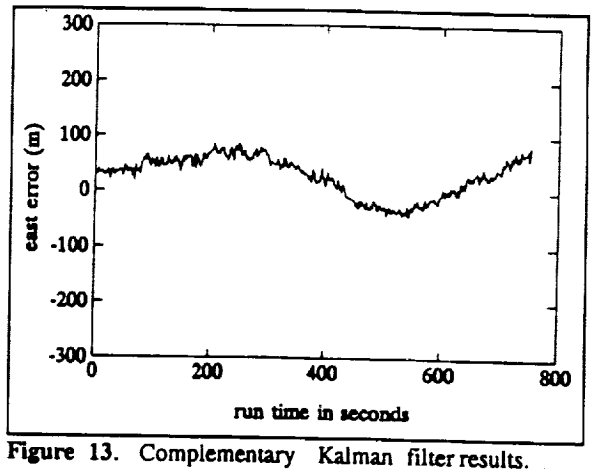


Figure 13. Complementary Kalman filter results.

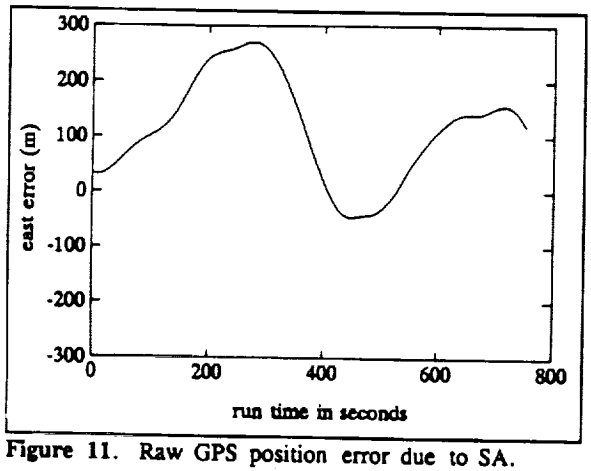


Figure 11. Raw GPS position error due to SA.

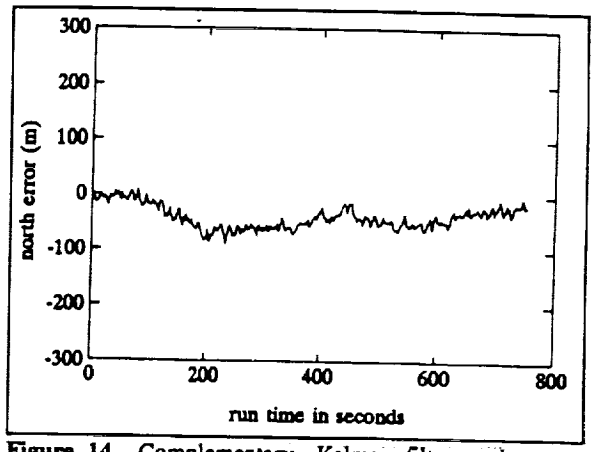


Figure 14. Complementary Kalman filter results.

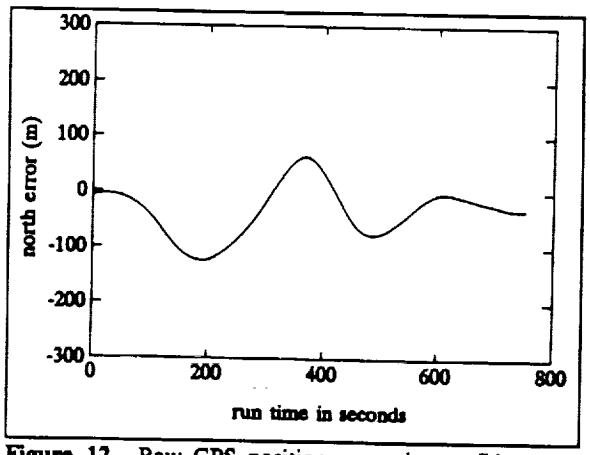


Figure 12. Raw GPS position error due to SA.

References

1. van Graas, F. "Hybrid GPS/Loran-C: A Next-Generation Of Sole Means Air Navigation," Ph.D. Dissertation, Department of Electrical and Computer Engineering, Ohio University, Athens, Ohio, November 1988.

2. Braasch, M. "A Signal Model for GPS," NAVIGATION: Journal of the Institute of Navigation, Vol. 37, No. 4, Winter 1990-91.

3. Wysocki, J. "GPS and Selective Availability - The Military Perspective," GPS World, July/August 1991.

4. Hwang, P. and R. G. Brown, "GPS Navigation: Combining Pseudorange with Continuous Carrier Phase Using a Kalman Filter," NAVIGATION: Journal of the Institute of Navigation, Vol. 37, No. 2, Summer 1990.

5. van Graas, F. and M. Braasch, "GPS Interferometric Attitude and Heading Determination: Initial Flight Test Results," NAVIGATION: Journal of the Institute of Navigation, Vol. 38, No. 4, Winter 1991-92.

6. Anonymous, "Antispoofing, Selective Availability Moves Surprise Civil Users," Global View, GPS World, November/December 1990.

58-04

207602

N94-27292

A GPS COVERAGE MODEL

Trent A Skidmore
Avionics Engineering Center
Ohio University

BIOGRAPHY

Dr. Trent A. Skidmore is on staff at the Ohio University Avionics Engineering Center where he is currently involved in the development of a Global Positioning System (GPS) coverage model. He also serves as principal investigator for the Visual EEG Tracking System (VETS) program and contributes to the improvement of the very-high-frequency omnirange (VOR) en route navigation system. He is a graduate of Ohio University (Ph.D., 1991) and Michigan Technological University (M.S., 1988 and B.S., 1986), all in electrical engineering. His research interests include digital signal processing, communication theory, and biomedical engineering.

ABSTRACT

This paper summarizes the results of several case studies using the Global Positioning System coverage model developed at Ohio University. Presented are results pertaining to outage area, outage dynamics, and availability. Input parameters to the model include the satellite orbit data, service area of interest, geometry requirements, and horizon and antenna mask angles. It is shown for precision-landing Category I requirements that the planned GPS 21 Primary Satellite Constellation produces significant outage area and unavailability. It is also shown that a decrease in the user equivalent range error dramatically decreases outage area and improves the service availability.

1. INTRODUCTION

An excellent summary documenting the impending need for a comprehensive Global Positioning System (GPS) satellite coverage model is given in [1]. To be complete, this summary is repeated here:

"The continuous movement of navigation satellites with respect to the surface of the earth results in continual changes of the system coverage. There may be times when the number of satellites in view of an aircraft near a particular airport would be less than that required for executing a precision approach. The periods of time when precision approach coverage will be inadequate at given airports must be known well in advance in order that operations may be restricted.

A satellite-based precision approach system requires a high level of availability within the service region to ensure operational suitability of the system. At the present time, a precise requirement for availability is not defined; however, preliminary studies indicate that system unavailability should be well below one minute per day. Critical sources of unavailability result from poor satellite geometry, planned satellite down time, known satellite failures, and planned ground equipment down time (e.g., a differential reference station). Thus the majority of the satellite system unavailability is predictable. The primary consequence of predictable unavailability is the need to schedule around the known outages. Since a single satellite covers a large geographical area, a satellite outage could potentially affect a large service area. This would result in major operation, capacity, and economic concerns. For instance, a one-hour outage in a metropolitan area would result in multiple simultaneous missed approaches and simultaneous replanning for many aircraft in the air.

Note that unpredictable outages are primarily a safety concern because of their significant effect on the guidance of aircraft during the approach and landing phase. The contribution of unpredictable outages to the overall system unavailability is anticipated to be small compared to the predictable outages, but this assumption must be verified.

A computer model would be used initially to characterize the coverage, and to analyze the size, duration, and dynamics of the outage areas under a wide variety of failure scenarios and for different system architectures. Input parameters to the model

Presented at the Institute of Navigation 1992 National Technical Meeting, Washington, DC, June 29 - July 1, 1992

would include the service area of interest, satellite orbit data, geometry requirements, horizon and antenna mask angles, and satellite reliability data.

The descent of the aircraft while on an approach, along with the movement of navigation satellites, may also result in different optimal sets of satellites guidance at the initiation and at the conclusion of the approach. The impact of using a four or five channel receiver with potential satellite switching during the approach versus an all-in-view receiver must be addressed. The computer model can then be used as a tool to evaluate different system architectures.

...The criticality of this issue is judged to be HIGH, since coverage definition is necessary for the assurance of adequate system performance."

In order to address the GPS coverage issues, the Ohio University Avionics Engineering Center has been developing a comprehensive GPS coverage model. The different modules that comprise the model have been used in various applications as documented in reference [2]. This paper summarizes the most recent developments and highlights the model's unique features and capabilities. The results of several case studies are presented in order to gain an appreciation for the types of parametric studies the model will facilitate. The presentation concludes with a brief summary of additional work that is necessary in order to allow the GPS coverage model to be used as a complete system-analysis tool. It should be emphasized that the current model is capable of evaluating not only the present satellite architecture, but will eventually become a tool for designing and evaluating alternative satellite-based navigation architectures to meet precision-approach requirements.

2. CASE STUDIES

2.1 Introduction

This section details the results of several case studies involving the various modules which comprise the coverage model. For each case study presented, the test conditions are stated and results given. This is followed by a discussion and summary of the important conclusions.

The two case-study scenarios analyzed by the coverage model are summarized in Table 2.1. The parameters displayed in the table were chosen for validation purposes and to determine a near-global perspective on the system performance. Although these were chosen to be representative of the model's capabilities, additional work is still needed in order to develop minimum standards for time and space increments.

Table 2.1 Case-Study Initial Conditions

Test Conditions	Test 1 (World)	Test 2 (North America)
Max. N Lat.	90°	75°
Max. S Lat.	-90°	15°
Max W Long.	-180°	-170°
Max. E Long.	0°	-50°
Increment	5°	6°
Min. HDOP	-	2.3
Min. VDOP	-	9.2
Min. PDOP	6	-
Analysis Time	24 hours	24 hours
Time Increment	5 minutes	6 minutes
Constellation	Optimal 21	21 Primary

2.2 Case Study I: Counting Outages

As a first attempt at outage characterization, the duration and number of zero-failed satellite outages at each location in the search grid were determined. Figure 2.1 shows the Test-1 (World) outage contours. This result is essentially the same as that presented by Jorgensen [3]. Note that, in this case, even the complete constellation results in substantial outages. This is due to the fundamental limitations imposed by using the Optimal 21 Satellite Constellation. The inclusion of this result is not intended to be an analysis of the Optimal 21 Constellation, but is presented because of its importance to the validation of our model. The case studies shown throughout the remainder of the paper will be concerned exclusively with the GPS 21 Primary Satellite Constellation [4].

2.3 Case Study II: Outage Areas versus DOP

Shown in Table 2.2 are the Vertical Dilution of Precision (VDOP) requirements for the various categories of approach assuming a 6-foot user equivalent range error (95%) [5]. Table 2.3 expands upon this for the Category I landing by showing the required VDOP for different values of user equivalent range error (UERE). Throughout the paper, the required maximum allowable HDOP was chosen to be four times the specified VDOP.

The effect of varying DOP requirements (VDOP and/or HDOP) is of particular importance, especially when considering precision-approach issues. To determine the impact that the DOP requirement has on outage area, the model was used to characterize the outages based on the initial conditions set forth in Test 2 (North America). For this test, the GPS 21 Primary Constellation (as shown in Figure 2.2) was analyzed at three different DOP values for up to three failed

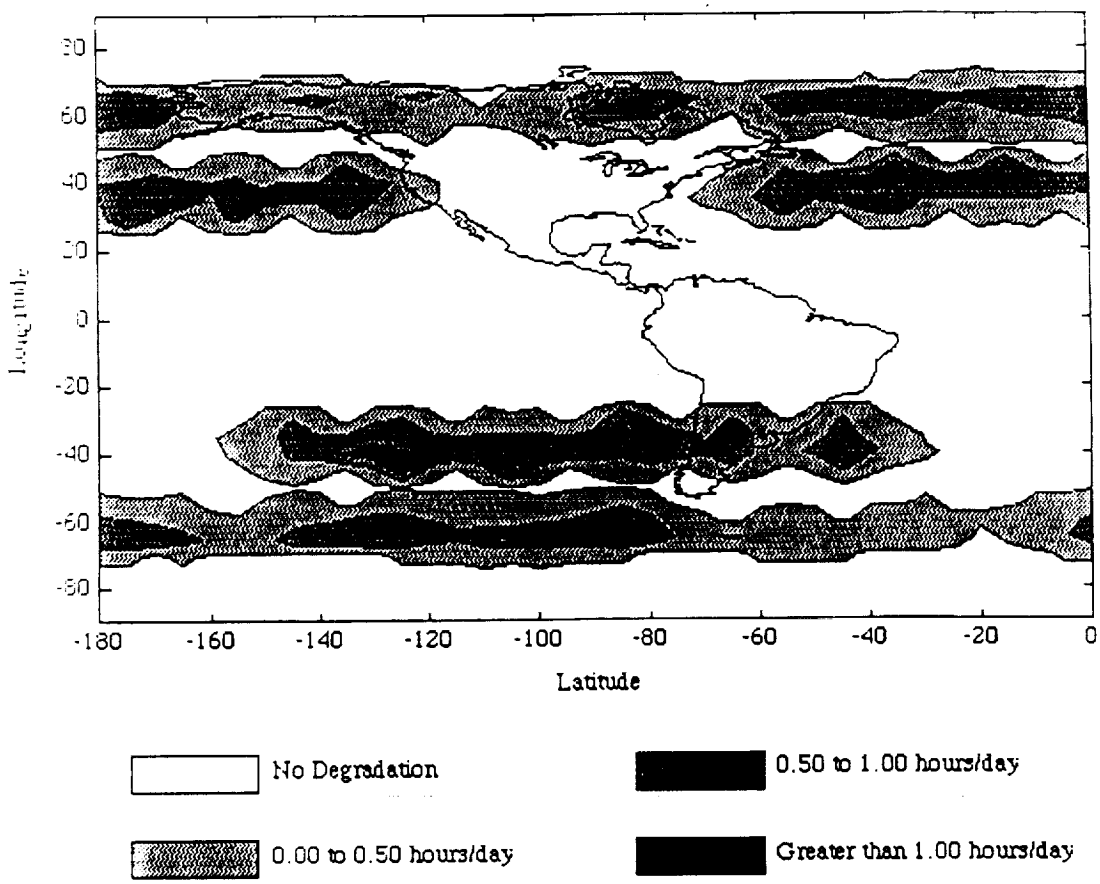


Figure 2.1 Accumulated Outages Per Day for Test 1. Contours show regions of degraded performance based on no failures in the Optimal 21 Satellite Constellation.

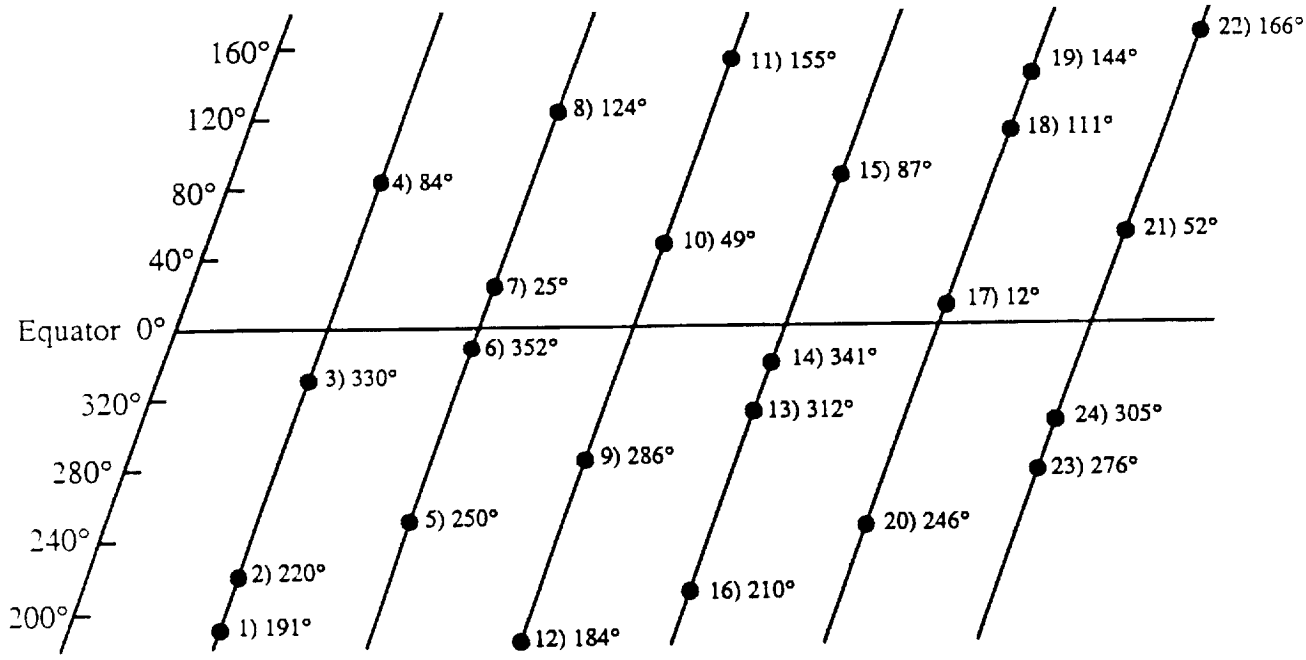


Figure 2.2 The GPS 21 Primary Satellite Constellation. The satellite numbering is based on the order in which the satellite ephemeris data is entered into the model.

satellites. Table 2.4 summarizes the worst and average satellite-failure combinations for each DOP condition.

Table 2.2 Vertical Accuracy as a Function of VDOP

Approach Category	Vertical Requirement (feet, 95%)	VDOP Requirement
I	13.5	2.3
II	5.6	0.9
III	2.0	0.3

Table 2.3 The Cat I Approach: VDOP and UERE

UERE (ft, 95%)	VDOP Requirement	Associated HDOP
6	2.3	9.2
3	4.6	18.4
1	13.8	55.2

Table 2.4 Worst and Average Satellite Failures

(HDOP, VDOP)	Worst Single Failure	Average Single Failure
(9.2, 2.3)	(22)	(11)
(18.4, 4.6)	(8)	(16)
(55.2, 13.8)	(22)	(20)
	Worst Double Failure	Average Double Failure
(9.2, 2.3)	(9, 22)	(5, 11)
(18.4, 4.6)	(6, 9)	(9, 20)
(55.2, 13.8)	(20, 22)	(5, 24)
	Worst Triple Failure	Average Triple Failure
(9.2, 2.3)	(10, 15, 22)	(1, 8, 23)
(18.4, 4.6)	(3, 6, 9)	(9, 10, 16)
(55.2, 13.8)	(1, 8, 22)	(8, 11, 19)

Note: Satellite numbering is based on Figure 2.2.

The most important observation that can be made based on an analysis of the data contained in Table 2.4 is that the worst- and average-failure combinations change as a function of HDOP and VDOP. This is evidence to the fact that the coverage provided by the 21 Primary GPS Constellation is highly nonlinear and nonuniform. Figure 2.3 also illustrates this nonlinearity by graphically displaying the average and worst-case outage area as a function of DOP and the number of failed satellites. To further illustrate this nonlinearity, consider Tables 2.5 and 2.6. Shown here are the ratios of outage-area decrease as a function of VDOP (and HDOP) relaxation. For example, from the worst-case single failure of Table 2.5 it can be seen that relaxing the (HDOP, VDOP) requirement by a factor of six decreases the corresponding outage area by a factor of 210, with pronounced area reductions occurring in the other cases as well. It is also interesting that DOP relaxation has a greater effect on the average failure combination than on the worst-case failure combination. This may be due to the fact that the worst-case failure combinations are highly sensitive to DOP requirements.

Table 2.5 VDOP Relaxation and Outage Area (The Worst-Case Satellite Failure)

Number of failed SVs	Relax VDOP (2.3 : 4.6) (1 : 2)	Relax VDOP (2.3 : 13.8) (1 : 6)
0	30.64	752.05
1	15.08	210.10
2	5.93	25.94
3	3.37	7.91

Table 2.6 VDOP Relaxation and Outage Area (The Average Satellite Failure)

Number of failed SVs	Relax VDOP (2.3 : 4.6) (1 : 2)	Relax VDOP (2.3 : 13.8) (1 : 6)
0	30.64	752.05
1	20.25	356.22
2	12.66	109.91
3	8.50	44.34

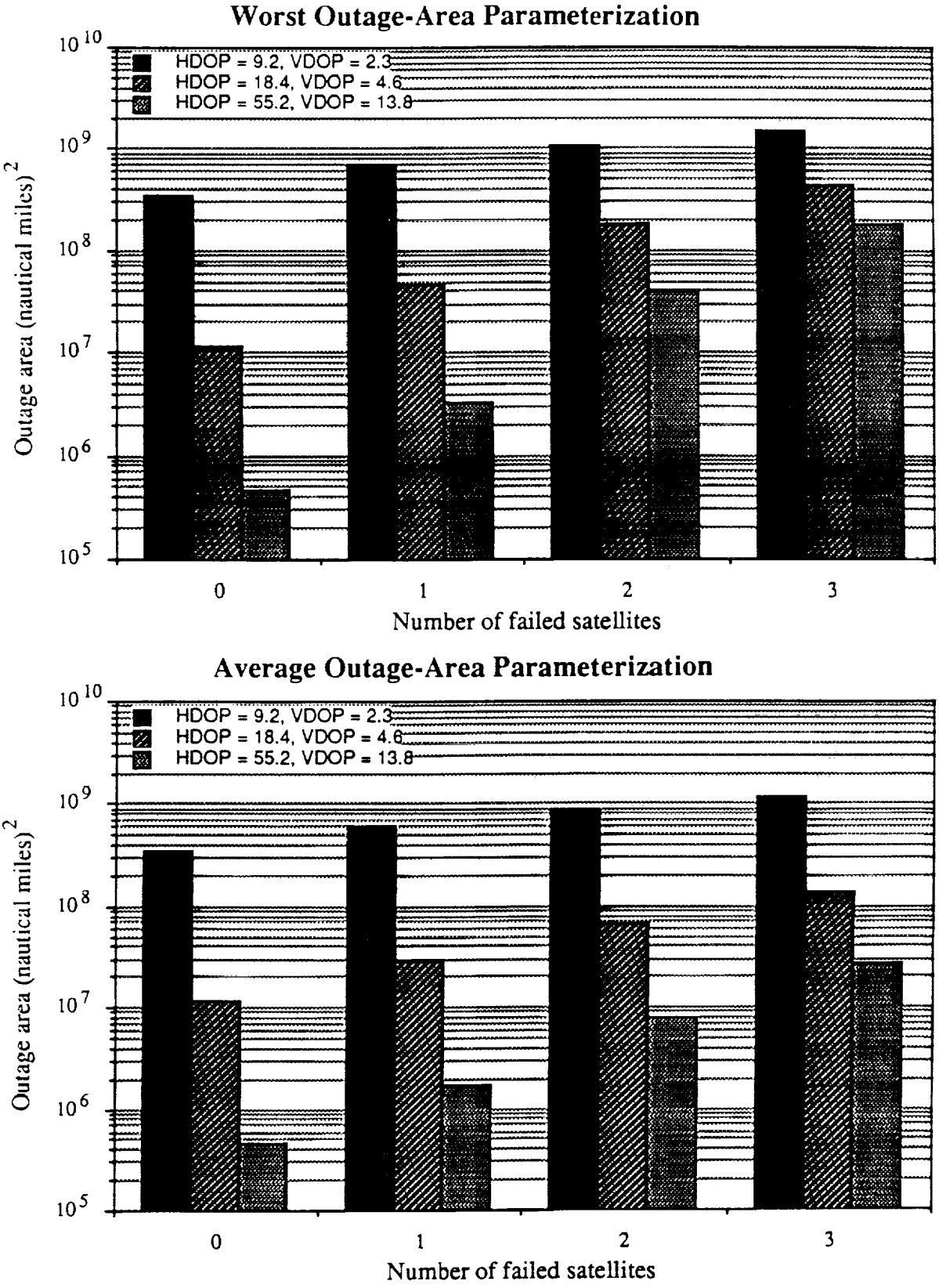


Figure 2.3 Worst-Case and Average Outage Area Parameterization. Shown is the zero-failure outage area along with the worst-case and average single-, double-, and triple-satellite failures as a function of DOP.

The histograms for the three-satellite failure case are shown in Figure 2.4. These demonstrate the distribution of failure combinations and the resulting outage areas. In each of the histograms it can be seen that the number of failure combinations close to the average is much greater than those close to the maximum. Thus, while averaging is not justified as a stand-alone means of system parametrization, it may be sufficient for a first-order approximation of system performance.

2.4 Case Study III: Outage Dynamics

A very important feature of the current model is the ability to generate movies, or slides, which give a time-dependent record of the outage areas. Figure 2.5 shows the outages generated during five consecutive five-minute intervals. The frames were generated using the Test-2 (North America) conditions. The outage contours are the result of failing satellites 1, 4, and 5 with a (HDOP, VDOP) requirement of (55.4, 13.8). The slides demonstrate that huge outages can appear and disappear very quickly. As one might expect, the outage contours generally exhibit an easterly movement. However, it also appears that the DOP requirements dictate outages more than the overall movement of the satellites. This is evidenced by the appearance and disappearance of the large outage areas.

The implication of these results is especially important when considering the outage-based alternate-airport issue. For an en route aircraft navigating in the middle of an outage area, the most probable course of action would be to maintain course and wait for the outage to dissipate. For an aircraft navigating in the terminal area, the issue of landing at an alternate airport is rather mute. Unless the primary airport is on the outside border of an outage, the probability of finding a suitable alternate airport is prohibitively low. Again, the most likely course of action would be to wait for the outage to dissipate. The situation is somewhat analogous to an aircraft entering a terminal with a single instrument landing system (ILS). If the ILS were to fail during an approach, a missed approach procedure would be taken, followed by further instructions from the ground station to the pilot. If an approach were being flown using GPS and an outage occurred, a monitoring station equipped with the GPS coverage model could quickly determine the nature and extent of the outage. Using a model capable of predicting the movement of the current outage, the pilot could be advised of the best course of action. While more work is undoubtedly needed in this area, the ability to track outages as a function of time should prove to be of great value in the evaluation of satellite-based navigation systems.

2.5 Case Study IV: Availability

To determine the availability (and subsequent unavailability) for each location in the North American search grid, the Markov state probabilities shown in Table 2.7 were used [5-8].

Table 2.7 Steady-State Markov Probabilities

Number of failed SVs	Markov Probabilities	Cumulative Probability
0	0.703	0.703
1	0.227	0.930
2	0.055	0.985
3	0.012	0.997
4	0.002	0.999
5	0.00042	0.99942
6	0.000071	0.999491

In order to create an outage record of workable size in the current computing environment, system availability was calculated by only considering up to three failed satellites. Thus, the maximum availability would necessarily be limited to 0.997, or an unavailability of 0.003 (1.000 - 0.997). This is equivalent to a minimum system unavailability of 4.32 minutes per day. While only allowing for three satellite failures seems somewhat prohibitive, it will be seen in the next section that the majority of unavailability is accounted for in just considering up to three satellite failures. Shown in Figure 2.6 are the unavailability contours for the North American continent. The plots dramatize the expected unavailability of the GPS 21 Primary Satellite Constellation as a function of location and dilution of precision. It is interesting to note the high degree of location dependency in each of the contours. From the contours it is evident that the stringent DOP requirement of (HDOP = 9.2, and VDOP = 2.3) results in substantial unavailability (2-7 hours) whereas the least stringent requirement (HDOP = 55.4, and VDOP = 13.8) results in significantly lower overall unavailability approaching the analysis-imposed 4.32 minutes-per-day limit.

The ratios of maximum and average unavailability as a function of DOP are summarized in Table 2.8. In comparing Table 2.8 with its counterparts for outage area (Tables 2.5 and 2.6), we see that relaxation of the DOP requirements has a similar nonlinear affect on unavailability. Also, it is again seen that the effect is greater on the average than at the extremes. In the

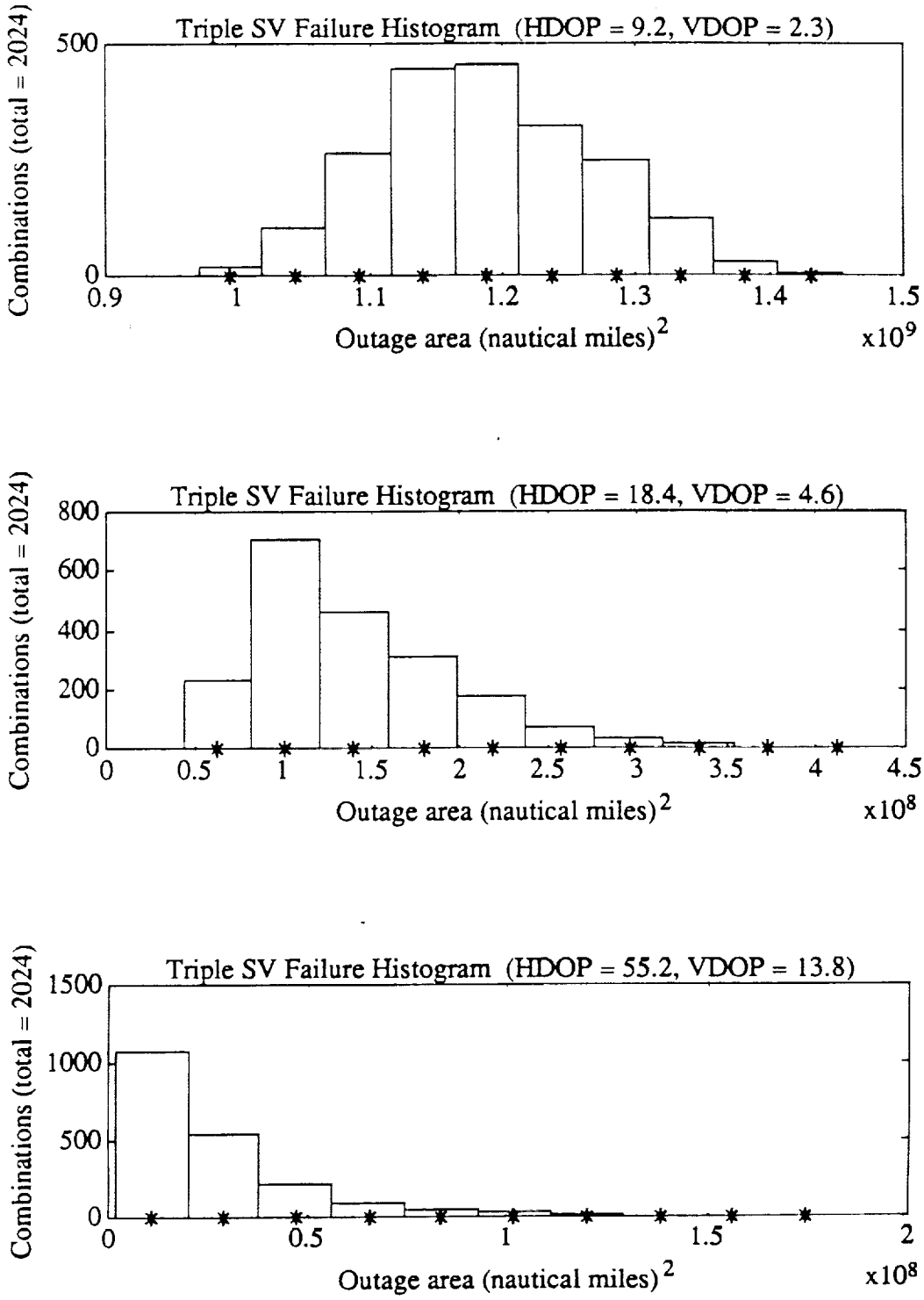


Figure 2.4 Triple Satellite Failure Histograms.
The number of triple satellite failures causing the specified outage area is given as a function of DOP.

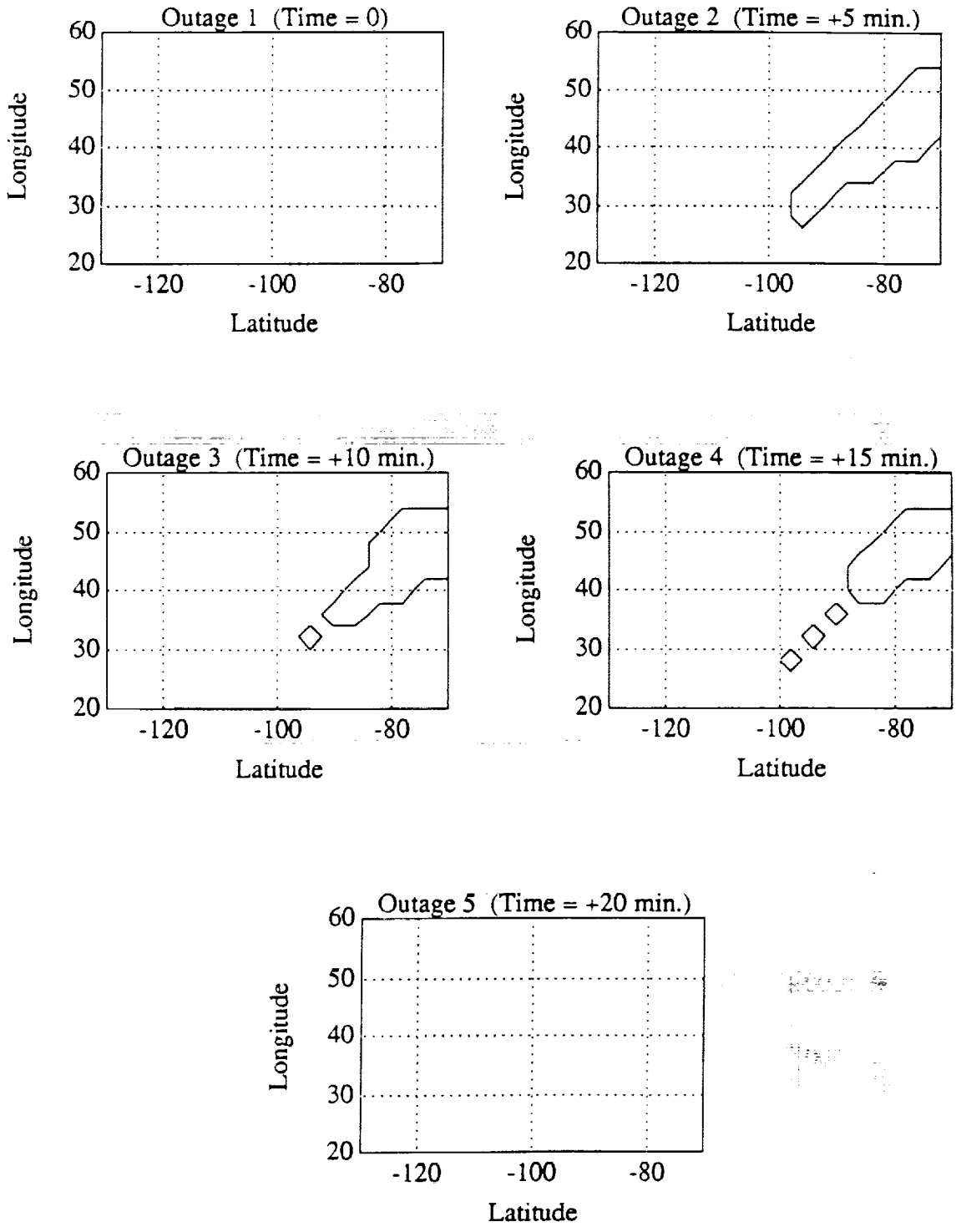


Figure 2.5 Outage dynamics as a function of time. Contours show outage regions as a result of failing satellites 1, 4, and 5 with a dilution of precision requirement of (VDOP = 13.8, HDOP = 55.2).

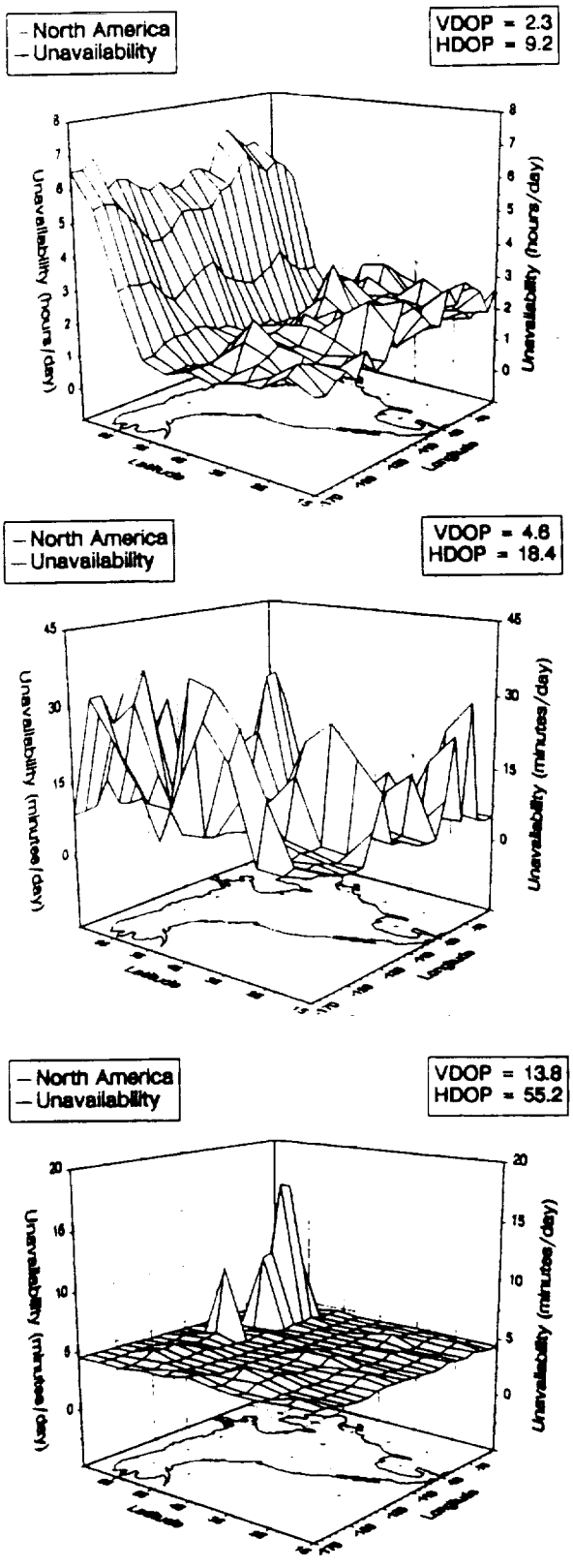


Figure 2.6 3-Dimensional Unavailability Contours. The three contour plots show unavailability as a function of location and dilution of precision (DOP).

next section, the model will be applied to a limited number of important locations in order to determine if they are representative of the entire North American search region.

Table 2.8 VDOP Relaxation: Worst and Avg. Unavail.

Number of failed SVs	Relax VDOP (2.3 : 4.6) (1 : 2)	Relax VDOP (2.3 : 13.8) (1 : 6)
Worst-Case Unavailability	10.92	24.92
Average Unavailability	14.65	33.11

2.6 Case Study V: Specific Airports

In an attempt to develop a benchmark for determining the number of satellite failures needed to best quantify the current system, the model was used to calculate the maximum and average unavailability experienced by the ten busiest airports in the United States (as reported by the Aircraft Owners and Pilots Association). Table 2.9 lists the name and location of these airports.

Table 2.9 The AOPA Ten Busiest U.S. Airports

Airport Name	Airport Location
1. Chicago O'Hare Int'l	41°58' N 87°54' W
2. Atlanta Int'l	33°38' N 84°35' W
3. Dallas/Ft. Worth Int'l	32°53' N 97°02' W
4. Los Angeles Int'l	33°56' N 118°24' W
5. Santa Ana	33°40' N 117°52' W
6. Van Nuys	34°12' N 118°29' W
7. Phoenix Sky Hrbr. Int'l	33°26' N 112°00' W
8. Long Beach	33°49' N 118°09' W
9. Denver Stapleton Int'l	39°46' N 104°52' W
10. Miami Int'l	25°47' N 80°17' W

Source: AOPA 1992 Fact Card (1990 Data), Aircraft Owners and Pilots Association, Frederick, Maryland.

For these locations, the model was run for up to and including six satellite failures. The maximum and average unavailabilities for failing 3, 4, 5, and 6 satellites for the ten busiest general aviation airports are shown in Table 2.10. These values are based on the Category I requirement of (HDOP = 9.2, VDOP = 2.3) assuming a UERE of 6 ft. (95%).

From Table 2.10 it can be seen that, on the average, the difference in the computed unavailability for running the model to three failures is only about two percent higher than running the model to its current limit of six failures. This appears to indicate that, in order to obtain a good approximation of system availability, the model need only consider up to three satellite failures. Naturally, as different constellations are evaluated, this assumption may no longer be valid and will require further investigation. It is worth noting that, while none of these airports represent the maximum unavailability discussed previously, they are representative of the average of the entire search grid to within 22%, or approximately 30 minutes of unavailability. While this seems to be a rough approximation, it may be justified in instances where a quick check is sufficient.

Table 2.10 Unavailability Calculations

Max. number of failed SVs	Maximum Unavailability (hours/day)	Average Unavailability (hours/day)
3	2.90	2.17
4	2.87	2.14
5	2.86	2.14
6	2.86	2.13

3. SUMMARY

The characterization of outages in the coverage provided by the Global Positioning System is of utmost importance when considering GPS as a sole means of navigation and as a navigation aid for flying a precision approach. The continued development of the Ohio University GPS coverage model will enable detailed parametric studies of various coverage issues. The application of the model was demonstrated through several representative case studies. These studies showed how various Dilution of Precision (DOP) requirements affect system performance.

Future work will include designing and running additional simulations to determine key system parameters. The model presented in this report will serve as a foundation for the development of a complete coverage model with the capacity of evaluating (and thus designing) a robust satellite-based navigation system.

ACKNOWLEDGEMENTS

This work was supported by the FAA under contract DTRS-57-87-c-00006, and the FAA and NASA through the Joint University Program (Grant NGR 36-009-017). The author wishes to express his thanks to Dr. Frank van Graas for his guidance and to Mohamed Shaltot for his assistance.

REFERENCES

[1] "Engineering and Operational Issues Associated with the Application of Satellite-Based Navigation to Precision Approach and Landing-Final Report," FAA's John A. Volpe National Transportation Systems Center, Cambridge, MA, December, 1991. Contracts DTRS-57-89-D00043 and DTRS-57-87-C-00006.

[2] van Graas, Frank. "Final Report, OU/AEC 17-89TM 00006/22, GPS/LORAN Interoperability Phase II," Avionics Engineering Center, Ohio University, April 1989.

[3] Green, G.B., Massatt, P.D., Rhodus, N.W. "The GPS 21 Primary Satellite Constellation," *Navigation*, Vol. 34, No. 4, Winter, 1987-88, pp. 9-24.

[4] Jorgensen, Paul S. "Achieving GPS Integrity and Eliminating Areas of Degraded Performance," *Navigation*, Vol. 34, No. 4, Winter, 1987-88, pp.297-306.

[5] Creamer, P.M., Geyer, E.M., Pisano, J.J. "GPS Availability and Reliability for Aircraft Precision Approach," ION National Technical Meeting, San Diego, CA, January, 1992.

[6] Durand, J-M and Caseau, A. "GPS Availability, Part I: Availability of Service Achievable for Different Categories of Civil User," *Navigation*, Vol. 37, No. 2, Summer, 1990., pp. 123-139.

[7] Durand, J-M and Caseau, A. "GPS Availability, Part II: Evaluation of State Probabilities for 21 Satellite and 24 Satellite Constellations," *Navigation*, Vol. 37, No. 3, Fall, 1990, pp. 285-296.

[8] Kline, Paul A. "Fault Detection And Isolation For Integrated Navigation Systems Using The Global Positioning System," Masters Thesis, Ohio University, November, 1991.

ION 1993 National Technical Meeting

GROUND STATION SITING CONSIDERATIONS FOR DGPS

James D. Waid

BIOGRAPHY

James Waid is a student at Ohio University currently completing his M.S.E.E. under Dr. Frank van Graas. He has been working on developing a GPS landing system funded under a grant from NASA Langley Research Center. He received a B.S.E.E. degree form Ohio University in 1990.

ABSTRACT

Aircraft guidance and positioning in the final approach and landing phases of flight requires a high degree of accuracy. The Global Positioning System operating in differential mode (DGPS) is being considered for this application. Prior to implementation, all sources of error must be considered. Multipath has been shown to be the dominant source of error for DGPS and theoretical studies have verified that multipath is particularly severe within the final approach and landing regions. Because of aircraft dynamics, the ground station segment of DGPS is the part of the system where multipath can most effectively be reduced. Ground station siting will be a key element in reducing multipath errors for a DGPS system. This situation can also be improved by using P-code or narrow correlator C/A-code receivers along with a multipath rejecting antenna. This paper presents a study of GPS multipath errors for a stationary DGPS ground station. A discussion of GPS multipath error characteristics will be presented along with some actual multipath data. The data was collected for different ground station siting configurations using P-code, standard C/A-code and narrow correlator C/A-code receiver architectures and two separate antenna constructions.

INTRODUCTION

GPS soon will have the capability to provide position information to users anywhere in the world nearly 24-hours per day. For applications requiring precise positioning (better than 100 meters (95%)), a stand alone installation is not sufficient to provide adequate positioning accuracy for civilian users. However, differential GPS (DGPS) can provide users with sub-meter level accuracies. Aircraft guidance and positioning in the final approach and landing phases of flight is a prime example of an application for DGPS.

At Ohio University's Avionics Engineering Center, the use of DGPS for guidance and positioning of aircraft during final approach and landing is being investigated. GPS by itself has many sources of error including Selective Availability (SA), ionospheric delay, tropospheric delay, receiver hardware errors, receiver noise and multipath. DGPS eliminates those errors which are common to both receivers. The single largest source of error that remains is the error due to multipath [1]. If DGPS is to be used for final approach and landing, the effects that multipath has on the GPS range measurements must be characterized and controlled to meet the required error budgets. This paper will present a discussion of different characteristics and multipath errors observed for various antenna and receiver configurations. The siting configurations include: ground level and ground plane mounted hangar rooftop antenna placements using a standard microstrip GPS antenna and an experimental helix antenna. The above antenna placements will be combined with separate receiver architectures that include: P-code, standard C/A-code and narrow correlator C/A-code receivers.

Presented at the Institute of Navigation 1993 National Technical Meeting, San Francisco, CA, January 20-22, 1993

BACKGROUND

The accuracy of GPS positioning depends on the accuracy of the pseudorange measurements. There are many error sources which cause erroneous range measurements. The major error sources are as follows:

- signal delay due to propagation through the troposphere
- signal delay due to propagation through the ionosphere
- error due to satellite clock offset and orbit uncertainty
- Selective Availability (SA)
- receiver inter-channel biases
- receiver measurement errors
- dynamics
- thermal noise
- specular multipath
- diffuse multipath

Although integrated carrier phase measurement accuracies are typically on the order of two centimeters, the code phase measurements are still required for ambiguity resolution. Therefore, this paper focuses on the code phase measurement error. The signal at the antenna is a combination of different types of signals: direct and non-direct. The direct signal is the signal received that travels the geometric distance from the satellite to the receiver. The non-direct or multipath signal is a signal that has been reflected or diffracted off an object and arrives at the receiver after the direct signal. In general, multipath signals are weaker than the direct signals. When the direct and the multipath signals combine, the result is a signal with the same frequency but having a relative phase difference with respect to the original direct signal. This phase error affects both the code measurement and the carrier phase measurement.

DGPS eliminates the errors in the measurements that are common to both receivers, but multipath has a different effect on each receiver. This is because multipath depends on the GPS antenna environment. For a typical DGPS system, the receivers are not close enough to each other to possess the same multipath characteristics. Three categories of multipath for the final approach and landing environment are [2]:

- Obstacle-based at the airborne receiver.
- Airframe-based at the airborne receiver.
- Obstacle-based at the ground reference station receiver.

The air and ground system obstacle-based multipath originates from the ground itself as well as from buildings or other structures on or near the ground.

The obstacle-based multipath at the ground reference station often arrives at the antenna from a direction below the horizon. An effective method for eliminating this multipath is to limit the antenna's gain pattern so that the antenna is only capable of receiving signals from above the horizon. This can be achieved in two ways: placing the antenna on a large ground plane or electrically adjusting the antenna gain pattern to attenuate any signals below the horizon. Both of these methods will be discussed later in the results segment of the paper.

DATA COLLECTION

GPS multipath data collection was performed at the Ohio University Airport (UNI) located near Albany, Ohio. The area surrounding UNI is flat and free of clutter. There are also two large fixed structures (hangars) that are capable of generating significant multipath. Data was collected at two sites: site one was located on top of the larger of the two aircraft hangars, site two was located in a field approximately 500 meters away from the hangars and the antenna was placed at ground level. Site one represents a typical DGPS reference station siting with the hangars being the leading multipath contributor. Site two can be considered a benign multipath environment because the antenna is being placed on a large ground plane and the leading multipath contributor is the ground itself because there are no fixed obstacles above the horizon that are generating multipath signals.

Two GPS antennas were used during the data collection, a dual-frequency microstrip antenna and an experimental helix antenna. The experimental helix antenna was provided by Mr. Don Spitzmesser of the Jet Propulsion Laboratories. The antenna consists of a 20 cm parabolic reflector and a thin wire helix placed in the center of the reflector dish. The helix is configured to receive both L1 and L2 frequencies. Because of the parabolic dish, the helix antenna is more directive and better masks signals that may arrive from below the horizon. There were two GPS receivers used for the data collection: an Ashtech P-12 GPS receiver and a Novatel GPS CARD receiver. The P-12 is capable of continuous tracking of L1 C/A-code and both L1 and L2 P-code. The Novatel GPS CARD is an L1 frequency, narrow correlator C/A-code receiver.

The measurement data for the P-12 and the GPS CARD was collected and recorded in real time using a 386 notebook computer and a 286 desktop computer respectively. Data was collected over a 120 minute time period. Five sets of data were collected for this analysis:

Hangar Roof:
- P-12 with microstrip antenna
- GPS Card with microstrip antenna
- P-12 with Helix antenna
- GPS Card with Helix antenna

Field Location:
- P-12 with microstrip antenna on the ground

DATA PROCESSING TECHNIQUES

The combination of multipath, thermal noise, unknown bias and receiver error was extracted from the data using the standard code-minus-integrated Doppler technique [3,4]. Equation 1 shows the result:

$$d_{code} - d_{phase} = 2d_{iono} + d_{code-meas} - d_{phase-meas} + d_{code-noise} - d_{phase-noise} + d_{code-mp} - d_{phase-mp} - \Delta + d_{other} \quad (1)$$

where:
- d_{code} is the code measurement
- d_{phase} is the carrier-phase (integrated doppler) measurement
- d_{iono} is the signal delay due to propagation through the ionosphere
- $d_{code-noise}$ is a combination of thermal noise and diffuse multipath on the pseudorange
- $d_{phase-noise}$ is a combination of thermal noise and diffuse multipath on integrated carrier phase
- $d_{code-meas}$ & $d_{phase-meas}$ is receiver measurement noise for code and phase measurements
- $d_{code-mp}$ & $d_{phase-mp}$ is specular multipath on the code and phase
- Δ is an integer wavelength ambiguity
- d_{other} includes receiver measurement error

For situations where the strength of the multipath is less than the direct signal, the carrier-phase multipath term (d_{phase}) will not exceed 4.8 centimeters [5]. It has been shown that state-of-the-art receivers exhibit phase-noise ($d_{phase-noise}$) values on the order of 0.1 millimeter (1-sigma) [6] allowing this term to be neglected as well. The receiver phase measurement errors ($d_{phase-meas}$) are also negligible [7]. When compared to the code-multipath error ($d_{code-mp}$), which is usually on the order of meter, the carrier-phase multipath ($d_{phase-mp}$) and the noise ($d_{phase-noise}$) terms are very small. For this reason they can be dropped from equation (1). The integer ambiguity (Δ) is a constant bias for the duration of the data collection, which is not of interest for this study. Equation (1) is then approximated by:

$$(d_{code} - d_{phase})' = 2d_{iono} + d_{code-meas} + d_{code-noise} + d_{code-mp} + d_{other} \quad (2)$$

The error due to the propagation delay through the ionosphere can be removed through the standard dual-frequency correction [8]:

$$d_{iono_{f1}} = \left(\frac{f_2^2}{f_2^2 - f_1^2}\right)\left(d_{code_{f1}} - d_{code_{f2}}\right) \quad (3)$$

Noise in the data is reduced by averaging (filtering) the code measurements against the stable carrier measurements. This is accomplished using a complementary Kalman filter [9]. After applying the ionospheric correction and the complementary Kalman filter, we arrive at the following:

$$(d_{code} - d_{phase})'' = d_{code-meas} + d_{code-mp} + d_{other} \quad (4)$$

The next section presents the results of the data collection and data analysis.

DISCUSSION OF RESULTS

The results are presented in the following figures and table. The filtered code-minus-carrier for satellites 3, 17, and 23 is shown in figures 2 through 25 for all the receiver and antenna configurations being considered. The three satellites were selected because they include the elevation angles of interest: SV17 exhibits the characteristics of a high elevation satellite, SV23 represents a medium elevation satellite and SV3 is indicative of a lower elevation satellite that vanishes below the horizon during the data collection. Figure 1 shows the elevation angles for the satellites during the data collection. As anticipated, the error levels are correlated to the lower elevation angles for all the test cases. Table I shows the root mean squared (rms) of the multipath error in meters for C/A-code, narrow correlator C/A-code and P-code for each satellite for data collected on the hangar roof and C/A-code and P-code for data collected at the site away from the aircraft hangars. The last row in the table represents the average for the three satellites for the receiver and antenna configuration listed in that column.

The best case for all the scenarios run was the P-code receiver operating out in the field away from all structures. The worst case was observed on the hangar roof using the standard C/A-code with the microstrip antenna. The contrast between the two results indicates that the multipath does indeed enter the antenna from below the horizon. These results are as expected. From the data presented it is easy to see that the lowest levels of multipath were experienced for high elevation satellites using the P-code measurements. This result is also expected.

In general, the measurement taken away from the hangar showed lower rms levels of multipath for all satellites. This kind of multipath environment may not be available for a typical DGPS reference station location. The hangar roof can be considered a more typical example of a DGPS reference station site. For this site the helix antenna produced results that were significantly better than the microstrip antenna.
The helix antenna has the limitation of only being able to track satellites down to an elevation angle of 10°. Another consideration for a DGPS landing system, P-code may not be available for all aircraft. In the case that P-code is not available, obviously C/A-code would have to be used. Looking at the comparison between C/A-code and narrow correlator C/A-code, the narrow correlator C/A-code exhibits multipath with less noise and having smaller magnitude than the standard C/A-code measurements.

Also it should be noted that the C/A-code errors measured in the field are mostly caused by high-frequency measurement noise, rather than by multipath. Integration over time of high-frequency noise gives rise to a random-walk error. It was found that the errors measured in the field exhibit insignificant correlations from one day to the next.

Although the helix antenna performed very well in a multipath environment, its gain at lower elevation angles is much less than that of the microstrip antenna. Another concern is the stability of the phase center of the helix antenna for carrier-phase tracking applications. For code-phase DGPS, however, this is not a significant problem.

Table I

	Field		Hangar Roof					
	Microstrip		Microstrip			Helix		
	C/A rms (meters)	P rms (meters)	C/A rms (meters)	N.C.C/A rms (meters)	P rms (meters)	C/A rms (meters)	N.C.C/A rms (meters)	P rms (meters)
SV3	0.4757	0.0802	1.2658	0.4516	0.3329	0.9232	0.2031	0.0996
SV17	0.4624	0.0456	0.8015	0.3115	0.3408	0.3504	0.0685	0.0417
SV23	0.4289	0.0397	0.6418	0.3463	0.2550	0.4438	0.1809	0.0445
average	0.4557	0.0552	0.9030	0.3698	0.3096	0.5725	0.1508	0.0619

Recommendations:

1.) Use a site out in the field for minimum multipath. A major draw back to this recommendation is that snow can cover the antenna and the area around the antenna when placed on the ground. This will seriously affect the performance of the antenna.
2.) The next best siting that was considered was the helix antenna placed at a location that provided visibility down to 5° (hangar roof). The same effect can be achieved by placing any antenna on a large ground plane.

For all siting options considered, the use of narrow correlator C/A-code or P-code significantly reduces the multipath error .

CONCLUSIONS

Multipath is the dominate error source for DGPS. A number of extreme siting scenarios were investigated with respect to multipath performance. It was found that a significant level of multipath enters the antenna pattern from below the horizon. Therefore it is recommended to either have a large ground plane or reduce the antenna pattern below the horizon.

ACKNOWLEGDEMENTS

This work was supported by NASA Langley Research Center (Grant NAG1-1423), and the FAA and NASA through the Joint University Program (Grant NGR 36-009-017). The P-12 GPS receivers were provided by Ashtech, Inc. The GPS Card receivers were provided by Novatel, Inc. The author would like to thank Dr. Frank van Graas for his helpful suggestions during the writing of the paper. The author would also like to thank Mr. Don Spitzmesser for providing the experimental helix antenna.

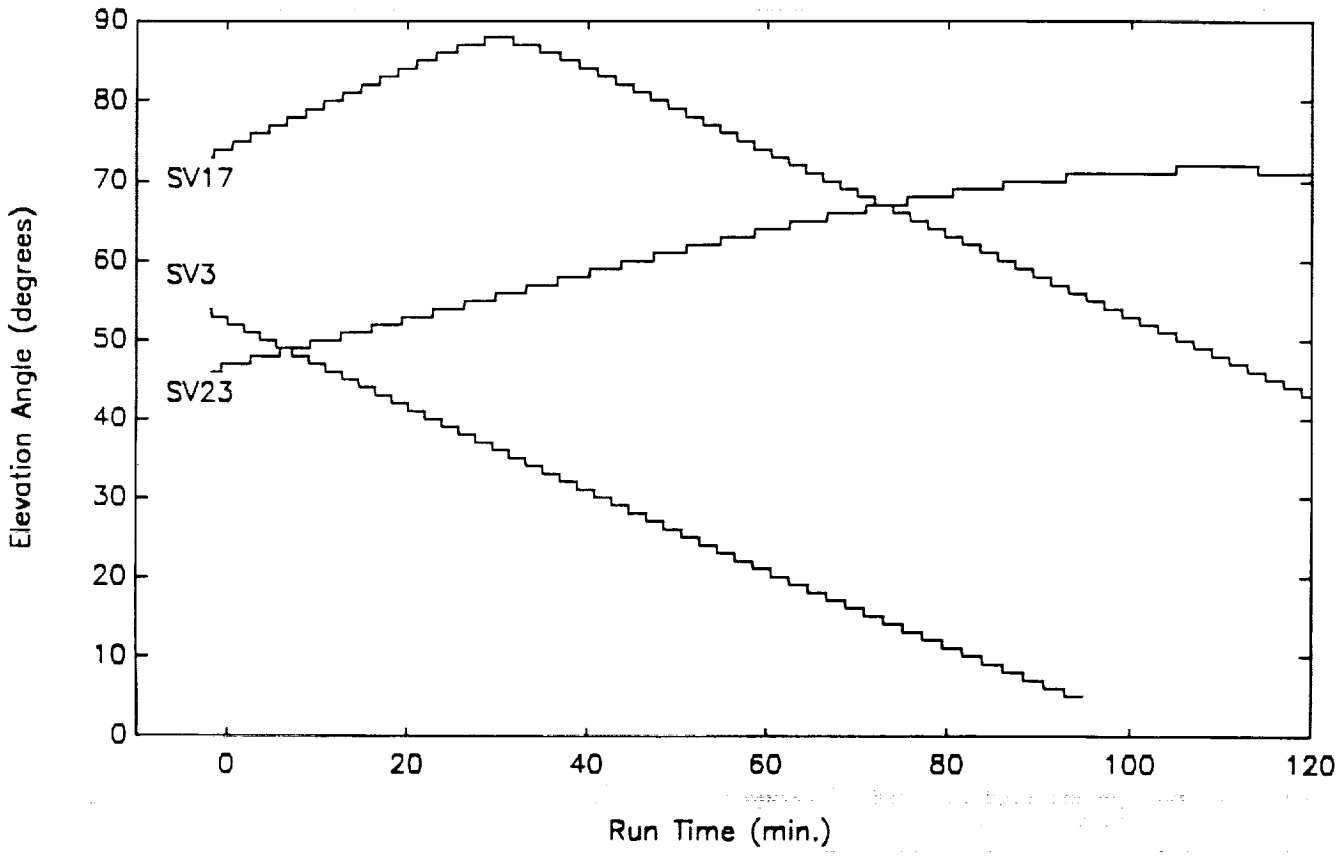


Figure 1. Satellite Elevation Angles

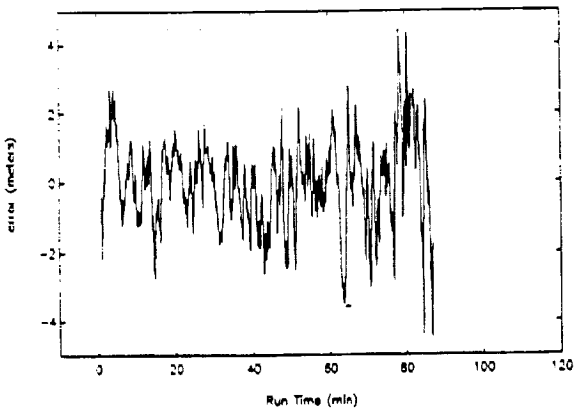


Figure 2. SV3: C/A-code

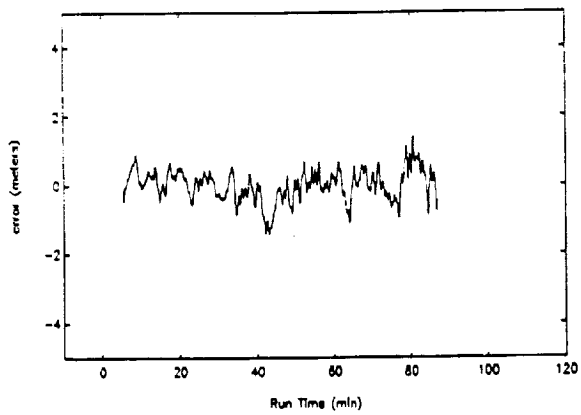


Figure 3. SV3: Narrow Correlator C/A-code

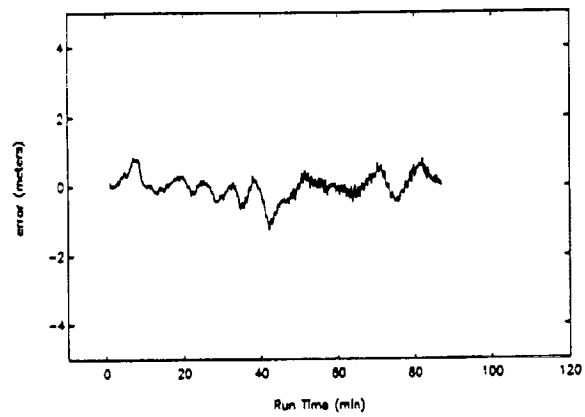


Figure 4. SV3: P-code

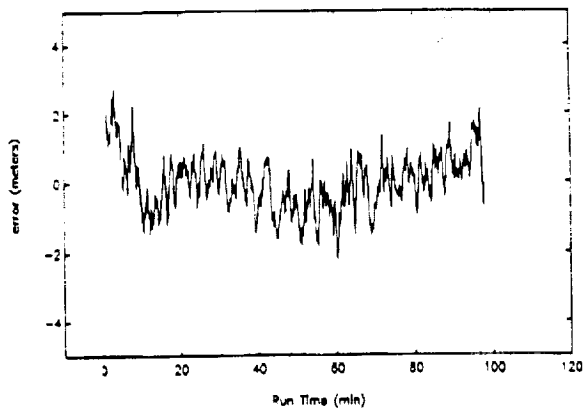


Figure 5. SV17: C/A-code

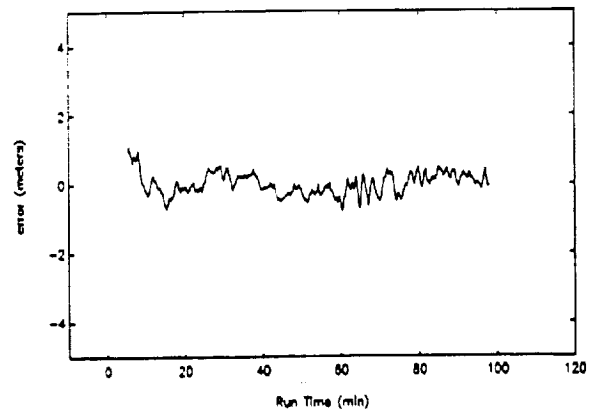


Figure 6. SV17: Narrow Correlator C/A-code

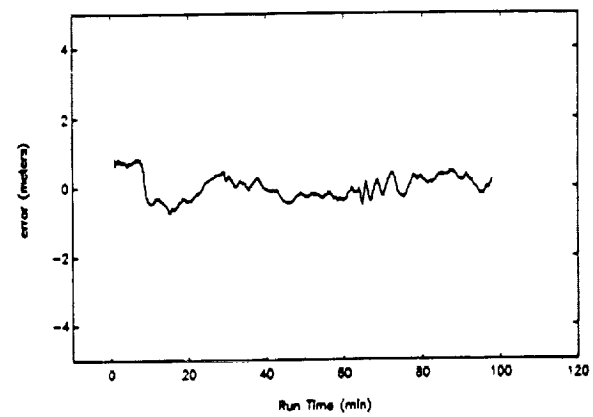


Figure 7. SV17: P-code

Figures 2-7: Hanger Roof with Microstrip Antenna - Multipath, Thermal Noise, Unknown Bias and Receiver Error

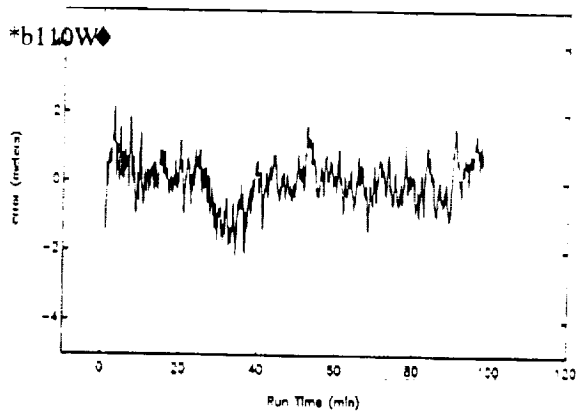


Figure 8. SV23: C/A-code

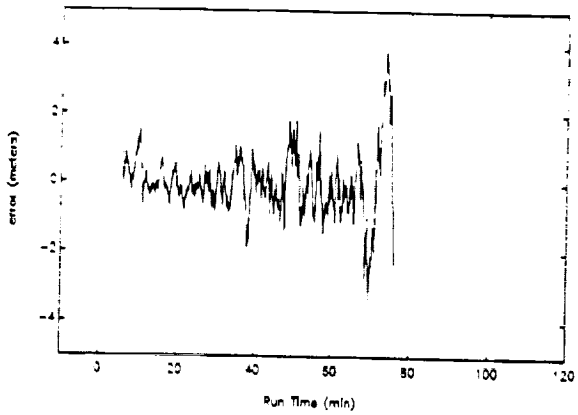


Figure 11. SV3: C/A-code

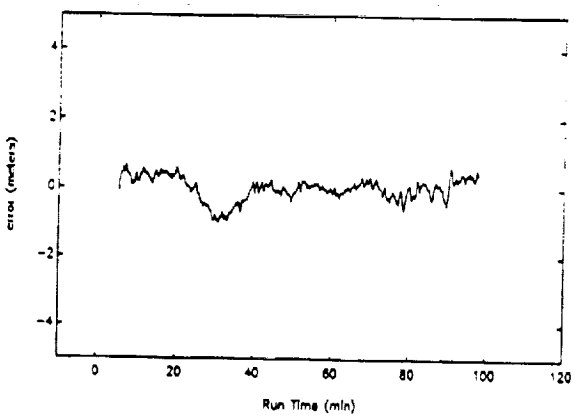


Figure 9. SV23: Narrow Correlator C/A-code

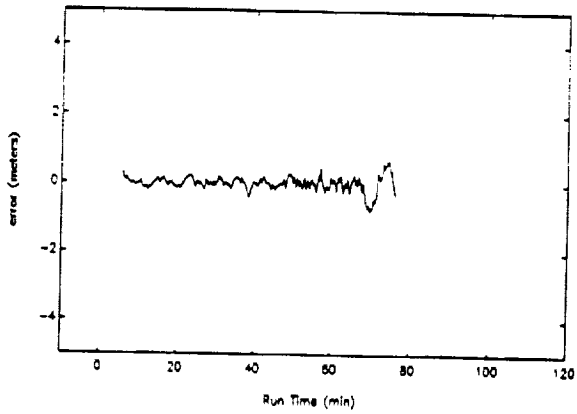


Figure 12. SV3: Narrow Correlator C/A-code

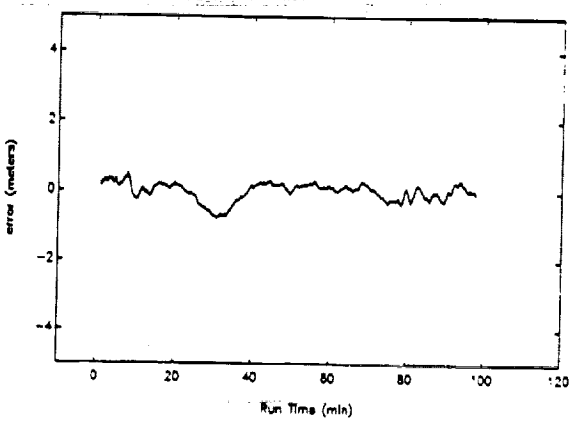


Figure 10. SV23: P-code

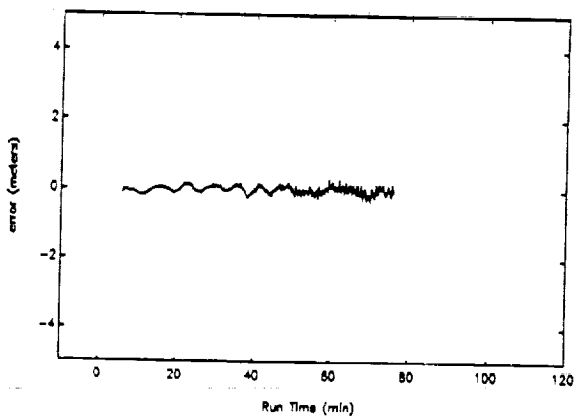


Figure 13. SV3: P-code

Figures 8-10: Hanger Roof with Microstrip Antenna - Multipath, Thermal Noise Unknown Bias and Receiver Error

Figures 11-13: Hanger Roof with Helix Antenna - Multipath, Thermal Noise, Unknown Bias and Receiver Error

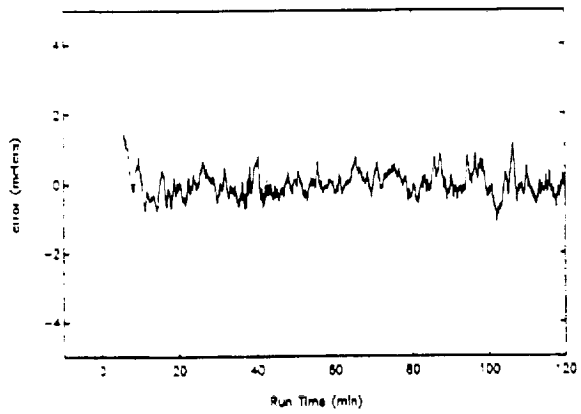


Figure 14. SV17: C/A-code

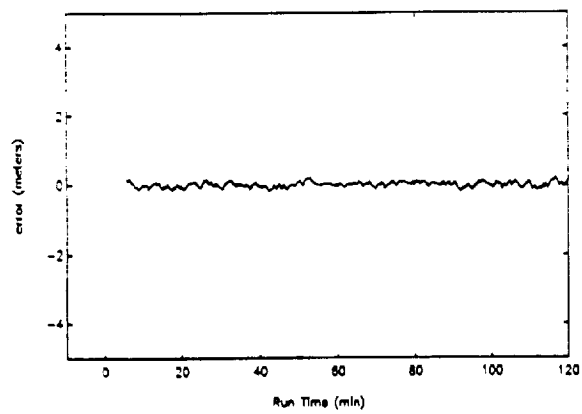


Figure 15. SV17: Narrow Correlator C/A-code

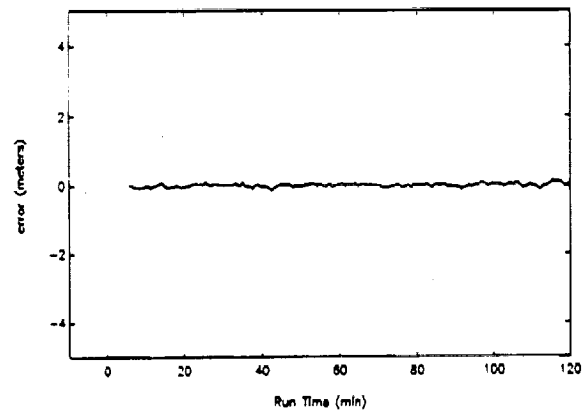


Figure 16. SV17: P-code

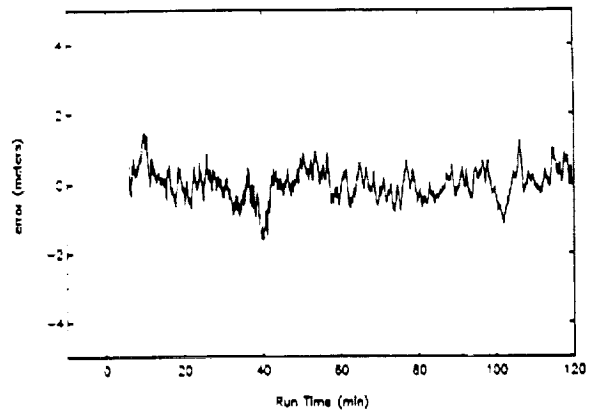


Figure 17. SV23: C/A-code

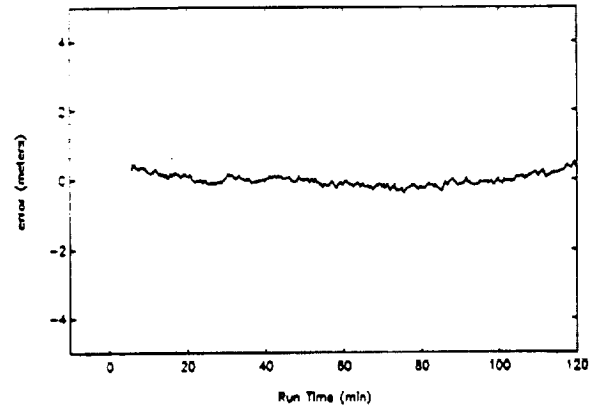


Figure 18. SV23: Narrow Correlator C/A-code

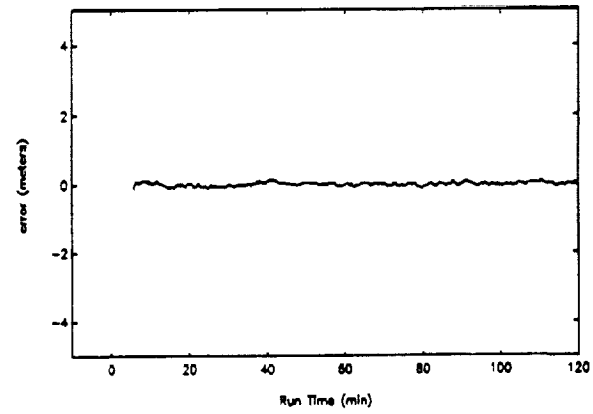


Figure 19. SV23: P-code

Figures 14-19: Hanger Roof with Helix Antenna - Multipath, Thermal Noise, Unknown Bias and Receiver Error

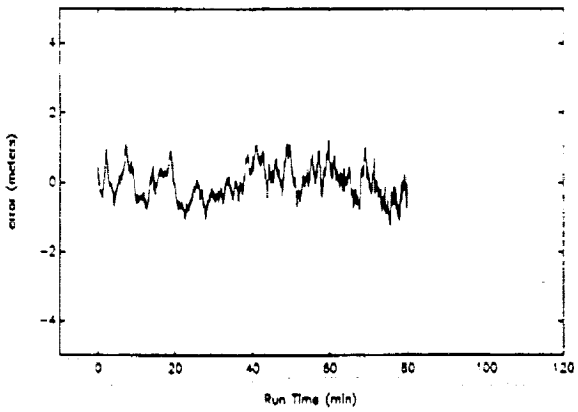


Figure 20. SV3:C/A-code

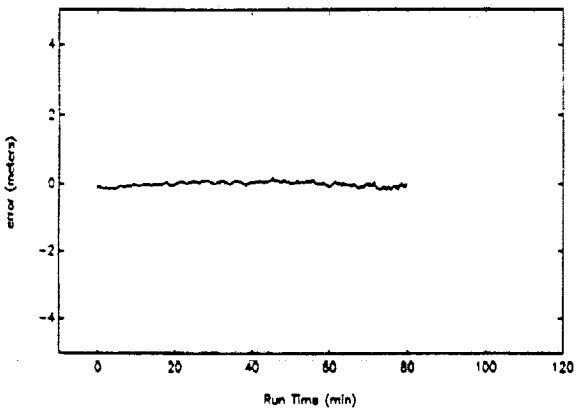


Figure 21. SV3:P-code

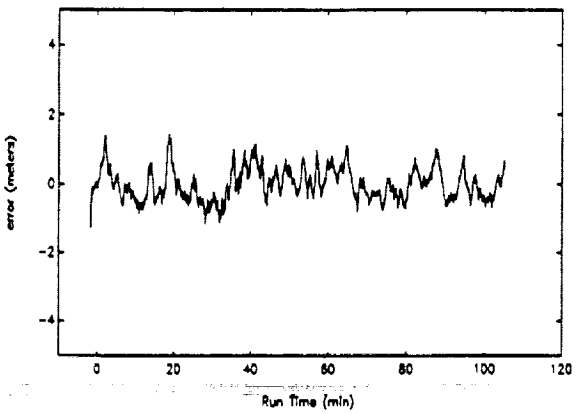


Figure 22. SV17:C/A-code

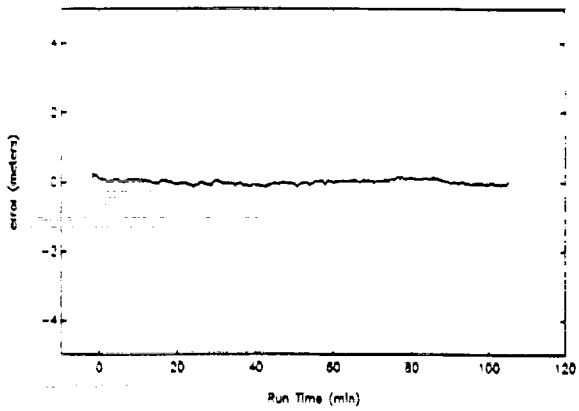


Figure 23. SV17:P-code

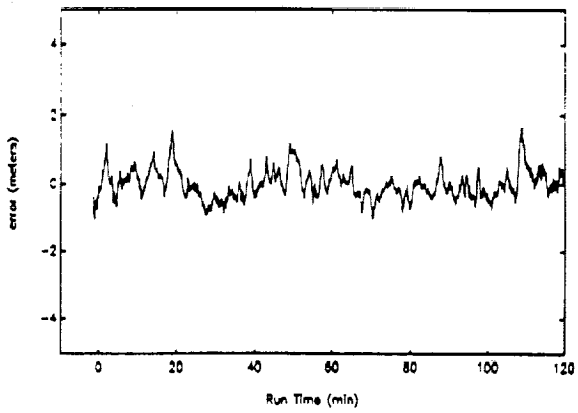


Figure 24. SV23:C/A-code

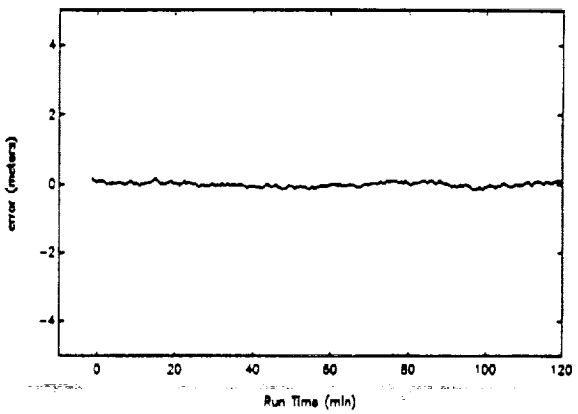


Figure 25. SV23:P-code

Figures 20-25: Field with Microstrip Antenna - Multipath, Thermal Noise, Unknown Bias and Receiver Error

REFERENCES

[1] Lachapelle, G., W. Falkenberg, D. Neufeldt and P. Keilland, "Marine DGPS Using Code and Carrier in a Multipath Environment," ION GPS-89, Colorado Springs, CO, September, 1989.

[2] Braasch, M., "On the Characterization of Multipath Errors In Satellite-Based Precision Approach and Landing Systems," Ph.D. Dissertation, Ohio University, Athens, OH, June, 1992.

[3] Evans, A., "Comparison of GPS Pseudorange and Biased Doppler Range Measurements to Demonstrate Signal Multipath Effects," International Telemetering Conference, Las Vegas, NV, October, 1986.

[4] General Dynamics, Electronics Division, "Final User Field Test Report for the NAVSTAR Global Positioning System Phase I, Major Field Test Objective No. 17: Environmental Effects, Multipath Rejection," Rep. GPS-GD-025-C-US-7008, San Diego, CA, March 28, 1979.

[5] Ibid 2.

[6] Ferguson, K., et al, "Three-Dimensional Attitude Determination with the Ashtech 3DF 24-Channel GPS Measurement System," Proceedings of the Institute of Navigation's National Technical Meeting, Phoenix, AZ, January, 1991.

[7] van Graas, F. and M. Braasch, "GPS Interferometric Attitude and Heading Determination: Initial Flight Test Results," NAVIGATION: Journal of the Institute of Navigation, Vol. 38, No. 4, Winter, 1991.

[8] Spilker, Jr., J., "GPS Signal Structure and Performance Characteristics," NAVIGATION: The Journal of the Institute of Navigation, Vol. 25, No. 2, Summer, 1978.

[9] Ibid 7.

omit

PRINCETON UNIVERSITY

INVESTIGATION OF AIR TRANSPORTATION TECHNOLOGY AT PRINCETON UNIVERSITY, 1992-1993

Professor Robert F. Stengel
Department of Mechanical and Aerospace Engineering
Princeton University
Princeton, New Jersey 08544

SUMMARY OF RESEARCH

The Air Transportation Research Program at Princeton University proceeded along five avenues during the past year:

- Flight Control System Robustness
- Microburst Hazards to Aircraft
- Wind Rotor Hazards to Aircraft
- Intelligent Aircraft/Airspace Systems
- Aerospace Optical Communications

This research has resulted in a number of publications, including theses, archival papers, and conference papers. An annotated bibliography of publications that appeared between June 1992 and June 1993 appears at the end of this report. The research that these papers describe was supported in whole or in part by the Joint University Program, including work that was completed prior to the reporting period.

Control system robustness is defined as the ability to maintain satisfactory stability or performance characteristics in the presence of all conceivable system parameter variations. While assured robustness may be viewed as an alternative to gain adaptation or scheduling to accommodate known parameter variations, more often it is seen as protection against uncertainties in plant specification. Consequently, a statistical description of control system robustness is consistent with what may be known about the structure and parameters of the plant's dynamic model. Rarely will there be a single "most robust" controller, as design tradeoffs must inevitably be considered. For example, stability, settling time, and control usage all may be of concern; controllers that favor one criterion over the other two have dramatically different characteristics.

Our initial research focused on probabilistic analysis of the stability and performance robustness of given controllers, while more recent research has shifted to designing robust controllers [1-6]. We have demonstrated that classical stability (i.e., gain and phase) margins are not good indicators of robustness, particularly when comparing compensators with different structures. Numerical search is shown to produce robust controllers based on proportional-filter linear-quadratic regulators with implicit model-following.

Severe downdrafts and resulting high velocity outflows caused by microbursts present a significant hazard to aircraft on takeoff and final approach. *Microbursts*, which are often associated with thunderstorm activity, also can occur in the vicinity of dissipating convective clouds that produce no rainfall at ground level. Microburst encounter is a rare but extremely dangerous phenomenon that accounts for one or two air carrier accidents and numerous general aviation accidents each year (on average). Conditions are such that an aircraft's performance envelope may be inadequate for safe penetration unless optimal control strategies are applied.

An expert system for wind shear avoidance that extends the FAA Microburst Windshear Guidelines can account for temporal and spatial variations in the evidence that wind shear is present [7, 8]. A Bayesian Belief Network relates information gathered from many sources to determine the probability of encountering a microburst on the intended flight path. Measurements made by a look-ahead sensor (e.g., Doppler radar or lidar) are processed by extended Kalman filters to develop a head-tailwind profile.

Real-time guidance for the case in which wind shear has been encountered is being investigated. Our emphasis has shifted from optimal strategies for abort and recovery [9] to strategies based on nonlinear-inverse-dynamic controllers [10]. The former approach seeks to minimize a path-following cost function that implicitly maximizes the minimum altitude during an aborted approach to landing. The latter approach prescribes a desired rate of climb once an abort has been declared, then generates the necessary control commands by inverting the aircraft's dynamic model.

The dynamics of a twin-jet transport encountering an intense wind "rotor" have been studied [11]. It was found that a physically realizable rotor could roll the aircraft to inverted attitude if left unopposed by lateral control. Similarly, unopposed full rudder deflection could invert the aircraft in its landing configuration. Conventional linear-quadratic flight control laws can maintain the wing's level through such encounters.

Advanced concepts for air traffic management are being developed by modeling aircraft and air traffic centers as intelligent agents that engage in principled negotiation [12]. Each agent is characterized as a dynamic system that carries out declarative, procedural, and reflexive functions [13]. Principled negotiation entails the proposal of alternative flight plans, evaluation of costs and constraints according to separate and shared interests, and conflict resolution. We are setting the groundwork for an *Intelligent Aircraft/Airspace System (IAAS)*. The goal is to identify means by which ground-based and airborne flight management systems can cooperate to produce a net gain in the efficiency and robustness of air transportation.

With growing demands on the radio spectrum, it is likely that currently unused alternatives could play important roles in the *IAAS*. *Optical sensing and communication* could shoulder a significant percentage of the overall load. Of course, there are weather and line-of-sight limitations on optical devices, so they may never be considered the sole means of providing services. From a global or national perspective, however, optical devices may prove useful in off-loading radio frequencies on a regional and/or altitude-dependent basis. The national airspace is rarely (if ever) "socked in" coast-to-coast, and even in areas of cloud cover, there are altitude strata (especially at cruising heights) within which visual line-of-sight is retained over long distances. By definition, very-low-altitude line-of-sight exists in the terminal area for Category I IFR conditions or better. Consequently, there are numerous situations in which precision measurements and high-bandwidth communication could be furnished by optical systems, including transmissions through communications satellites. (Optical systems can be substantially *more precise* and allow much *higher data rates* than radio transmissions.) The *IAAS* would adapt to changing weather and traffic conditions, always maximizing allowable flight operations subject to practical constraints; consequently, on average, airspace capacity could be greatly increased. An annotated bibliography is contained in [14].

In addition to the research noted above, two publications related to the Joint University Program appeared during the reporting period. A book chapter describing an architecture for intelligent flight control was published [15]. Notes and homework assignments for an undergraduate course on aerospace guidance and control were included in a book describing educational applications of the MATLAB programming language [16].

ANNOTATED BIBLIOGRAPHY OF 1992-1993 PUBLICATIONS

1. R. F. Stengel and C. I. Marrison, "Robustness of Solutions to a Benchmark Control Problem," *J. Guidance, Control, and Dynamics*, Vol. 15, No. 5, Sept.-Oct. 1992, pp. 1060-1067.

The robustness of ten solutions to a benchmark control design problem presented at the *1990 American Control Conference* has been evaluated. The ten controllers have second- to eighth-order transfer functions and have been designed using several different methods, including H_∞ optimization, loop transfer recovery, imaginary-axis shifting, constrained optimization, structured covariance, game theory, and the internal model principle. Stochastic Robustness Analysis quantifies the controllers' stability and performance robustness with structured uncertainties in up to six system parameters. The analysis provides insights about system response that are not readily derived from other robustness criteria, and it provides a common ground for judging controllers produced by alternative methods. One important conclusion is that gain and phase margins are not reliable indicators of the probability of instability. Furthermore, parameter variations actually may improve the likelihood of achieving selected performance metrics, as demonstrated by results for the probability of settling-time exceedance.

2. L. R. Ray and R. F. Stengel, "Stochastic Measures of Performance Robustness in Aircraft Control Systems," *J. Guidance, Control, and Dynamics*, Vol. 15, No. 6, Nov.-Dec. 1992, pp. 1381-1387.

Stochastic robustness, a simple technique used to estimate the robustness of linear, time-invariant systems, is applied to a twin-jet transport aircraft control system. Concepts behind stochastic *stability* robustness are extended to stochastic *performance* robustness. Stochastic performance robustness measures based on classical design specifications and measures specific to aircraft handling qualities are introduced. Confidence intervals for comparing two control system designs are presented. Stochastic performance robustness, the use of confidence intervals, and tradeoffs between performance objectives are applied to a twin-jet aircraft example.

3. L. R. Ray and R. F. Stengel, "A Monte Carlo Approach to the Analysis of Control System Robustness," *Automatica*, Vol. 29, No. 1, Jan. 1993, pp. 229-236.

Stochastic robustness, a simple technique used to estimate the stability and performance robustness of linear, time-invariant systems, is described. The scalar *probability of instability* is introduced as a measure of stability

C-2

robustness. Examples are given of stochastic performance robustness measures based on classical time-domain specifications. The relationship between stochastic robustness measures and control system design parameters is discussed. The technique is demonstrated by analyzing an LQG/LTR system designed for a flexible robot arm. It is concluded that the analysis of stochastic robustness offers a good alternative to existing robustness metrics. It has direct bearing on engineering objectives, and it is appropriate for evaluating robust control system synthesis methods currently practiced.

4. L. R. Ray and R. F. Stengel, "Computer-Aided Analysis of Linear Control System Robustness," *Mechatronics*, Vol. 3., No. 1, Jan. 1993, pp. 119-124.

Stochastic robustness is a simple technique used to estimate the stability and performance robustness of linear, time-invariant systems. The use of high-speed graphics workstations and control system design software in stochastic robustness analysis is discussed and demonstrated.

5. C. I. Marrison and R. F. Stengel, "Gain and Phase Margins as Indicators of Robustness," *Proceedings of the 1992 IEEE Regional Control Conference*, New York, July 1992, pp. 5-8.

A Monte Carlo analysis of scalar compensators designed for a benchmark problem shows that there is very little correlation between classical stability margins and the likelihood that plant parameter variations will lead to instability. The principal reason is that parameter variations change the shape of the Nyquist plot as well as the gain and phase margins; hence, the branch of the nominal Nyquist plot or critical frequency that determines stability margins may not be the one that produces instability as parameters vary. This result also calls into question the use of singular values as measures of stability robustness, because transfer-function amplitude ratio is equivalent to the singular value in the scalar case.

6. R. F. Stengel, L. R. Ray, and C. I. Marrison, "Probabilistic Evaluation of Control System Robustness," presented at the *IMA Workshop on Control Theory and Its Applications*, Minneapolis, Oct 1992.

Practical control systems must operate satisfactorily with uncertain variations in plant parameters (i.e., control systems must be *robust*), but there are limits to the degree of robustness that may be considered desirable. Tolerance to parameter variations that never occur is not useful, and it could lead to closed-loop systems whose normal performance has been compro-

mised unnecessarily. A probabilistic definition of robustness based on expected parameter variations is consistent with accepted design principles, and it is readily evaluated by simulation. *Stochastic Robustness Analysis* predicts the effects of likely parameter variations on closed-loop stability and performance through evaluation of commonly accepted criteria. Competing control designs are judged by the likelihood that system response and design metrics will fall within desired bounds. Together with numerical search, probabilistic evaluation is a powerful approach not only for comparing alternative controllers but for designing control systems that satisfy robustness and performance requirements.

7. D. A. Stratton and R. F. Stengel, "Stochastic Prediction Techniques for Wind Shear Hazard Assessment," *J. Guidance, Control, and Dynamics*, Vol. 15, No. 5, Sept.-Oct 1992, pp. 1224-1229.

The threat of low-altitude wind shear has prompted development of aircraft-based sensors that measure winds directly on the aircraft's intended flight path. Measurements from these devices are subject to turbulence inputs and measurement error, as well as to the underlying wind profile. In this paper, stochastic estimators are developed to process on-board doppler sensor measurements, producing optimal estimates of the winds along the path. A stochastic prediction technique is described to predict the hazard to the aircraft from the estimates as well as the level of uncertainty of the hazard prediction. The stochastic prediction technique is demonstrated in a simulated microburst wind shear environment. Use of the technique in a decision-making process is discussed.

8. D. A. Stratton, "Aircraft Guidance for Wind Shear Avoidance: Decision Making Under Uncertainty," Ph. D. Thesis, Princeton University, Department of Mechanical and Aerospace Engineering, Oct. 1992.

Severe low-altitude wind shear poses a significant threat to air transportation safety. Concepts for assisting critical decision making under uncertainty are advanced to promote the avoidance of hazardous weather, particularly microburst wind shear. Computational strategies founded on probability and optimal estimation theories enable flight deck integration of diverse forecasting and detection systems, from airport weather information services to airborne forward-looking wind sensors.

A decision-making policy for wind shear is developed from a comprehensive investigation of microburst phenomenology, its observed characteristics, and its effects on aircraft flight. Existing avoidance guidelines for

wind shear are extended to exploit the latest available technology, such as Doppler weather radar and lidar. Theories for probability-based decision making facilitate real-time computer reasoning with dynamic, conflicting data from a wide array of sources. Bayesian neural networks fused with multivariable estimators account for the limited precision, reliability, and timeliness of correlated sensor measurements. Monte Carlo analyses are conducted to refine Kalman filters for forward-looking sensors, with statistical results completing their incorporation into Bayesian reasoning.

Symbolic and numerical processes for a Wind Shear Safety Advisor are implemented and evaluated. A risk assessment model based on empirical and analytical results is used to compare the relevance of available wind shear information sources. Simulations of the risk-assessment model show its insensitivity to parameter variations. Validations of overall Wind Shear Safety Advisor logic illustrate how it conveys beneficial advance warnings in rapidly developing microburst-encounter situations. These results prove that intelligently-integrated detections systems can warn pilots of threatening wind shear sooner, more frequently, and more effectively than isolated systems can.

9. S. S. Mulgund and R. F. Stengel," Optimal Recovery from Microburst Wind Shear," *Proceedings of the AIAA Atmospheric Flight Mechanics Conference*, Hilton Head, Aug. 1992.

The flight path of a twin-jet transport aircraft is optimized in a microburst encounter during approach to landing. The objective is to execute an escape maneuver that maintains safe ground clearance and an adequate stall margin during the climb-out portion of the trajectory. A cost function penalizing rate of climb deviations from a nominal value and the rate of elevator deflection produces qualitatively good results in a variety of microburst encounters. The optimal maneuver is a gradual pitch-up that ceases near the core of the microburst, followed by a slight reduction in pitch attitude in the tailwind area of the microburst. A minimum airspeed constraint in the optimization prevents excessive airspeed loss in very severe microbursts. The aircraft equations of motion include short-period dynamics, so that the optimization solves directly for the control surface deflections required to achieve the optimal flight paths.

10. S. S. Mulgund and R. F. Stengel, "Aircraft Flight Control in Wind Shear Using Partial Dynamic Inversion," *Proceedings of the 1993 American Control Conference*, San Francisco, June 1993, pp. 400-404.

A flight control law based on partial inversion of the longitudinal dynamics of a twin-jet transport aircraft is presented. The controller is partitioned into a slow-time-scale and a fast-time scale to simplify its design. Three types of controllers are developed: airspeed/climb rate, groundspeed/climb rate, and throttle/climb rate. For microburst encounters during approach to landing, it is found that a combination of airspeed and groundspeed regulation is quite effective for controlling the flight path to touchdown. Regulation of groundspeed to a nominal value in the performance-increasing region of the microburst prevents an inadvertent reduction in thrust, while regulation of airspeed to a nominal value in the performance-decreasing area of the microburst prevents excessive airspeed loss. The throttle/climb rate controller is used for aborted-landing encounters. The combination of groundspeed and airspeed control is used until the decision is made to abort the landing, at which point maximum throttle and a specified positive climb rate are commanded.

11. D. R. Spilman, "Dynamic Response and Control of a Jet Transport to a Single-Axis Wind Vortex," M. S. E. Thesis, Princeton University, Department of Mechanical and Aerospace Engineering, Jan. 1993.

The dynamic response and control of a twin-jet transport aircraft encountering a single-axis wind vortex on final approach to landing is investigated. A horizontal wind vortex, or wind rotor, is formed by strong winds that flow over a mountain range and roll up on the leeward side of the mountain, forming a rotating airmass. If proper control action is not taken immediately after encountering a rotor, then severe performance degradation and possible ground impact may result.

A complete six-degree of freedom jet transport aircraft model that includes nonlinear aerodynamic data, unsteady aerodynamic effects, and wind-gradient effects over the aerodynamic surfaces is used to simulate an aircraft-vortex encounter. Dynamic simulations are used to determine the effects of vortex strength, vortex length, lateral entry position, vertical entry position, and encounter incidence angle on the aircraft response parameters. Roll angle and sideslip angle are primary response parameters because they may introduce performance degradation and control hazards. A large induced roll angle results from a co-axial encounter in which the vortex axis is aligned with the flight path and the wind-shear gradient is directly incident

over the aircraft wing. An encounter with a rotor oriented at a 60° angle to the flight path produces a severe sideslip angle response which in turn causes a large roll-angle response. In this case the response is highly dependent on the precise initial conditions of the encounter.

Both rudder and aileron controls are useful in alleviating vortex-induced roll; however, rudder control may excite lightly damped Dutch-roll dynamics. A simple lateral-directional linear-quadratic controller that uses rudder to control sideslip and aileron to control roll successfully controls the simulated aircraft through strong wind vortices without exciting unwanted dynamics. In addition to demonstrating the value of using automatic control to reduce the wind vortex hazard, such a control system has benefits beyond its immediate design goals. Because of the similarities between wake vortex flows and mountain-wave vortex flows, the controller may be used to reduce required separation distances at airports. It also may prove superior to a human pilot in preventing catastrophic low-altitude control system failures.

12. R. F. Stengel and J. P. Wangermann, "Air Traffic Management as Principled Negotiation Between Intelligent Agents," presented at the AGARD Guidance and Control Symposium, *Machine Intelligence in Air Traffic Management*, Berlin, May 1993.

Air transportation provides the backbone for passenger transport over moderate to long distances in the U.S. and much of the world, and it is becoming an increasingly important mode for short-range travel and cargo transport as well. There is a growing demand for use of available airspace and a heightened concern for on-time performance. Demand frequently exceeds available capacity of the airspace system, causing flight delays, negative economic impact, and passenger inconvenience [1, 2]. New technologies are emerging that will make flight operations both simpler and more complex. On the one hand, advances hold promise for increasing the productivity, reliability, and safety of the air transportation system. On the other, advances in technology introduce uncertainty, increase human workload (if not properly implemented), increase the potential for dispute, and present new challenges for both certification and day-to-day operations. This paper presents a concept for an *Intelligent Aircraft/Airspace System (IAAS)* that could be a focal point for developing air traffic management in the coming decades. The *IAAS* would integrate the capabilities of all ground-based and airborne components of the system (identified as *Intelligent Agents*) in order to provide increased capacity and maintained or improved safety. *Principled Negotiation* is proposed as a framework for interactions between intelligent agents.

13. R. F. Stengel, "Intelligent Flight Control Systems," presented at the *IMA/RAS Conference on Aerospace Vehicle Dynamics and Control*, Cranfield, UK, Sept. 1992.

The capabilities of flight control systems can be enhanced by designing them to emulate functions of natural intelligence. Intelligent control functions fall in three categories. *Declarative* actions involve decision-making, providing models for system monitoring, goal planning, and system/scenario identification. *Procedural* actions concern skilled behavior and have parallels in guidance, navigation, and adaptation. *Reflexive* actions are spontaneous, inner-loop responses for control and estimation. Intelligent flight control systems learn knowledge of the aircraft and its mission and adapt to changes in the flight environment. Cognitive models form an efficient basis for integrating "outer-loop/inner-loop" control functions and for developing robust parallel-processing algorithms.

14. R. F. Stengel, "Aerospace Optical Communications Abstracts," Princeton University, Department of Mechanical and Aerospace Engineering, Princeton, NJ, May 26, 1993.

Over 100 abstracts related to the possible application of optical communication to aircraft were drawn from the AIAA Aerospace Abstracts. The abstracts describe papers published between 1989 and 1983. Although the papers focus primarily on space applications, several address aircraft-to-aircraft and aircraft-to-satellite communications.

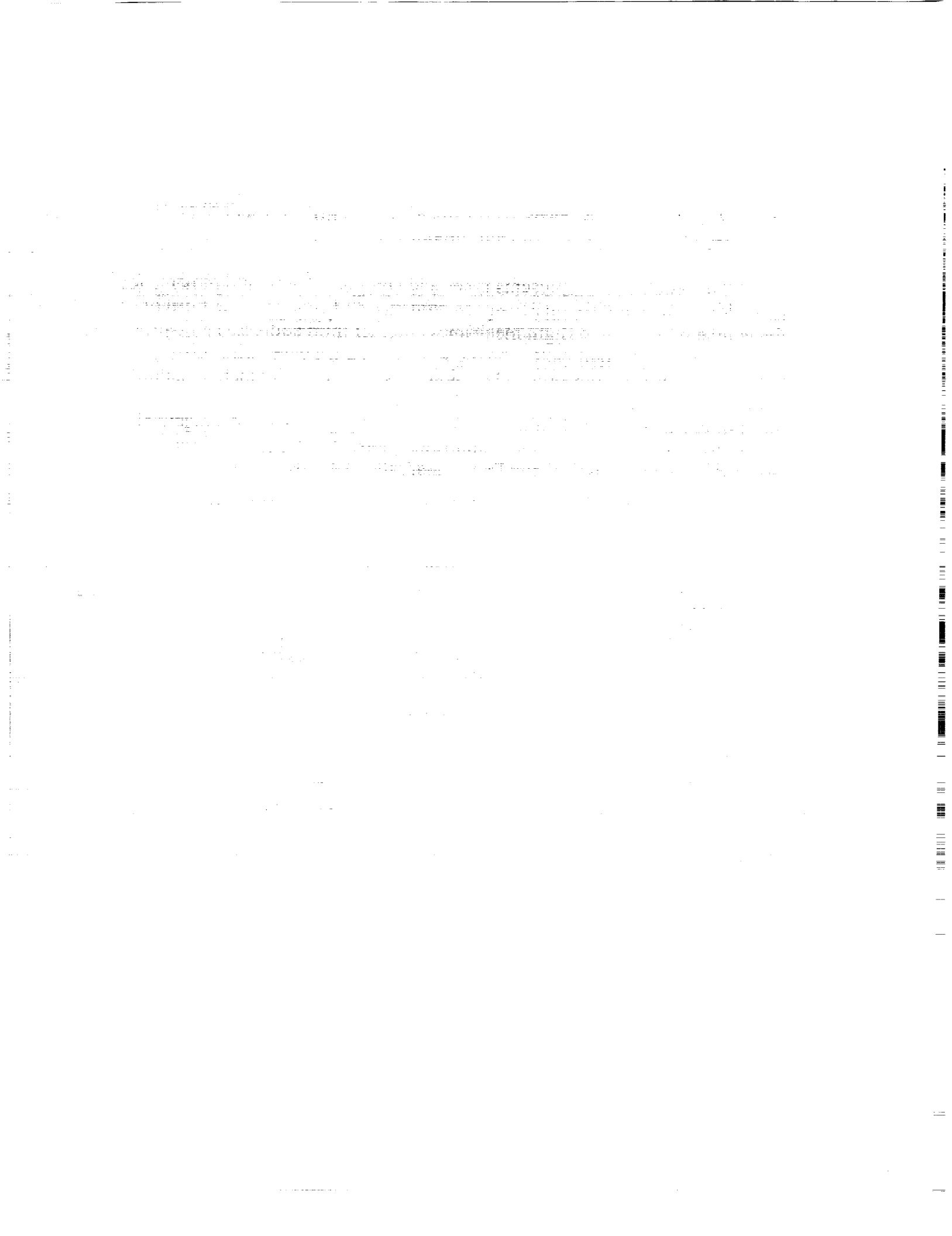
15. B. L. Belkin and R. F. Stengel, "AUTOCREW: A Paradigm for Intelligent Flight Control," *An Introduction to Intelligent and Autonomous Control*, P. Antsaklis and K. Passino, ed., Kluwer Academic Publishers, Norwell, MA, 1993, pp. 371-400.

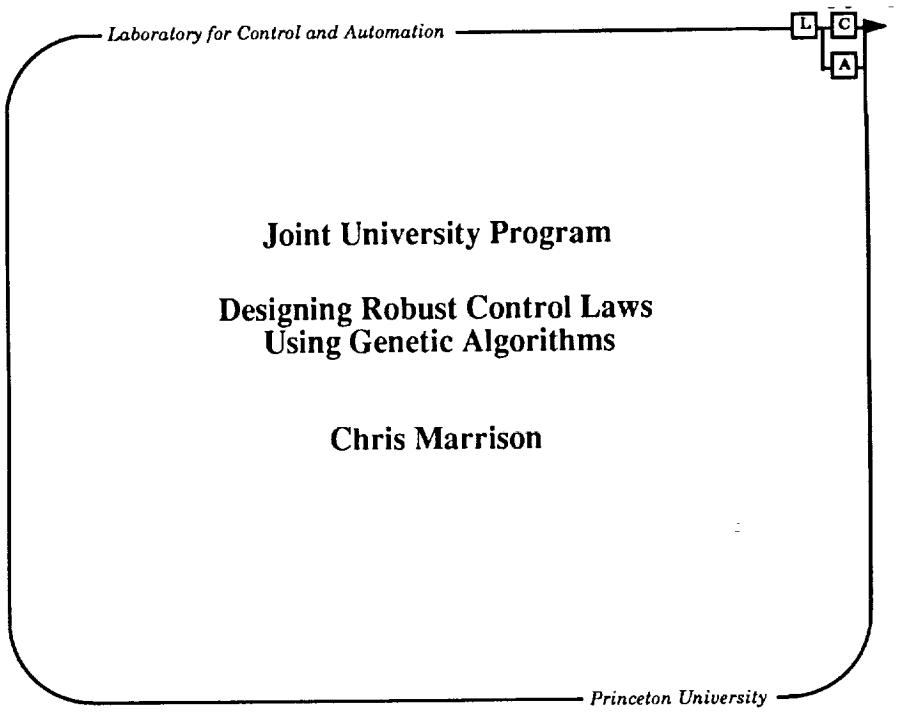
An expert system Pilot-Aid is envisioned to automate many functional and low-level decision-making tasks in future high performance and jet transport aircraft to help alleviate pilot workload. Nine modular rule-based systems, collectively called AUTOCREW, were designed to automate functions and decisions associated with a combat aircraft's subsystems. The knowledge bases were designed individually; areas of cooperation between the knowledge bases were identified, and common information was designated as "shared" information. An interactive graphical simulation testbed was developed to demonstrate and test the cooperating AUTOCREW ensemble's performance. Workload metrics were formulated to quantify AUTOCREW's performance in terms of the ensemble's efforts in assisting the Pilot. The workload metrics give reasonable results for the comparison

of workloads among AUTOCREW's experts, as well as comparative results among task groups within a single knowledge base. The applicability of the methods utilized to design AUTOCREW for other applications is also discussed.

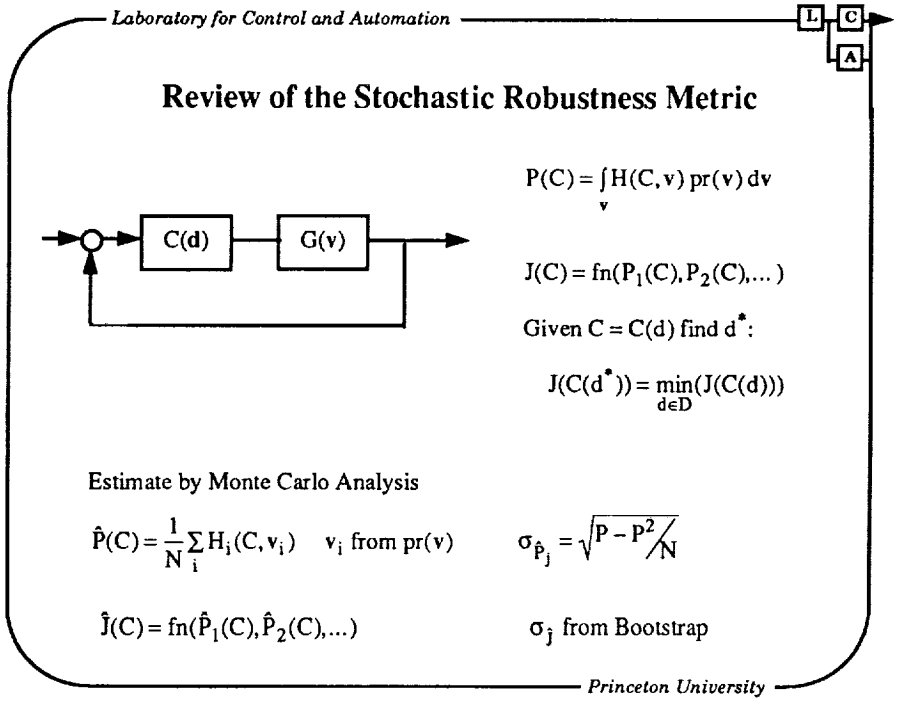
16. R. F. Stengel, "Aerospace Guidance and Control," *Using MATLAB in the Classroom*, Prentice Hall, Englewood Cliffs, 1993, pp. 3-26.

This book chapter presents notes and computer-based assignments for an undergraduate course on aerospace guidance and control. One third of the course is devoted to flight mechanics, another third addresses guidance and control of the flight path, and the remaining third deals with instrumentation for measuring position and motion. The course assignments include computational flight tests with a six-degree-of-freedom simulation of a light aircraft; calculations of stability- and control-derivative matrices, eigenvalues, transfer functions; root locus and Bode plots; and design of flight control systems using classical and linear-quadratic methods.





The purpose of this research is to create a method of finding practical, robust control laws. The robustness of a controller is judged by Stochastic Robustness metrics and the level of robustness is optimized by searching for design parameters that minimize a robustness cost function.

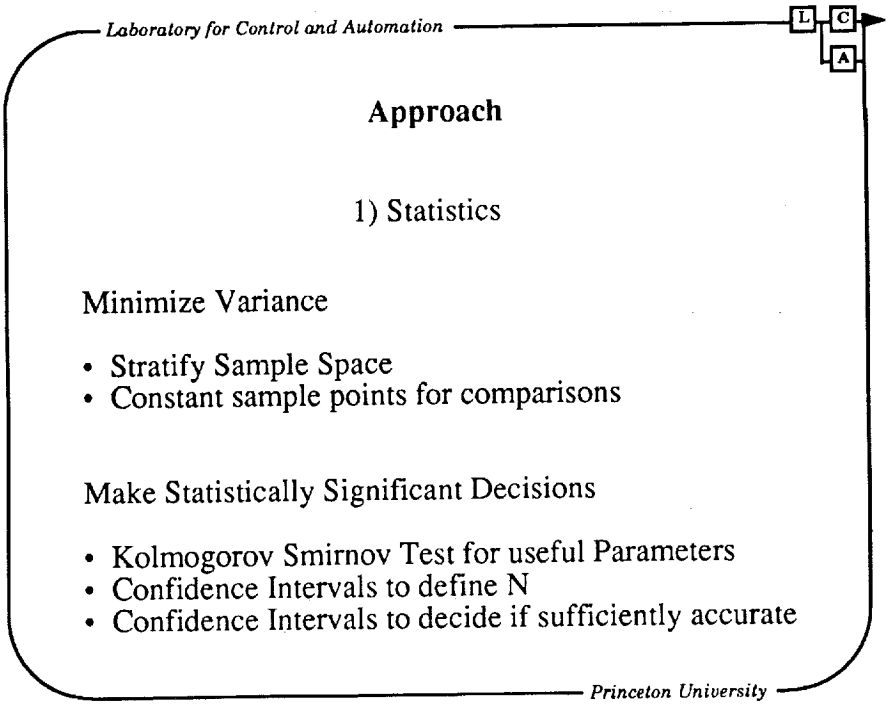


Given the expected variation of the plant parameters, a Stochastic Robustness metric characterizes a compensator by giving the probability that the compensator will fail to perform acceptably. The definition of what is unacceptable is left to the designer but will normally include such features as instability and slow response time. To calculate the probability of unacceptability, P, the indicating function, H(C,**v**) must be integrated over the space of expected parameter variations. H is a function of both the compensator, C and the plant parameter values, **v**. H equals one when the metric is violated and zero otherwise.

Normally, more than one metric will be of importance in a given application. In such a case it may be necessary to make a trade-off between the metrics. The trade-off can be formalized by combining the probabilities into a scalar cost function, J. Weights within the cost function can then be used to reflect the importance to the designer of each metric.

Once J is defined, the task is to find the set of plant design parameters, **d**, to minimize J.

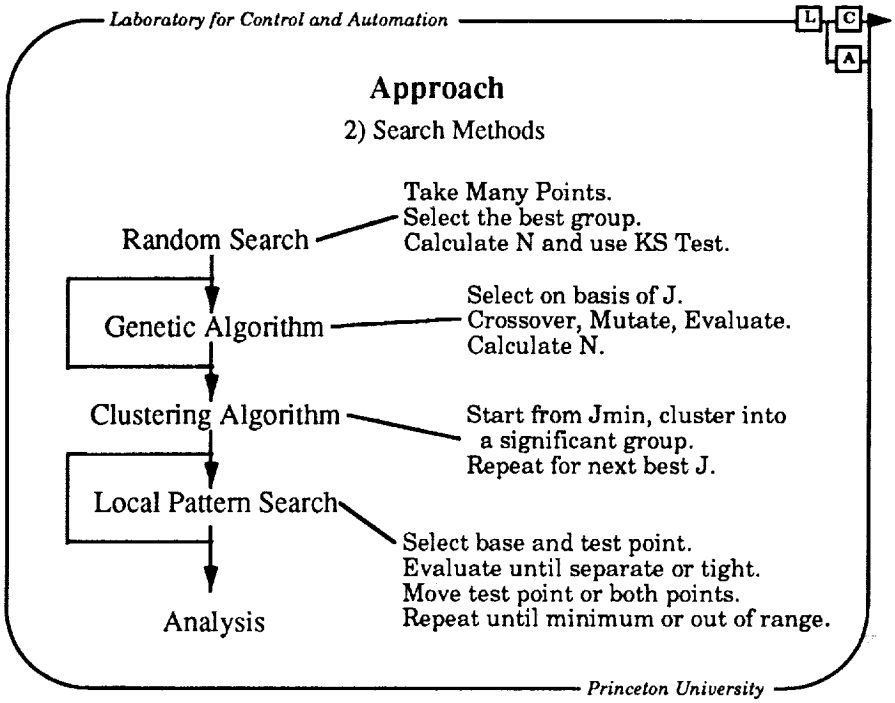
This task is hindered by the fact that it is normally impossible to evaluate the probabilities analytically. An alternative evaluation method is to use Monte Carlo Analysis; this has the disadvantage that errors can be expected in the estimate of P. The expected error reduces as the inverse of the square root of the number of evaluations. There is therefore a trade-off between the accuracy and of the evaluations and the computation time.



The approach to finding a stochastic global optimization method has two main thrusts. The first is to understand the statistical effects of the Monte Carlo Analysis and exploit them to reduce the number of evaluations necessary. The second approach is to identify suitable search algorithms.

The variability in the estimates of P has been reduced significantly by stratifying the sample space and by using the same sample points when comparing two compensators. An understanding of these statistical mechanisms has allowed a significant reduction in the number of evaluations which must be carried out to compare two compensators.

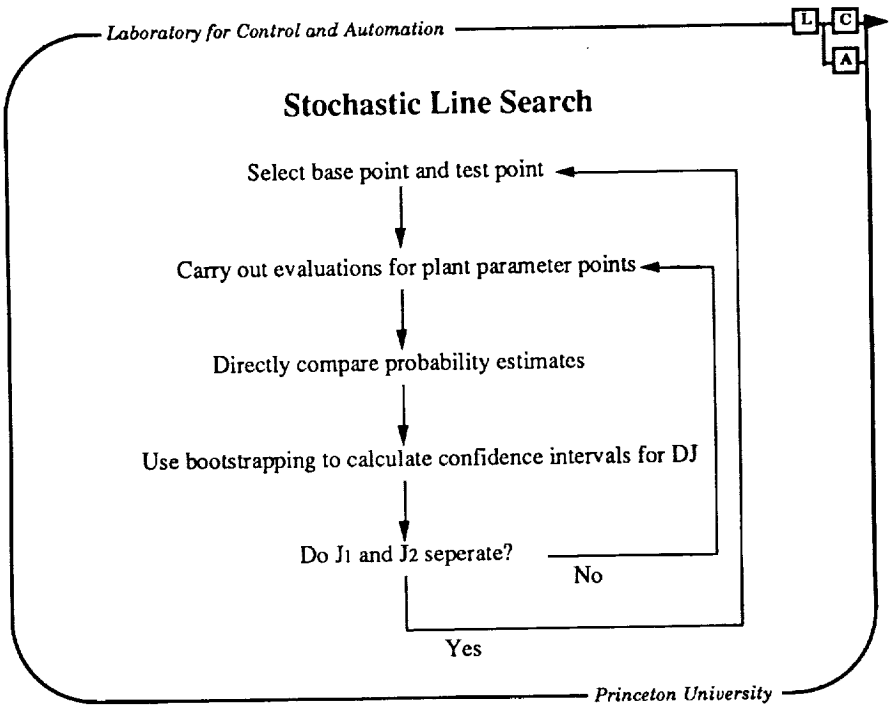
By using the Kolmogorov-Smirnov test it is possible to identify parameters that have a significant effect on J, allowing the search to concentrate on these parameters. The establishment of confidence intervals on the estimates of P provide a basis for making statistically significant search decisions and also to fix the number of further evaluations that must be required if the results are not yet statistically significant.



A wide range of modern search methods were screened for their possible use in searching stochastic space. The most efficient method combines the best qualities of several different methods.

The proposed search method begins by taking a broad, completely random, search across the design space. A few evaluations are made at each point and the best points are then presented as the starting population for a genetic algorithm. The genetic algorithm carries out the bulk of the search and later will be described in detail.

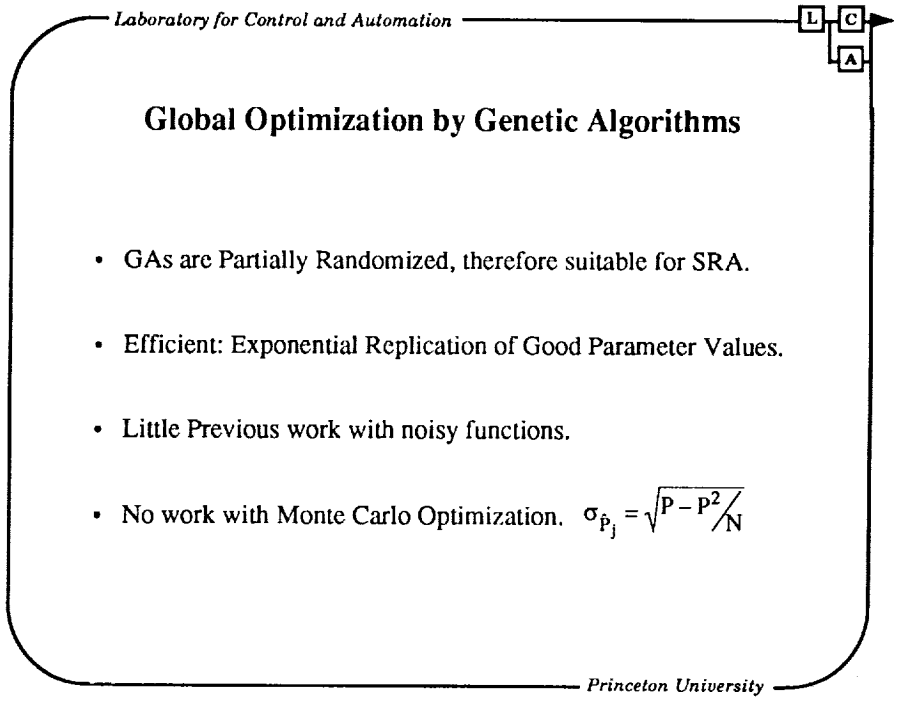
The result of the genetic algorithm is a set of candidate solutions, most of which should be close to the global minimum. A clustering algorithm is then used to identify groups of good solutions and a local line search is carried out from the centroid of each cluster.



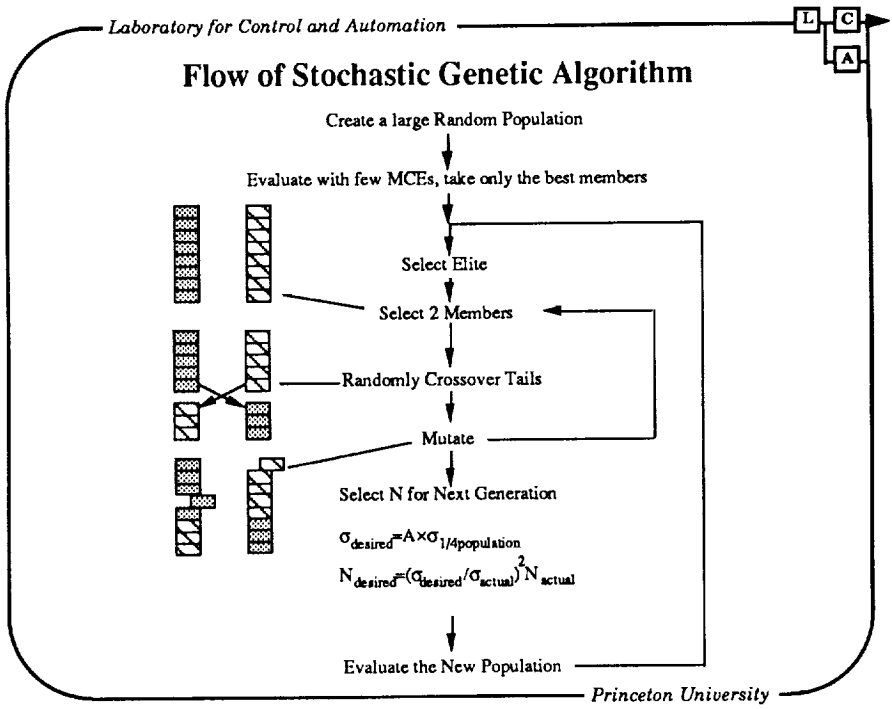
The line search is based on a pattern search with additional logic to deal with the uncertainties introduced by the Monte Carlo Evaluation.

The search moves along the line, comparing two points at a time. A set amount of Monte Carlo Evaluations are carried out and then a decision is made as to where along the line the next evaluations should be made. The decision is based on an estimate of the likely error in J. If the errors are relatively small then we can be confident that there is a true difference between the compensators and a new search point can be chosen. If the error is relatively large, more evaluations need to be carried out.

This search method has been implemented, and is effective in finding the minimum along a line in design parameter space.



Genetic Algorithms (GAs) were chosen as the main global optimization method. These algorithms have several attributes that make them well suited to searching a stochastic space. They rely on a partially randomized comparison of many points and are therefore insensitive to errors due to Monte Carlo Evaluation and they process information efficiently. However, little previous work has been done in using GAs to optimize noisy functions. This work must be carried out before using GAs for the synthesis of robust control .



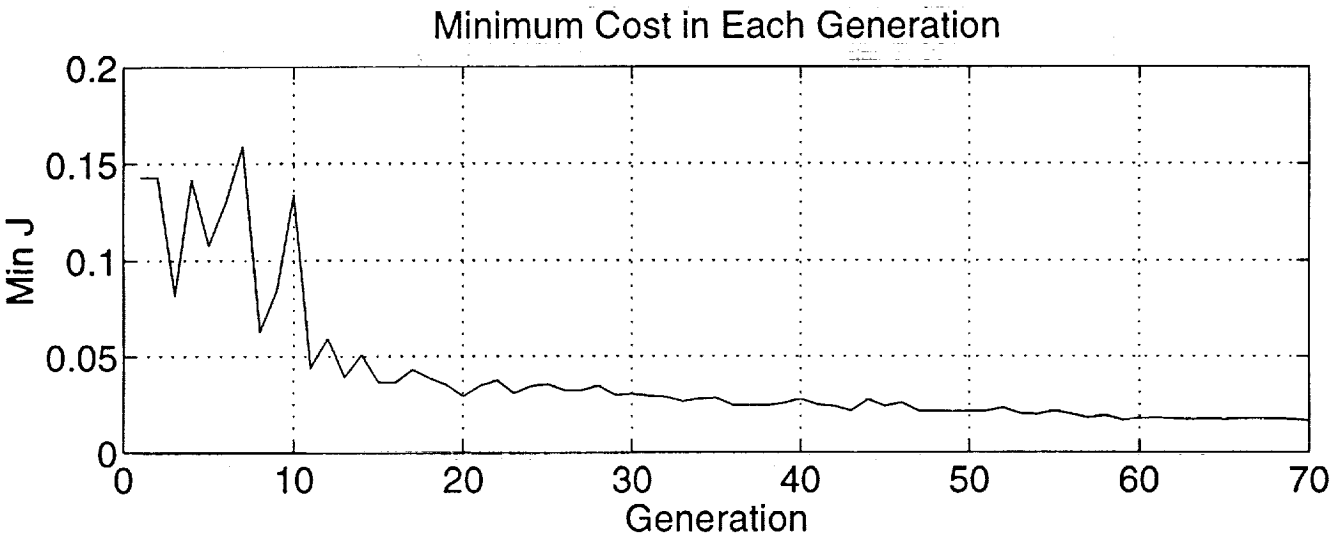
The Stochastic Genetic Algorithm (SGA) is currently being researched. The basic structure of the SGA is shown above. The SGA is similar to normal GAs except for 3 points:

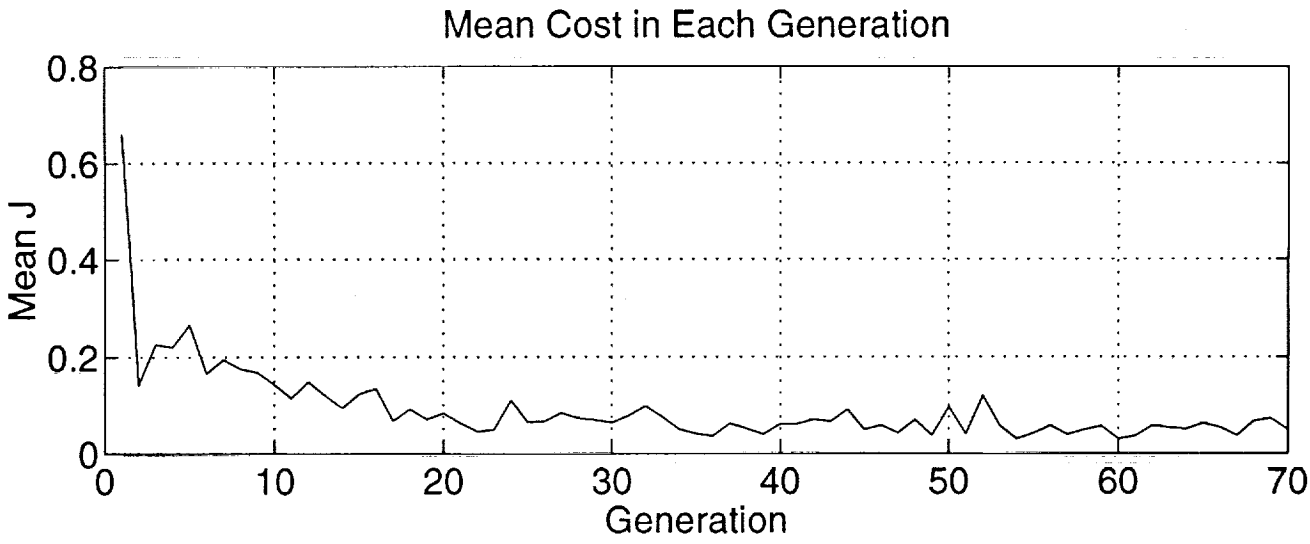
1) The search begins with a random search, using a few Monte Carlo Evaluations at each point, and using a small proportion of the random points as the initial population to 'kick-start' the SGA.

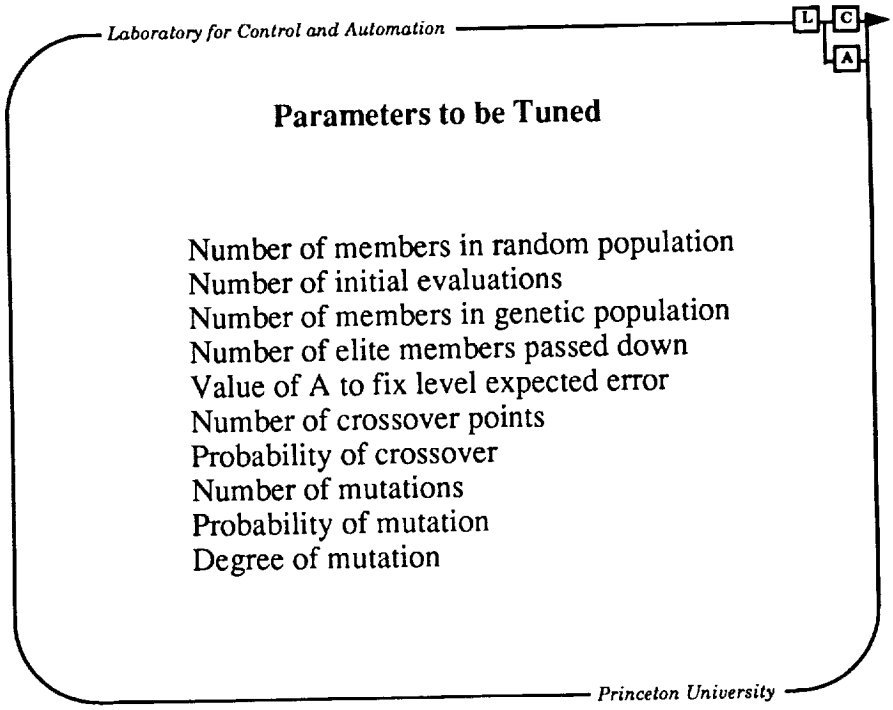
2) The Kolmogorov-Smirnov test is used to determine which design parameters are most important in affecting J. These parameters are used to cluster the best members of the population to form one averaged member. This is passed as an elite member into the next generation.

3) The number of evaluations per point, N, is fixed before each set of Monte Carlo Evaluations. This is done by comparing the expected error in the estimated cost of the best member of the population with the mean difference between the costs of the rest of the population. The ratio of the error allowed in the estimate to the difference in the cost of the population can be varied to improve the performance of the search. Here, this parameter is referred to as "A".

The next graphs show the results of a typical run of the SGA. The first graph shows the values of J for the best member in the population of each generation as the population evolves to a low value of J. The second graph shows the mean value for J for each generation.

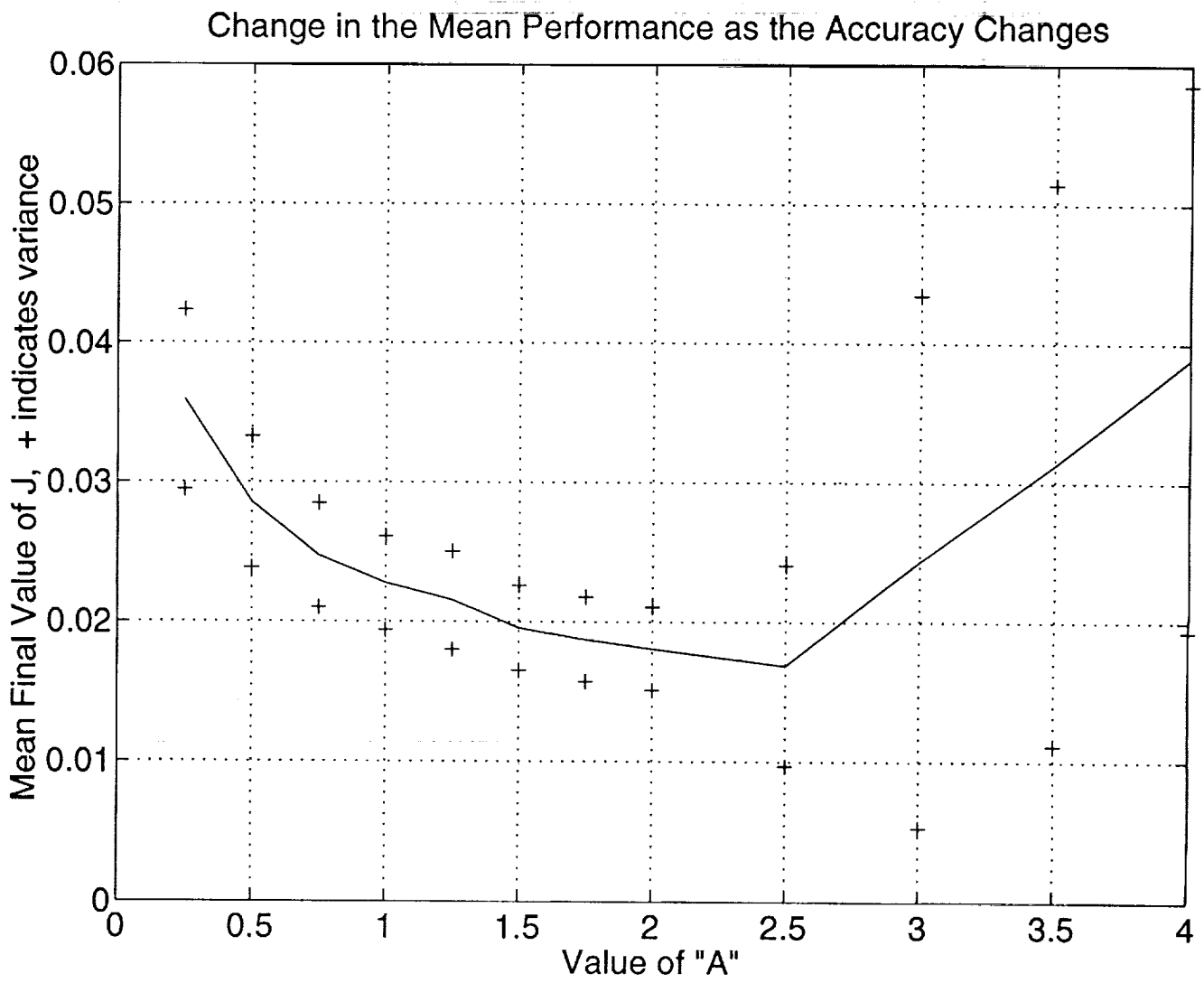

Minimum Cost in Each Generation
Min J
0.2
0.15
0.1
0.05
0
0
10
20
30
40
50
60
70
Generation

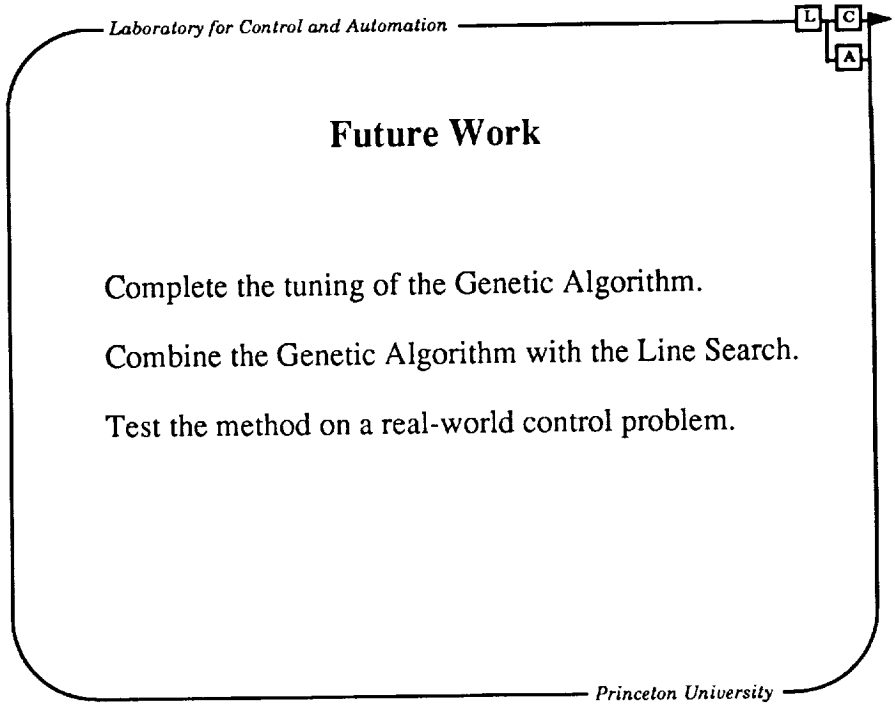

Mean Cost in Each Generation
Mean J
0.8
0.6
0.4
0.2
0
0
10
20
30
40
50
60
70
Generation



Within the SGA there are several parameters that must be carefully chosen to ensure that the search is efficient. These parameters are being tuned by running the SGA repeatedly on a test function, adjusting the parameter, and running the SGA again.

The next graph shows the effect of changing the value of A. Here the SGA was run 150 times for each of 12 different values of A. At low values of A, few evaluations are carried out per point and the SGA does not have information of sufficient quality to converge well; with high values of A the information is of higher quality than needed and the computational effort would be better spent searching more points. The optimum value is between 2 and 3. With A = 3 the performance is occasionally very good but on average the result is mediocre. With A = 2.5 the performance will on average be the best but there is a relatively wide variability. With A = 2 the average performance is not quite so good but the search is more robust; the variability is less and the search is less likely to result in a poor outcome.


Change in the Mean Performance as the Accuracy Changes
Mean Final Value of J, + indicates variance
0.06
0.05
0.04
0.03
0.02
0.01
0
0
0.5
1
1.5
2
2.5
3
3.5
4
Value of "A"



Future work will complete the tuning of the SGA and combine it with the line search. The overall algorithm will then be tested against real world control synthesis problems.

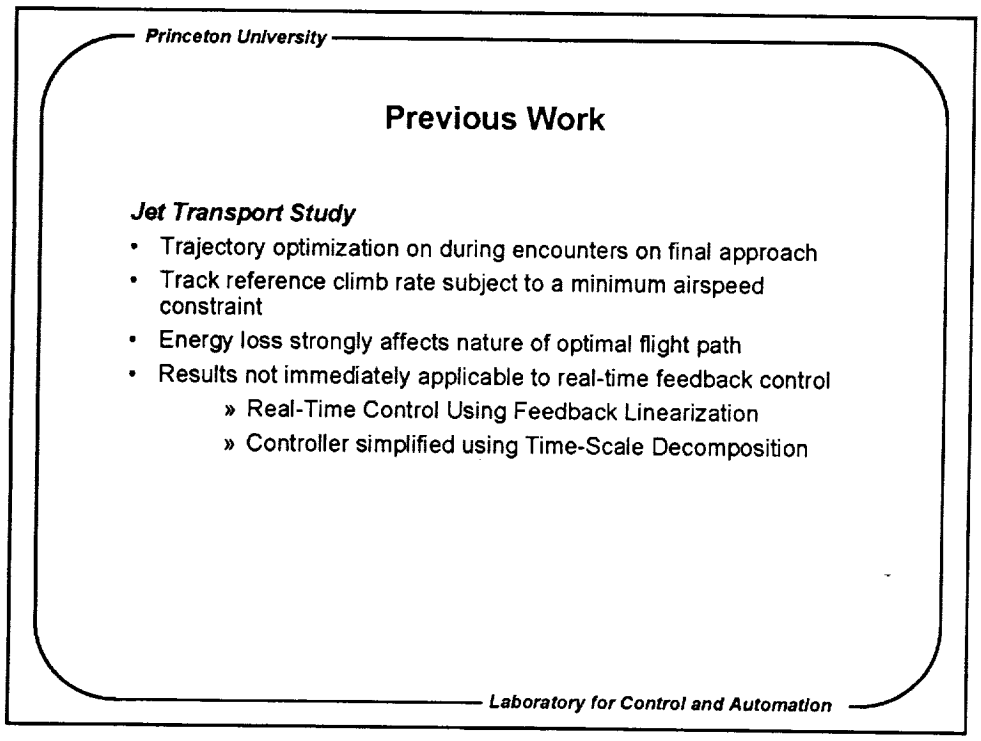
Princeton University

Optimal Nonlinear Estimation for Aircraft Flight Control in Wind Shear

Sandeep S. Mulgund
Department of Mechanical and Aerospace Engineering
Princeton University

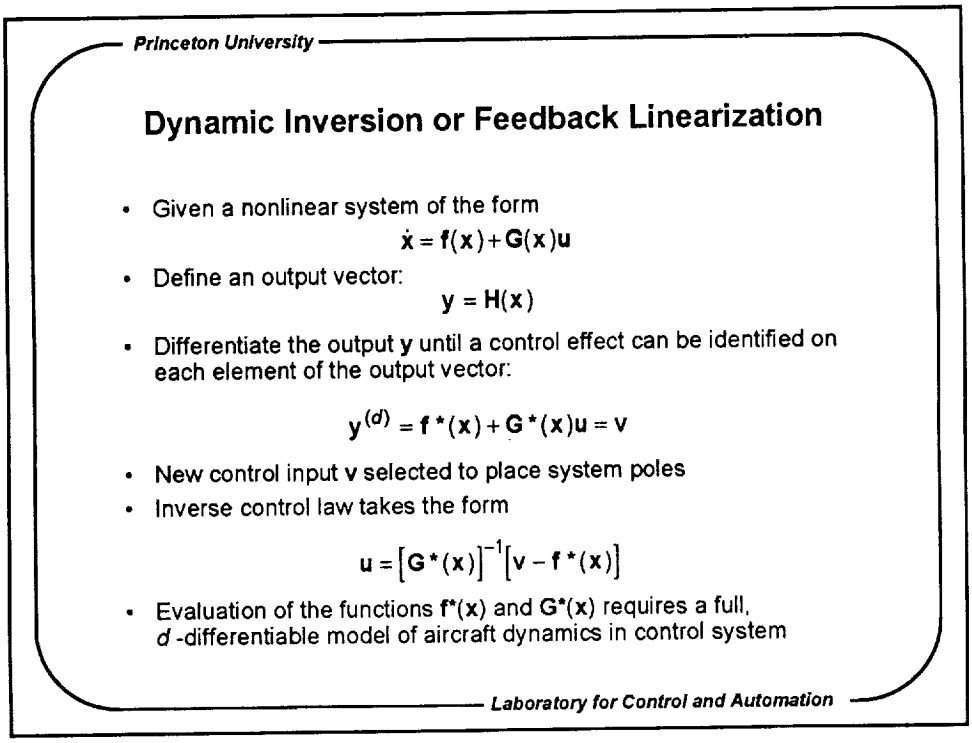
JUP Quarterly Review
April 1-2, 1993
Princeton, NJ

Laboratory for Control and Automation

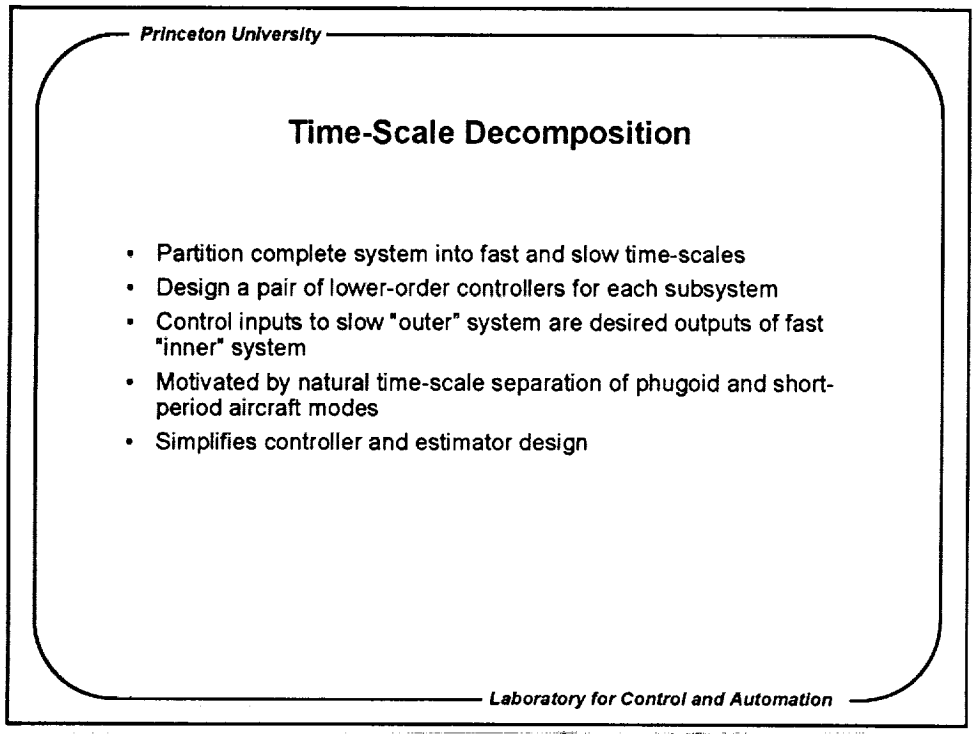


This presentation describes the most recent results in an ongoing research effort at Princeton in the area of flight dynamics in wind shear. The first undertaking in this project was a trajectory optimization study. The flight path of a medium-haul twin-jet transport aircraft was optimized during microburst encounters on final approach. The assumed goal was to track a reference climb rate during an aborted landing, subject to a minimum airspeed constraint. The results demonstrated that the energy loss through the microburst significantly affected the qualitative nature of the optimal flight path. In microbursts of light to moderate strength, the aircraft was able to track the reference climb rate successfully. In severe microbursts, the minimum airspeed constraint in the optimization forced the aircraft to settle on a climb rate smaller than the target. A tradeoff was forced between the objectives of flight path tracking and stall prevention.

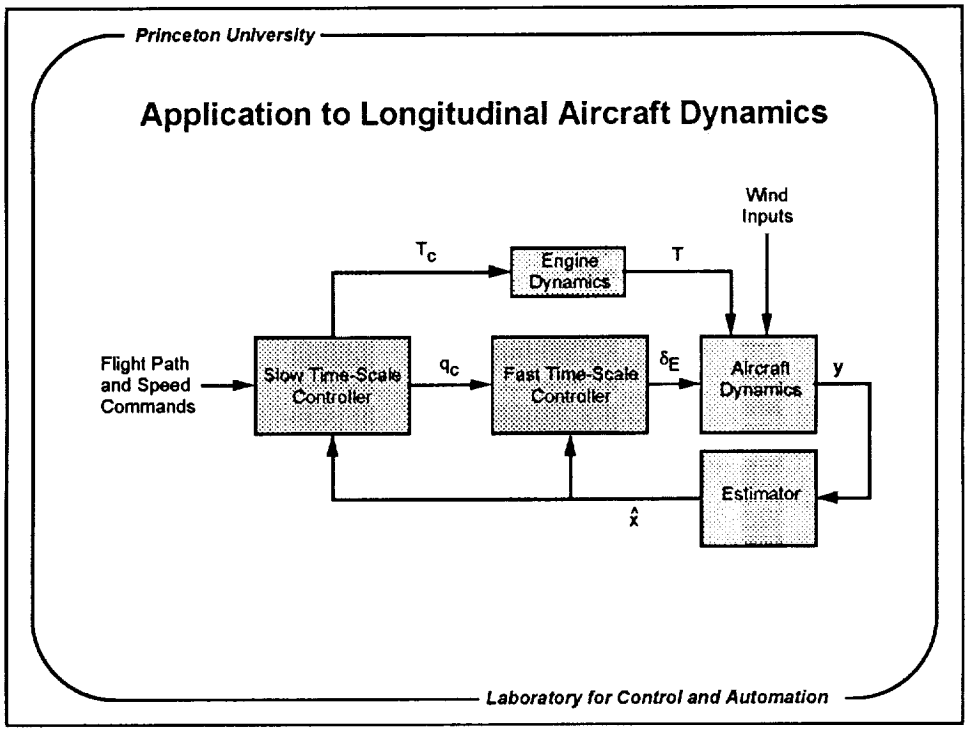
Although the results provided a qualitative picture of the nature of an optimal control strategy in wind shear, they were not immediately applicable to real-time control. Optimization is an iterative process requiring global knowledge of the flow field. Therefore, an initiative was undertaken to develop feedback control methods that approximated the performance realized in the optimal trajectories. The technique of nonlinear inverse dynamics or feedback linearization was used to develop a control law for a nonlinear model of the aircraft dynamics. The control design was simplified using *Time-Scale Decomposition*, which permitted the partitioning of the controller into a slow outer loop and a fast inner loop.



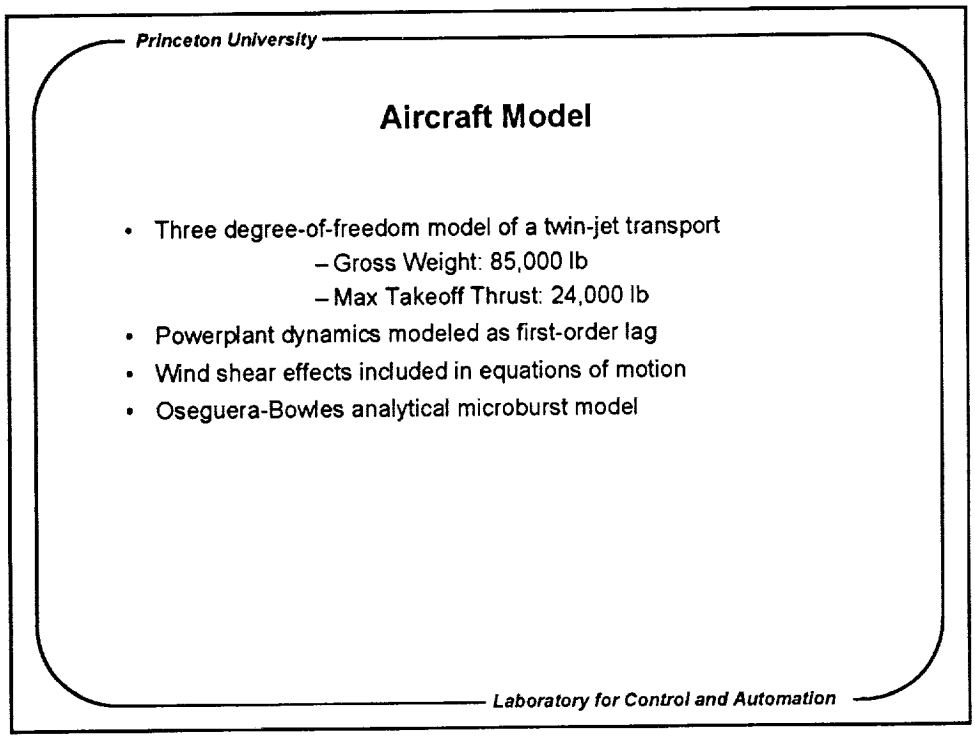
The control law designed for the aircraft model was based on the technique of dynamic inversion or feedback linearization. Given a nonlinear system of the form shown, it is possible to define an output vector **y** which is a known function of the system state **x**. This output is differentiated with respect to time until a control effect can be identified on each element of the output vector. The d^{th} derivative of the output is then equated to a new control input **v**. This control input can be selected to place the system poles in designer-specified locations, subject to the controllability of the original system. Although the form of the resultant nonlinear control law appears simple, the evaluation of its components requires that a full, d-differentiable model of the plant dynamics be included in the control system.



The control law based on nonlinear inverse dynamics can be simplified if it is possible to partition the original system into fast and slow time scales. If this is feasible, it is possible to design a pair of lower-order controllers for each subsystem. The control inputs to the slow "outer" system are the desired outputs of the fast "inner" system. For the aircraft problem, the time-scale decomposition is motivated by the time-scale separation that exists between the phugoid and short-period modes. The application of this technique simplifies both the controller and estimator design. Two lower-order controllers can be designed, and fewer system state derivatives must be estimated.



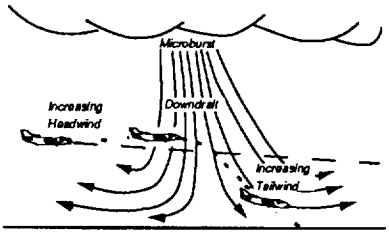
The structure of the nonlinear control law using time-scale decomposition is illustrated here for our aircraft study. The slow outer controller accepts flight path and speed commands. It generates a throttle and pitch rate command. The throttle command is passed on to the engine dynamics. The pitch rate command becomes the desired response of the fast inner controller. This controller generates the elevator deflection required to achieve the desired pitch rate. This controller is designed to have a response time at least 3 to 5 times faster than the outer controller. Thus from the perspective of the outer controller, the necessary pitch rate is achieved almost instantaneously. The elevator deflection calculated by the fast controller is fed into the aircraft dynamics, as is the actual thrust level produced by the engine dynamics. The output of the aircraft sensors is fed into an estimator, which generates the aircraft state estimate needed to accomplish the inversion. The design of this estimator and the performance of the controller/estimator pair are the subject of the rest of this presentation.



A three degree-of-freedom model of a twin-jet transport aircraft was used for this study. The aircraft has the given gross weight and maximum takeoff thrust. The powerplant dynamics are modeled as a first-order lag, and thrust lapse with mach number and altitude is also modeled. Wind shear effects are incorporated into the equations of motion, and the Oseguera-Bowles microburst model (developed at NASA Langley Research Center) provides the wind inputs used in simulated microburst encounters.

Princeton University

Control Strategies in Microburst Wind Shear



Airspeed Control

- Undesirable thrust reduction in headwind region
- Maintains airspeed in tailwind

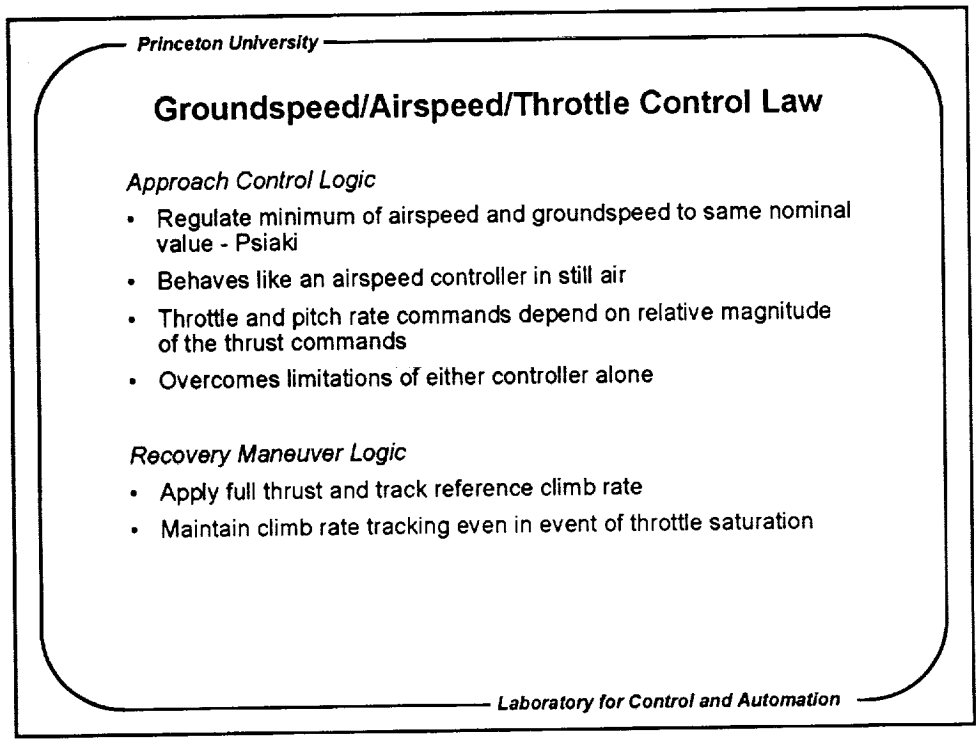
Groundspeed Control

- Maintains thrust in headwind region
- Airspeed loss in tailwind

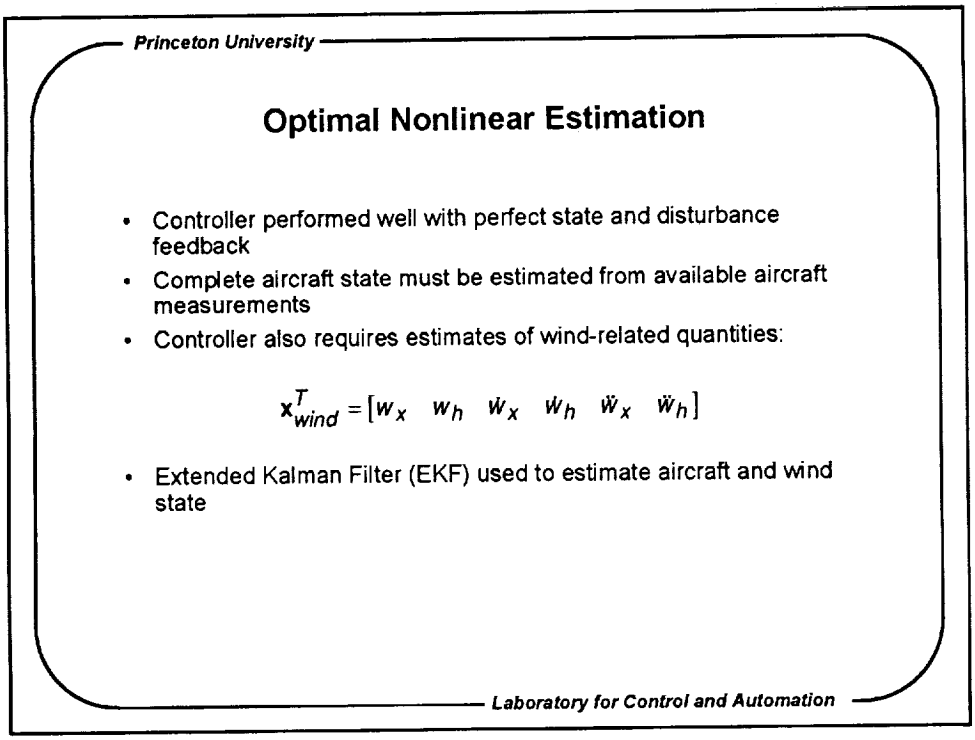
Laboratory for Control and Automation

The control law described earlier is designed to track reference speed and flight path inputs. It is worthwhile to consider what types of guidance strategies are suitable in a microburst environment. In a classical microburst encounter, the aircraft first encounters an increasing headwind. The airspeed increases, and the aircraft may balloon above the nominal flight path. If the flight crew is not alert to the fact that a microburst is present, they may take action to prevent the plane from climbing by throttling back and/or lowering the aircraft's nose. This headwind soon transitions to a downdraft, which may result in an increased sink rate. The subsequent tailwind causes an airspeed loss, and ground impact may result if the pilot does not apply an effective recovery technique.

Regulating airspeed about a nominal value causes an undesirable reduction in thrust in the headwind region of the shear to prevent an unwanted airspeed increase. This may leave the aircraft in a precarious state once it enters the performance-decreasing downdraft and tailwind. However, airspeed is maintained in the tailwind region, subject to the powerplant performance limits. Conversely, regulation of groundspeed maintains thrust in the headwind region. A thrust increase is typically required in the headwind region to maintain a nominal groundspeed. In the tailwind region, however, groundspeed regulation results in an airspeed loss and may lead to stall if the airspeed becomes too low. Taken together, these observations suggest that an effective strategy might be one that combines the desirable traits of groundspeed and airspeed control.



The guidance strategy used with the nonlinear control law was adapted from one developed by Mark Psiaki of Cornell University. The approach control logic regulates the minimum of airspeed and groundspeed to the same nominal value. This behaves like an airspeed controller in still air. In the current implementation, the throttle and pitch rate commands passed onto the aircraft dynamics depend on the relative magnitudes of the thrust commands generated by an airspeed/climb rate and a groundspeed/climb rate controller. This control logic overcomes the limitations of either airspeed or groundspeed control alone. During a recovery maneuver (where a decision is made to abort an approach and execute an escape trajectory), full thrust is applied directly together with a climb rate command.



The control logic described earlier was found to perform well with perfect aircraft and wind state feedback. The time-scale separation assumption was demonstrated to be valid, and the controller provided good recovery performance in a broad spectrum of microbursts. In practice, however, the complete aircraft state must be estimated from the available air-data and inertial measurements. The controller also requires feedback of the two wind components (horizontal and vertical) together with their first and second time-derivatives. The *Extended Kalman Filter* (EKF) was postulated as a candidate estimator structure for this problem.

Princeton University

Continuous-Discrete Extended Kalman Filter

- Given a system of the form

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] + \mathbf{L}\mathbf{w}(t)$$

where

$$E[\mathbf{w}(t)] = \mathbf{0}$$

$$E\left[\mathbf{w}(t)\mathbf{w}^T(\tau)\right] = \mathbf{Q}_c\delta(t-\tau)$$

- Discrete Measurement Model:

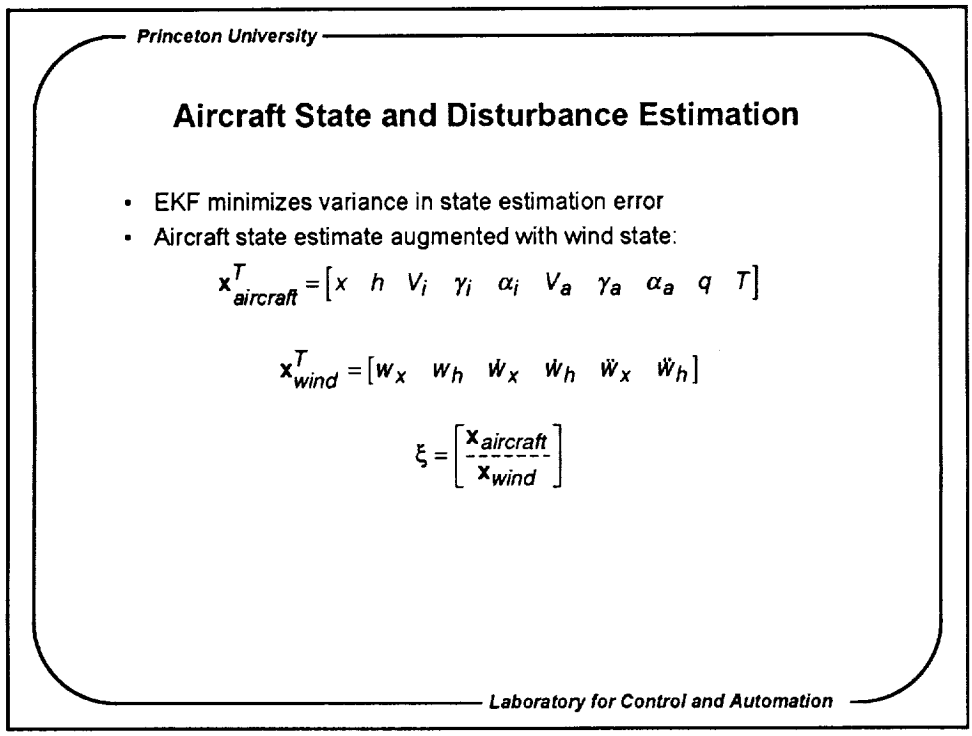
$$\mathbf{z}_k = \mathbf{h}[\mathbf{x}(t_k)] + \mathbf{n}_k \quad k = 0, 1, 2, \ldots$$

$$E(\mathbf{n}_k) = \mathbf{0}$$

$$E\left(\mathbf{n}_k\mathbf{n}_k^T\right) = \mathbf{R}_k$$

Laboratory for Control and Automation

This form of the EKF is based upon a continuous model of plant dynamics and a discrete measurement model. The disturbance $\mathbf{w}(t)$ influencing the plant dynamics is assumed to be a zero-mean Gaussian white noise process with a known spectral density matrix $\mathbf{Q}$. The measurements $\mathbf{z}_k$ are made at discrete instances t_k, and are known functions of the plant state. The measurement noise vector $\mathbf{n}_k$ is assumed to be a zero-mean Gaussian white noise process with known covariance $\mathbf{R}$.



The form of the EKF for the aircraft problem is now described. The nonlinear control law requires feedback of the wind state in addition to that of the aircraft state. This is achieved by defining the system state to consist of the aircraft and wind state. The wind state is defined to be the horizontal and vertical wind components, together with their first two time-derivatives.

Princeton University

Complete System Equations

- Wind dynamics:

$$\begin{bmatrix} \dot{w}_x \\ \dot{w}_h \\ \dot{\dot{w}}_x \\ \dot{\dot{w}}_h \\ \dddot{w}_x \\ \dddot{w}_h \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_x \\ w_h \\ \dot{w}_x \\ \dot{w}_h \\ \ddot{w}_x \\ \ddot{w}_h \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$\dot{\mathbf{x}}_{wind} = \mathbf{F}_{wind}\mathbf{x}_{wind} + \mathbf{L}_{wind}\mathbf{w}$$

- Aircraft Dynamics:

$$\dot{\mathbf{x}}_{aircraft} = \mathbf{f}(\mathbf{x}_{aircraft}, \mathbf{x}_{wind}, \mathbf{u})$$

- Complete System Equations:

$$\dot{\xi} = \begin{bmatrix} \dot{\mathbf{x}}_{aircraft} \\ \hline \dot{\mathbf{x}}_{wind} \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mathbf{x}_{aircraft}, \mathbf{x}_{wind}, \mathbf{u}) \\ \mathbf{F}_{wind}\mathbf{x}_{wind} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{L}_{wind} \end{bmatrix} \mathbf{w}$$

Laboratory for Control and Automation

The wind dynamics are modeled as a linear system driven by an external input **w**. The components of **w** are thus the third time-derivatives of w_x and w_h. The complete system equations shown here become the basis of the Kalman filter equations presented earlier.

Princeton University

Simulation Examples

- Measurements:

$$\mathbf{z}^T = \begin{bmatrix} h & V_i & V_a & \alpha_a & \theta & q & \dot{h} & \dot{x} & \ddot{h} \end{bmatrix}$$

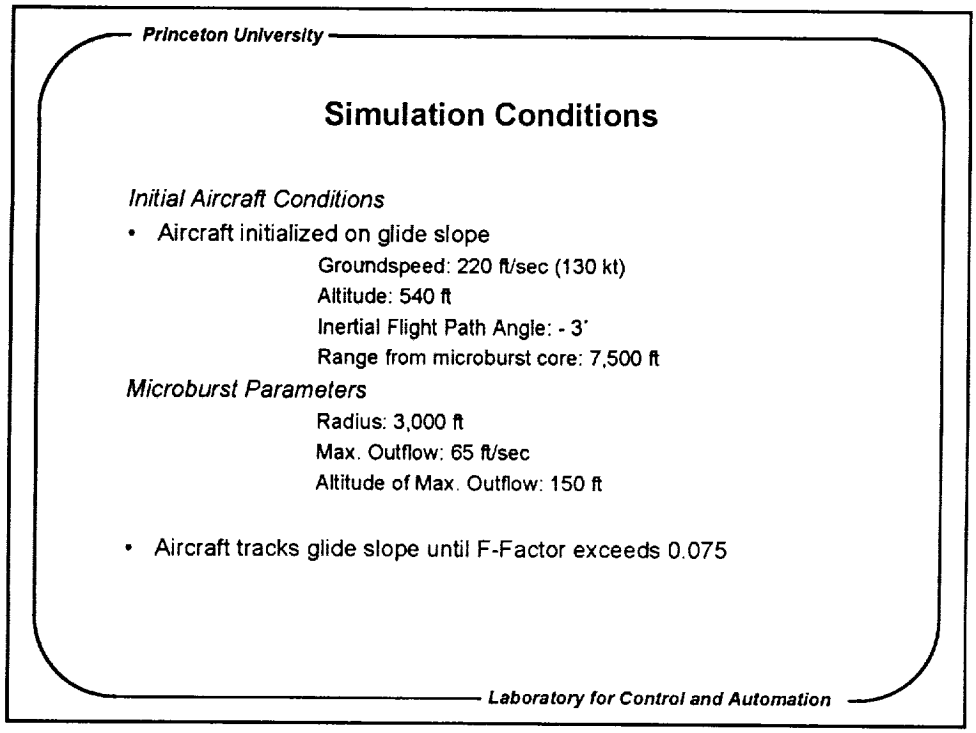
$$\mathbf{R_o} = diag\left(5,\ 4,\ 4,\ 0.025^2,\ 0.025^2,\ 0.025^2,\ 2,\ 2,\ 2\right)$$

- Sensor noise statistics based on conservative estimates of expected accuracy

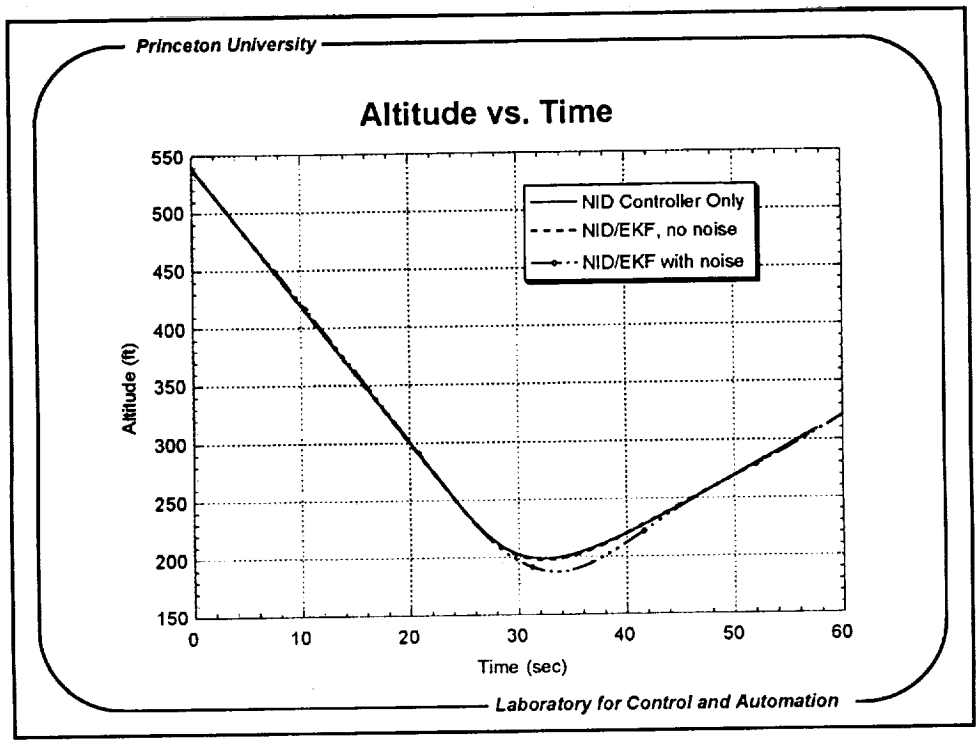
Case Number	Simulation Parameters	Measurement and Control Model
1	NID only	$\mathbf{u} = \mathbf{g}(\mathbf{x})$
2	NID and EKF; Perfect measurements	$\mathbf{u} = \mathbf{g}(\hat{\mathbf{x}})$ $\mathbf{z} = \mathbf{h}(\mathbf{x}) \Leftrightarrow \mathbf{R} \approx \mathbf{0}$
3	NID and EKF; Noisy measurements	$\mathbf{u} = \mathbf{g}(\hat{\mathbf{x}})$ $\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{n}$

Laboratory for Control and Automation

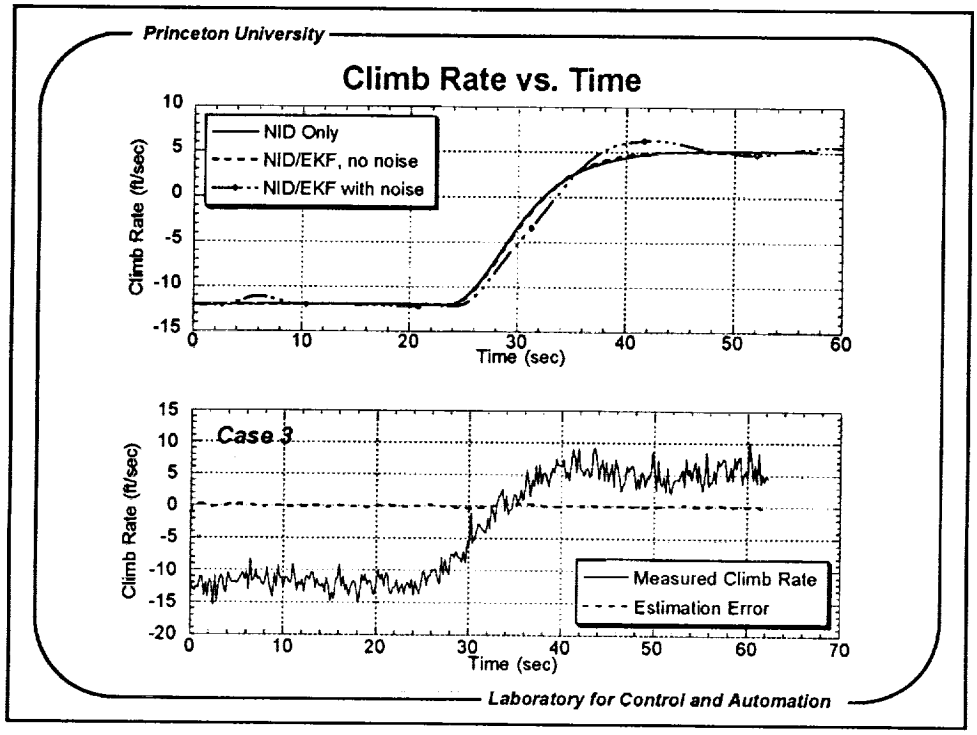
A set of nine measurements were postulated for the simulation examples. The assumed sensors were altitude, groundspeed, airspeed, angle of attack, pitch attitude and rate, climb rate, and horizontal and vertical acceleration. The sensor noise statistics were based on conservative estimates of the expected accuracy of those sensors. Three simulations were conducted using the same initial conditions and microburst wind profile. The simulations were structured in such a way to illustrate the degradation in controller performance caused by removing the assumption of perfect state feedback. In the first case, the NID controller was driven by perfect state feedback. In the second and third cases, the controller was driven by the output of an EKF that utilized the measurement vector shown. The difference between Cases 2 and 3 was that in Case 2, there was no noise in the sensor measurements. The performance realized here would thus be indicative of the theoretical limit of the performance of the NID/EKF combination. In Case 3, the measurements were noisy and had the statistics indicated by the matrix $\mathbf{R}_o$.



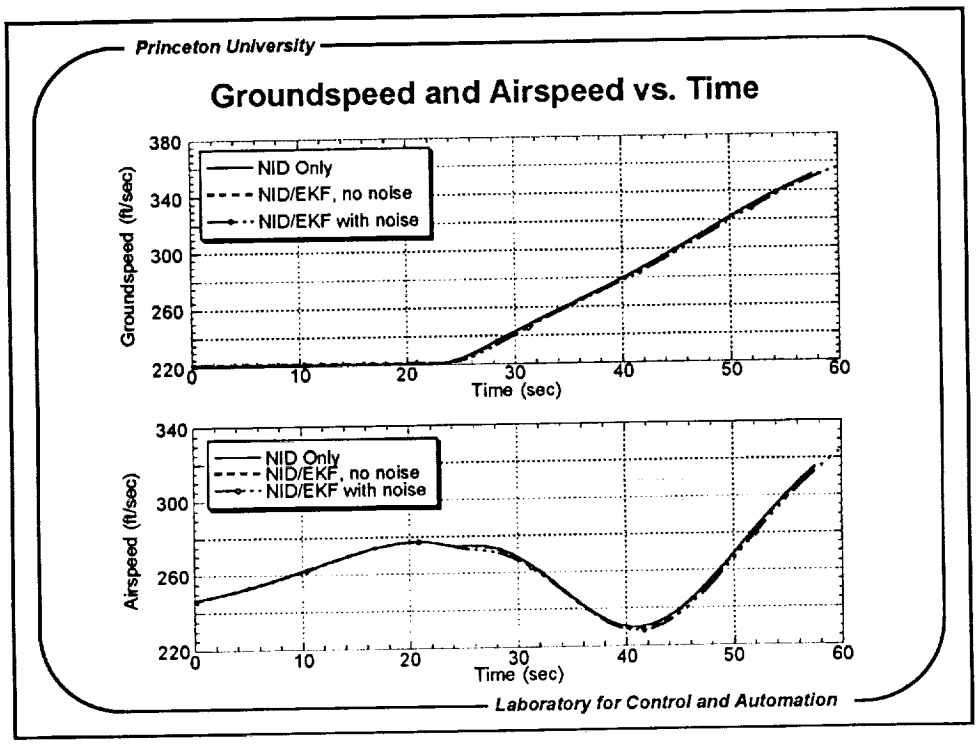
In all of the simulations conducted, the same initial conditions and microburst parameters were used. The aircraft was placed in an approach configuration a fixed distance away from the microburst core. The aircraft tracked the glide slope until the F-Factor exceeded a preset threshold, at which point a recovery was commanded using full thrust and a nominal climb rate of 5 ft/sec. For Cases 2 and 3 where the EKF was in use, the recovery was triggered on the basis of an estimate of the F-Factor.



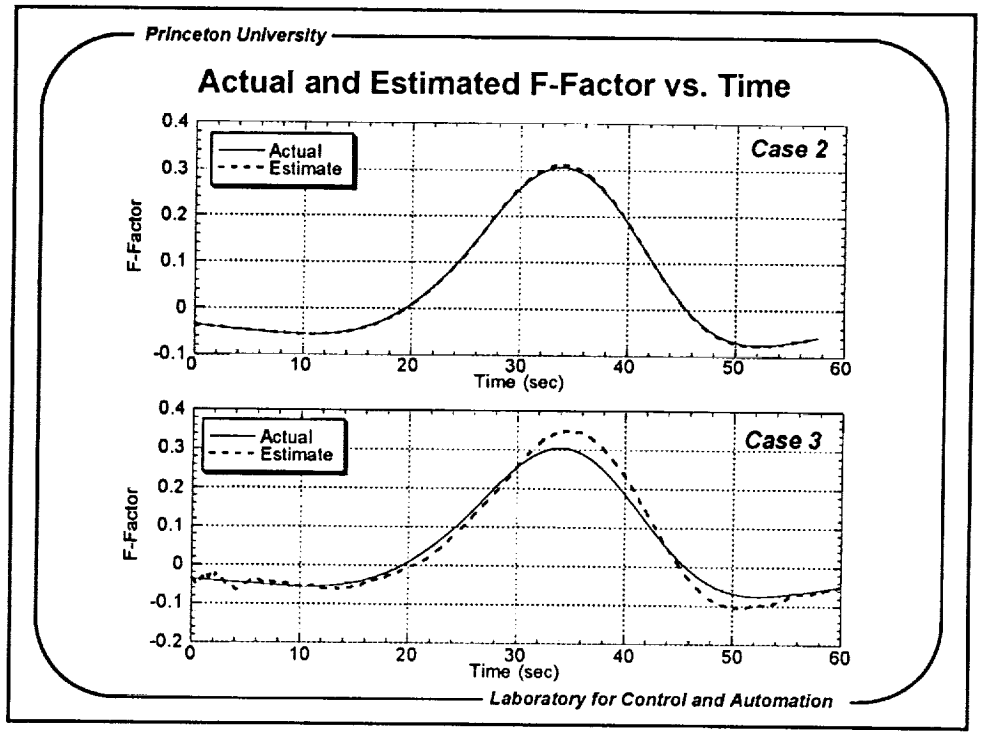
The altitude time-histories are shown here for the three simulation examples. It is apparent that there is little to no controller performance degradation between Cases 1 and 2. This suggests that in the limit as aircraft sensors become more and more accurate, the baseline performance realized using perfect state feedback can be achieved. There is only a slight loss in performance in Case 3, where the controller is driven by state estimates derived from noisy measurements.



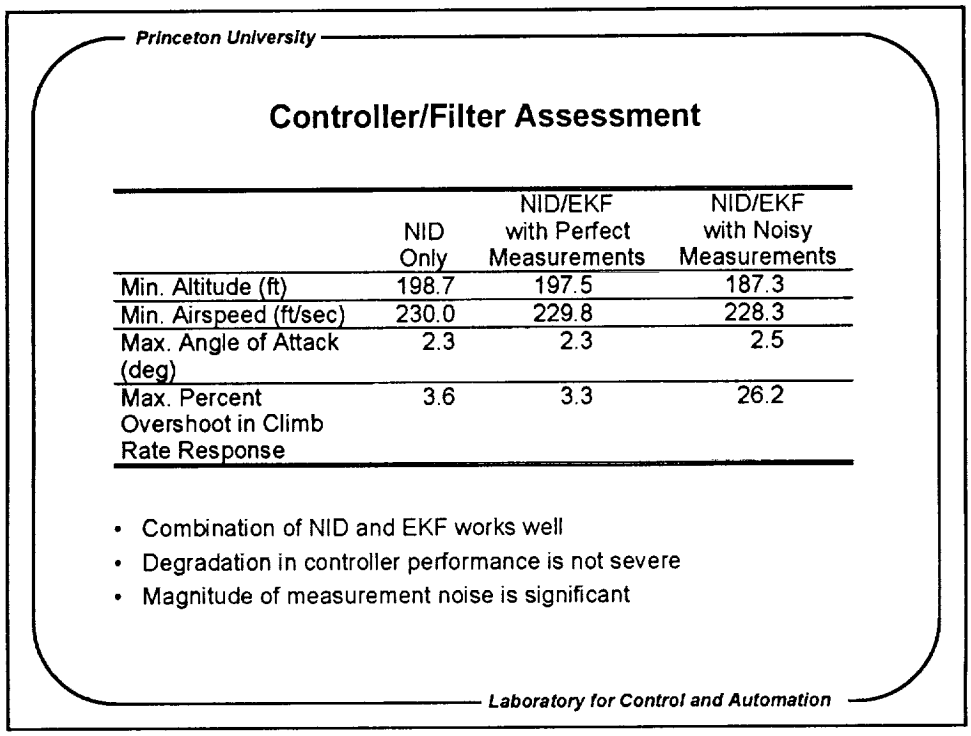
The climb rate histories for the 3 cases are shown in the top figure. Those from Cases 1 and 2 are quite similar to one another. In Case 3, there is more overshoot in the response of the aircraft. The performance of the EKF is indicated in the bottom figure. The output of the climb rate sensor is shown for Case 3, together with the resultant estimation error in climb rate. The magnitude of the estimation error is much smaller than the apparent level of noise in the sensor output. This indicates that the EKF is effective in eliminating the effects of measurement noise in the estimation of climb rate.



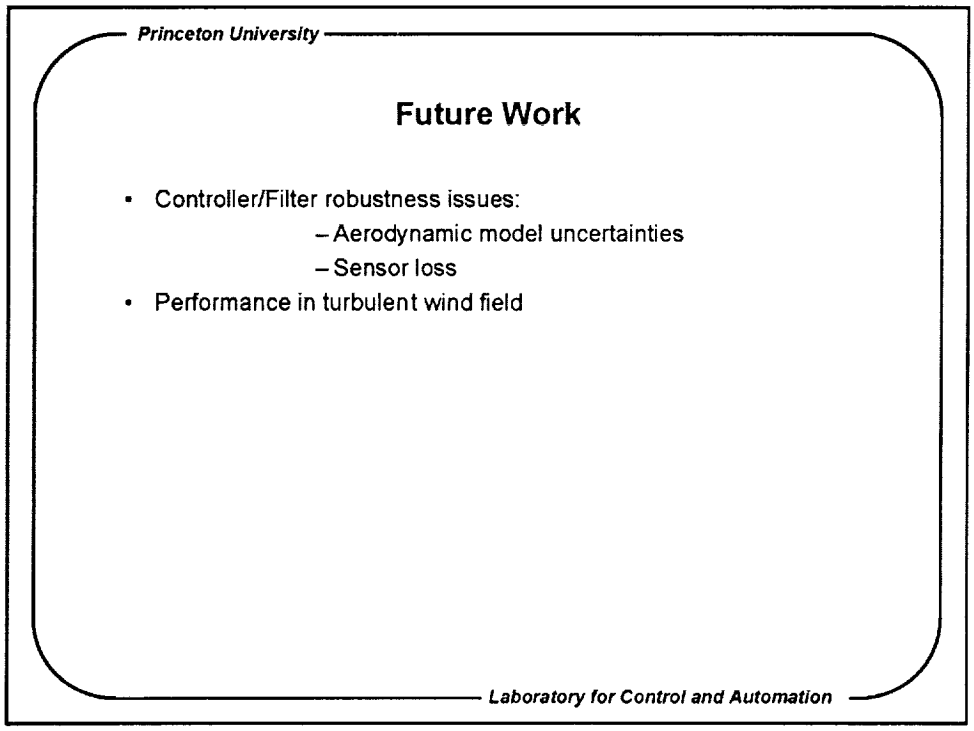
Groundspeed and airspeed climb rate histories are shown here for all three cases. They are virtually identical to one another. In the approach portion of the trajectory, the aircraft is able to track the reference groundspeed extremely well even when driven by optimal estimates derived from noisy measurements of groundspeed.



The ability of the EKF to estimate the F-Factor hazard index is illustrated here for Cases 2 and 3. In Case 2 where the EKF uses perfect measurements, the F-Factor is estimated very accurately. When noisy measurements are introduced in Case 3, some estimation lag becomes noticeable in the EKF output. The F-Factor estimates seem to lag the most when the sign of the F-Factor's time-derivative changes sign. The peak F-Factor is actually overpredicted by the EKF.



A summary is provided here of some salient features of each of the three cases. The difference in minimum altitude between Cases 1 and 3 is only 10.4 ft. The minimum airspeed is only 2 knots lower in Case 3 as compared to Case 1. This would suggest that in terms of maintaining safety margins, the EKF/NID combination is almost as effective as the NID alone driven by perfect state feedback. There is almost no difference in maximum angle of attack as well. The principal difference between Cases 1 and 3 is in the climb rate response of the aircraft. In Case 3, there is much more response overshoot than in Case 1. This is likely due to filter lags arising from uncertainty in the accuracy of the measurement vector.



There are a number of unresolved issues to be addressed in this work. The robustness of the NID/EKF combination to aerodynamic modelling errors will be studied. The system performance with a reduced sensor suite will also be investigated. The ability of the controller to track flight path command through a turbulent wind field will be investigated. It may be necessary to tune the EKF parameters to reduce unwanted control activity in wind fields containing high-frequency components.

AIR TRAFFIC MANAGEMENT AS PRINCIPLED NEGOTIATION BETWEEN INTELLIGENT AGENTS

J P Wangermann
Department of Mechanical and Aerospace Engineering
Princeton University

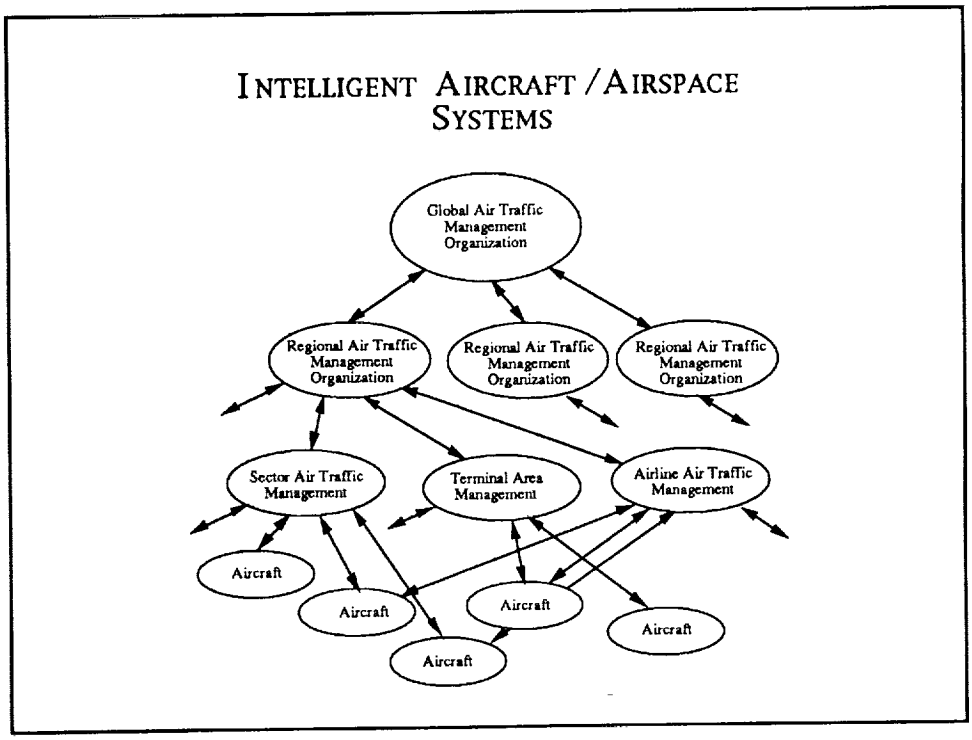
The major challenge facing the world's aircraft/airspace system (AAS) today is the need to provide increased capacity, whilst reducing delays, increasing the efficiency of flight operations, and improving safety. Technologies are emerging that should improve the performance of the system, but which could also introduce uncertainty, disputes, and inefficiency if not properly implemented.

The aim of our research is to apply techniques from intelligent control theory and decision-making theory to define an Intelligent Aircraft/Airspace System (IAAS) for the year 2025. The IAAS would make effective use of the technical capabilities of all parts of the system to meet the demand for increased capacity with improved performance.

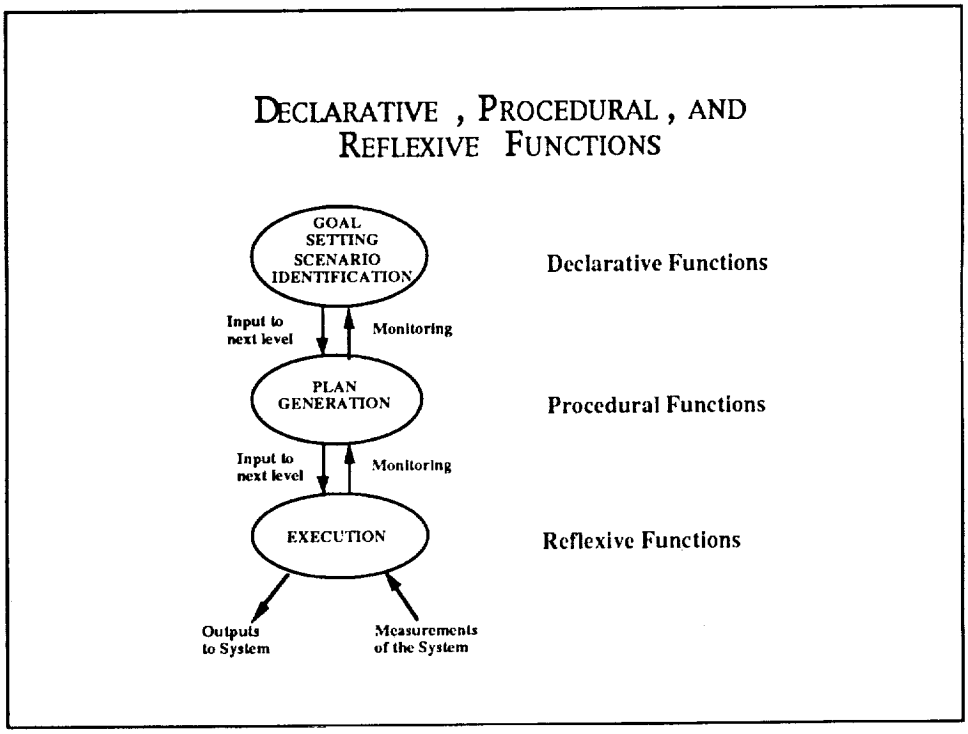
An Intelligent Aircraft Airspace System (IAAS) would allow each of these agents to interact in a way that:

- makes full use of the differing capabilities of all the agents
- allows each agent to obtain data residing in other parts of the system
- imposes as few restrictions as possible on aircraft operations in order to meet system performance requirements
- provides system robustness through dissimilar redundancy
- allows graceful degradation of system performance if any part should fail.

The Aircraft Airspace System consists of a variety of agents, operating in a broadly hierarchical structure. At the lowest level are the individual aircraft, from general aviation to commercial traffic; at the highest level are global organizations such as ICAO. At intermediate levels not only are there the various parts of today's air traffic management system, such as sector air traffic management (ARTCCs in the US), but also the airlines who already cooperate with flow control, and provide an increasingly important role in supporting aircraft in flight.



Each agent in the system is itself intelligent; it does more than execute instructions generated by the superior agent in the hierarchy. An Intelligent Agent performs a hierarchy of functions, bounded on one end by declarative functions, which typically involve decision-making, and on the other by reflexive functions, which are more-or-less spontaneous reactions to external or internal stimuli. An intermediate level, procedural functions, may also be defined. Like reflexive functions, these have a well-defined input-output characteristic, but have a more complicated structure.



This model of an intelligent agent can be used to describe any of the agents within the IAAS. Intelligent agent descriptions of a traffic control agent and a pilot/aircraft agent are given for illustration. The effect of emerging technologies will be to enhance the capabilities of the agents in all these functions. This will increase the overlap in capabilities of the agents.

As an example, collision avoidance systems (CAS) provide the pilot/aircraft agent with traffic situation data, previously only available to traffic control agents. These systems should provide increased safety, but have also on occasions caused conflict, when the CAS has issued instructions that conflicted with what the traffic control agent had planned.

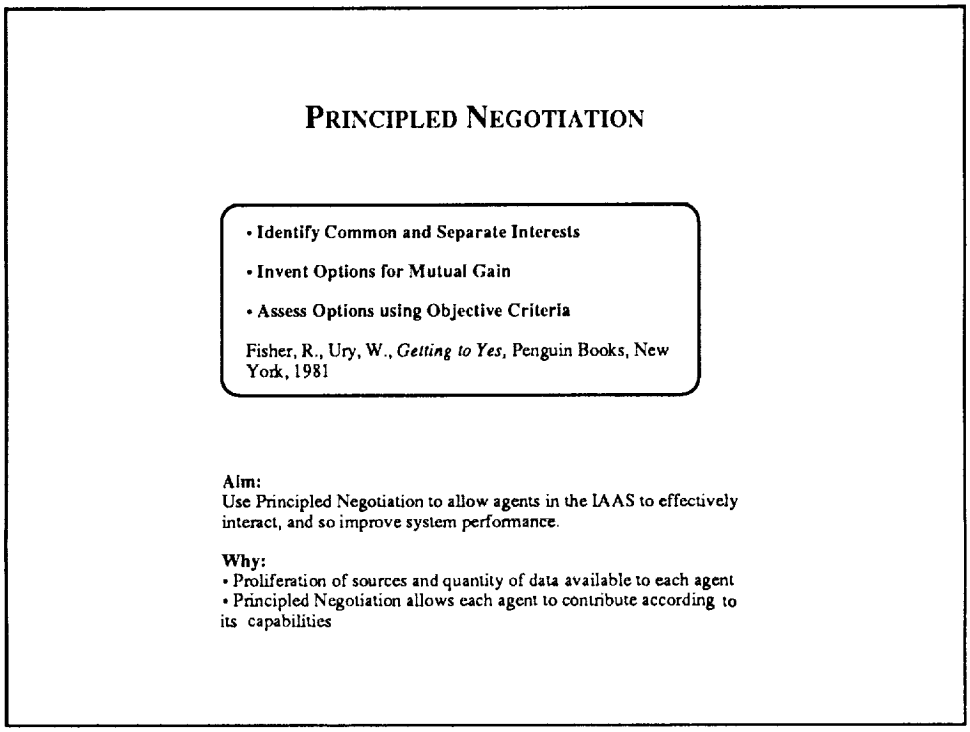
The IAAS must be able to overcome these types of potential problems, while exploiting the possibilities provided by the enhanced and overlapping capabilities of the agents.

FUNCTIONS OF INTELLIGENT AGENTS IN IAAS

TRAFFIC CONTROL AGENT	PILOT/AIRCRAFT AGENT
Declarative Functions	Declarative Functions
Sector allocation	System monitoring
Traffic monitoring	Goal planning
Conflict detection/prediction	System/scenario identification
Constraint monitoring	Choice of operating mode
Hazard detection	
Route assignment	Procedural Functions
	Adaptation
Procedural Functions	Guidance
Conflict resolution	Navigation
Flight path adaptation	Crew coordination
Networking	Networking
Assessment of pilot requests	
Flow control	Reflexive Functions
	Measurement
Reflexive Functions	State Estimation
Display update	Control
Communications	Communication
State vector processing	
Aircraft handover	

Each agent, either through its own sensors or through communications, will have the data and the computational ability to carry out strategic functions such as flight path modification, taking into account the interests of other agents as well as its own. By Inventing Options for Mutual Gain, and Assessing Options using Objective Criteria, agreements should be reached that benefit both parties. If more of the agents' interests are satisfied, the system is performing better.

Principled Negotiation is proposed as the structure for interaction of agents in the IAAS. Air traffic management can be viewed as a negotiation process; as the agents interact each is trying to best satisfy their own interests. Principled Negotiation exploits the fact that two parties in negotiation will have common interests on which an agreement that benefits both parties can be reached. Each agent in the IAAS has a different set of interests, but many interests are held in common.



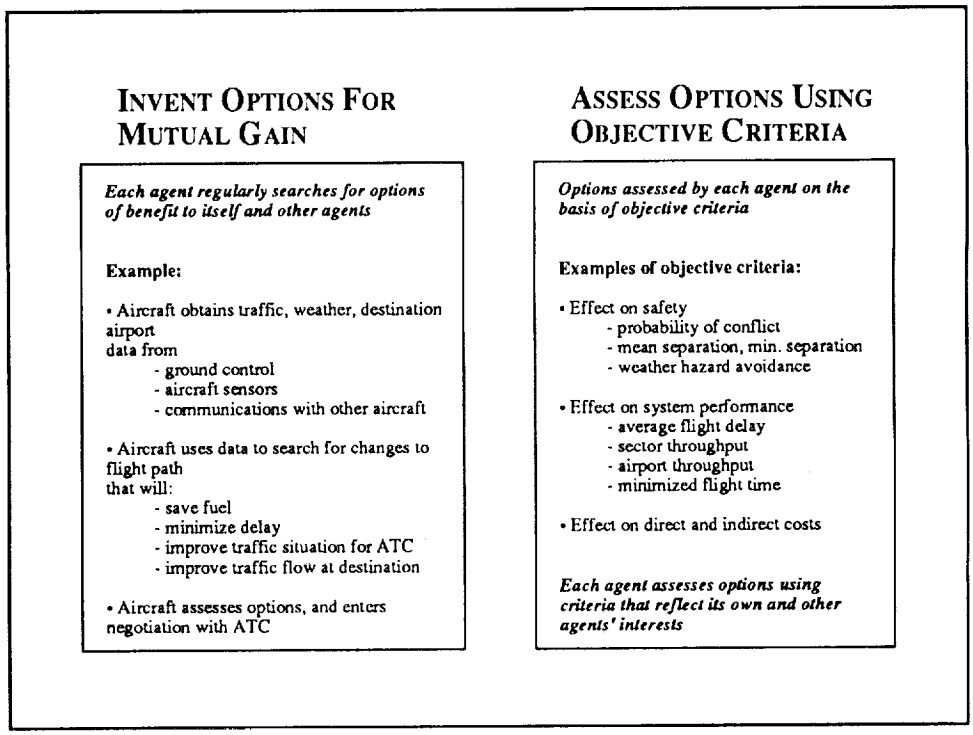
Negotiation is a viable model for cooperative decision-making in the IAAS, because of the large areas of common interest between the agents. Given a set of alternative decisions, two agents may often regard different decisions as optimal, because each agent will weight each factor differently. However it should be possible for the two agents to identify a single decision that, though not ideal for both agents, does better meet the interests of both agents than the status quo. Principled Negotiation provides a method by which this beneficial agreement can be achieved effectively.

EXAMPLES OF COMMON AND SEPARATE INTERESTS

	Pilot/ Aircraft	En-route Controller	Airline	Airport Operator
Safety	✓	✓	✓	✓
Fuel Costs	✓		✓	(✓)
Delays	✓	✓	✓	
Profit	(✓)		✓	✓
Throughput		✓	(✓)	✓
Scheduling Freedom			✓	

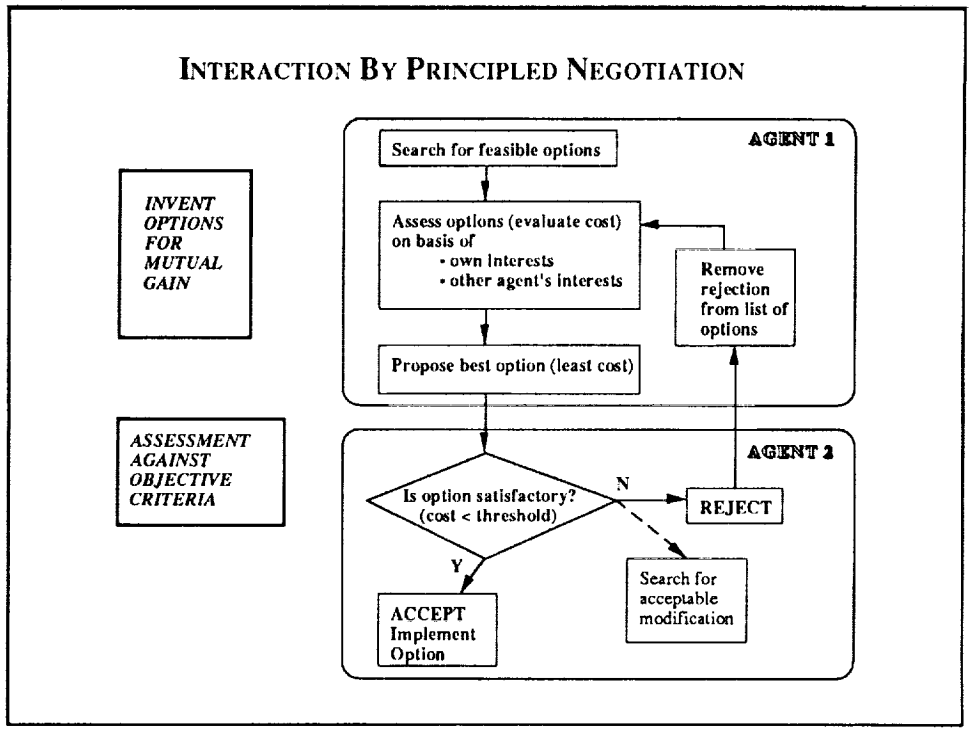
Each agent regularly searches for Options for Mutual Gain. It should consider the interests of the other agents in the system, not just its own. If a pilot/aircraft agent is searching for possible improvements to its flight path, it will be able to draw on data that describes the local traffic and weather situation, and may well have access to data on sectors further into the flight path, as well as the predicted situation at the destination airport at its planned arrival time. In assessing various options it should consider not only its own interests (fuel usage, time of arrival etc.), but also the interests of other agents. Does the option reduce traffic in an overloaded sector? Would arrival at the airport at a different time reduce a predicted peak in runway demand?

Once an agent has generated an option that provided mutual benefit, it would propose the option to the superior agent in the hierarchy. In the case of a proposed change in the flight path this would be the traffic control agent. The superior agent should assess any proposal using objective criteria. In the IAAS, objective assessment of a proposed flight path change would involve not only local analysis, but assessment of the impact of the change over as long a time scale and as wide a geographical area as possible.



The assessment of flight path changes is just one example of the many tasks that are undertaken in the IAAS. Most of these tasks involve the interaction of two or more agents, and Principled Negotiation should be applicable in all cases. These could be tasks occurring over time scales of months or years (such as airport slot allocation, or flight scheduling) or over short time scales (such as scheduling inbound streams of traffic in a terminal area).

This slide shows an algorithm that could be applied in any of these cases. Agent 1 would regularly conduct a search for options that provided mutual benefit. That benefit would probably be on the basis of a cost function that reflected the interests of itself and other agents. The best option would then be proposed to agent 2 (the superior agent). Agent 2 would make its own evaluation of the cost of the option, using its own data and possibly a different cost function. Different criteria could be used for accepting a proposal; one might be to accept a proposal if its cost was below a certain threshold. If the option was unacceptable, agent 2 might propose a modification to agent 1, or agent 1 might suggest an alternative.



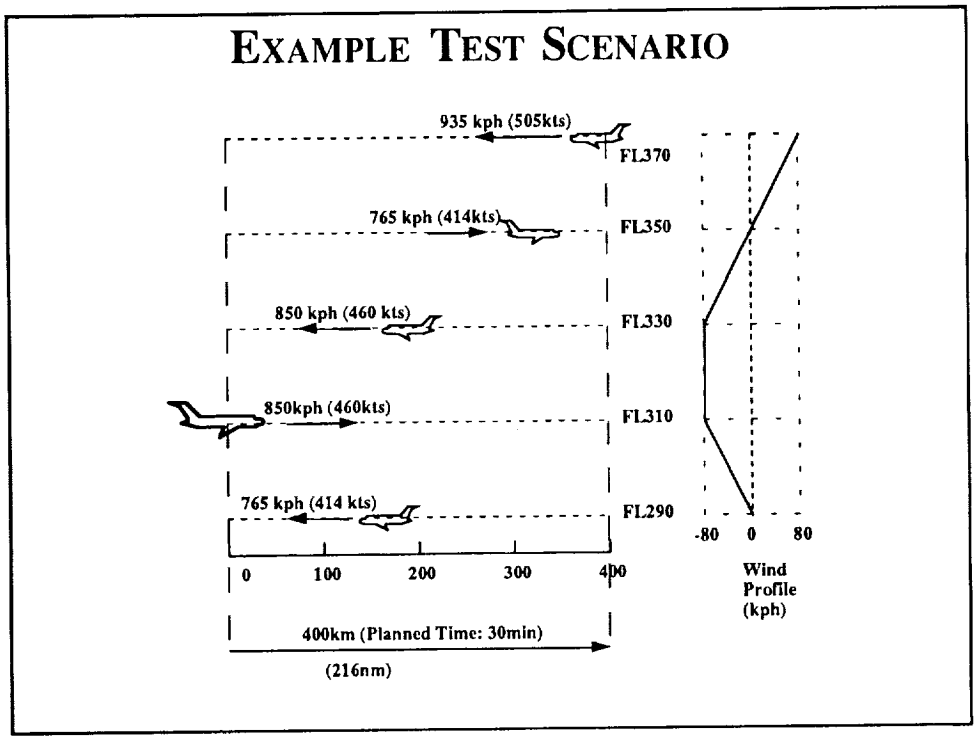
Current research is focussed on applying these ideas to a test scenario, and evaluating the concept. The initial test scenario is a 2D high level (FL290 - 370) sector. Although superficially a simple scenario, it provides a rich set of variables which can be analyzed. Some examples of effects which can be studied are:

- effect of different agent cost functions
- effect of conflicting aircraft data
- effect of wind distributions and other weather phenomena
- effect of different negotiation algorithms.

The decision-making system can be tested on various traffic distributions, and the effectiveness of the system analyzed in terms of:

- safety
- efficiency of operations
- capacity of the system
- punctuality (accuracy of aircraft at 4D waypoints)

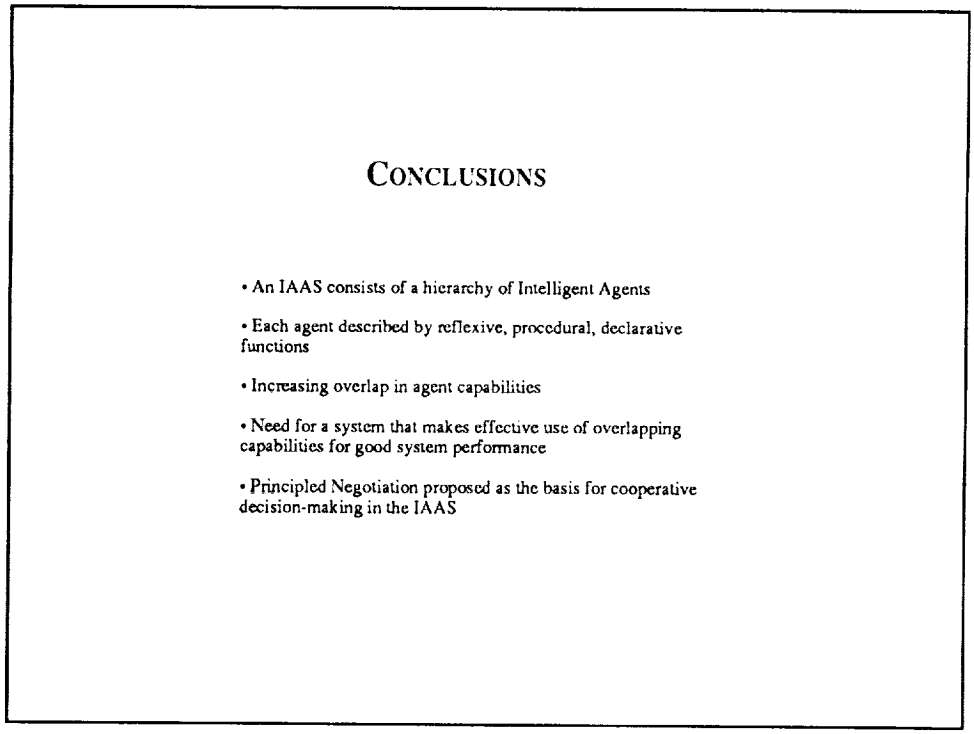
This scenario mainly deals with pilot/aircraft - traffic control agent interactions. The scenario can easily be made more complex , and eventually it is hoped to examine the possibilities of such a system in terminal airspace.



In summary, the capabilities of the agents in a future AAS will overlap to a much greater degree than at present. As each agent becomes increasingly intelligent, the declarative functions of the agents will have increasing commonality. The key to improved performance of a future AAS will be the effective use of these overlapping capabilities.

Overlapping capabilities can provide increased redundancy and flexibility for AAS operations, and effective combination of these overlapping yet distinct capabilities should give an IAAS improved performance compared to today's system. Principled Negotiation is proposed as a form of agent interaction that allows each agent to use its capabilities to ensure that decisions taken better meet each agents interests, and so improve system performance.

Work validating the concept in a 2-D en-route traffic scenario is progressing.



OPTICAL COMMUNICATIONS FOR TRANSPORT AIRCRAFT

ROBERT STENGEL
Department of Mechanical and Aerospace Engineering
Princeton University

THE PROBLEM

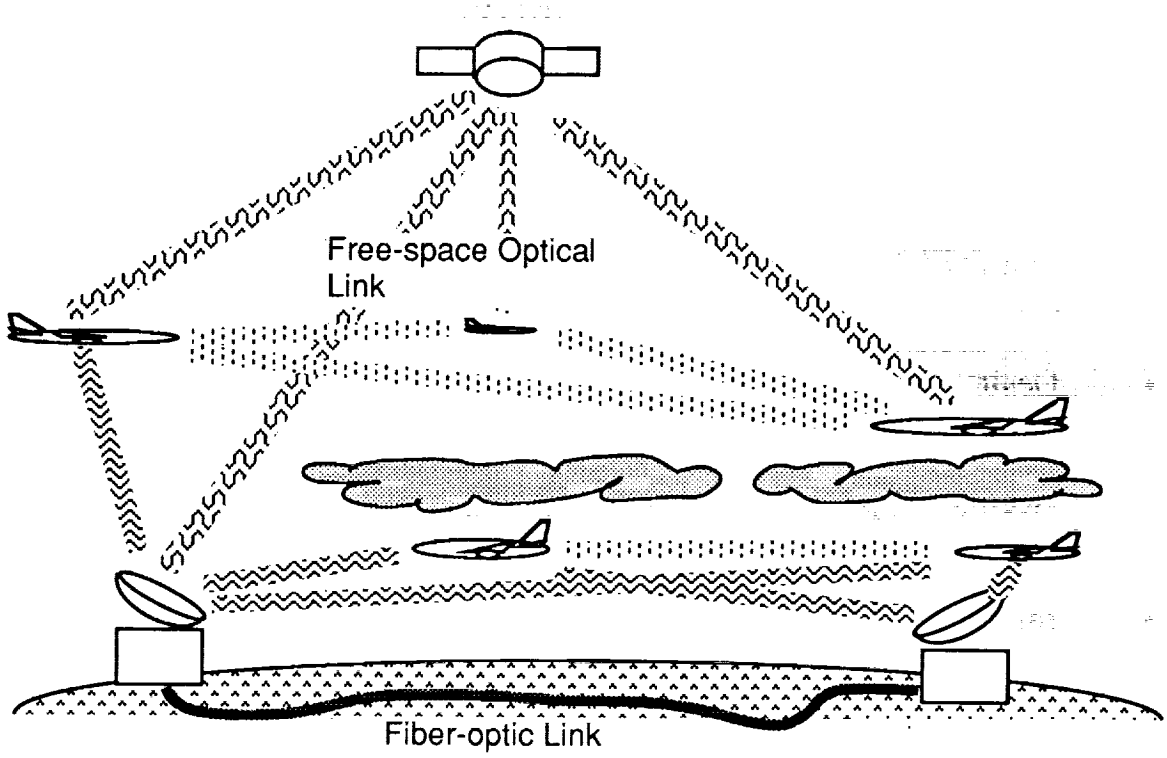
- Increasing demand for radio-frequency bands from an enlarging pool of users (aircraft, ground and sea vehicles, fleet operators, traffic control centers, commercial radio and television)
- Desirability of providing high-bandwidth, dedicated communications to and from every aircraft in the National Airspace System
- Need to support communications, navigation, and surveillance for a growing number of aircraft
- Improved meteorological observations by use of probe aircraft

THE SOLUTION

- Optical signal transmission support very high data rates
- Optical transmission of signals between aircraft, orbiting satellites, and ground stations, where unobstructed line-of-sight is available
- Conventional radio transmission of signals between aircraft and ground stations, where optical line-of-sight is unavailable
- Radio priority given to aircraft in weather

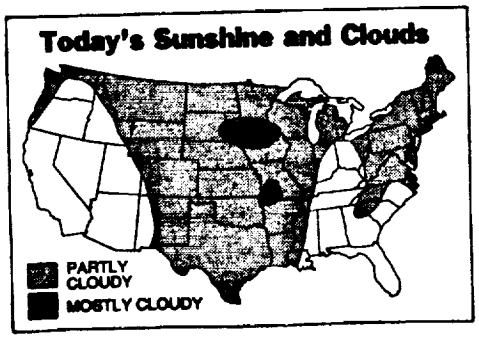
AEROSPACE OPTICAL COMMUNICATION

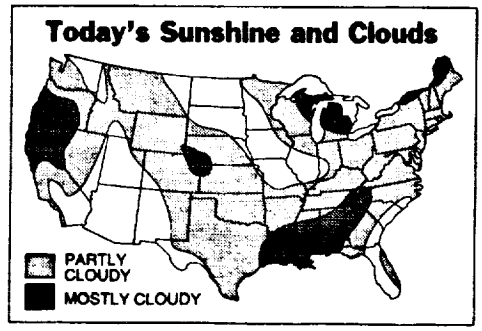
Data communications between aircraft and ground stations could be supported with direct and relayed signals. Aircraft at altitude typically would have unobstructed line of sight to an overhead spacecraft and frequently could communicate with other aircraft at similar altitude. Fiber-optic links on the ground complete the data path for air-ground links obscured by clouds through unobscured air-satellite-ground links.

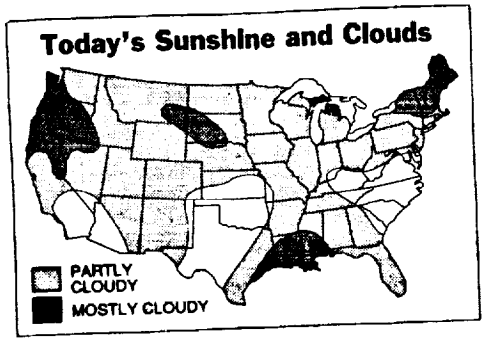


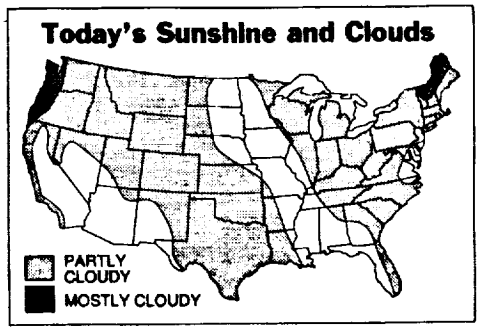
- Opportunistic Optical Transmission
- Distributed Network containing Free-Space and Fiber-Optic Links
- Radio Transmission where Optical Link is Unavailable

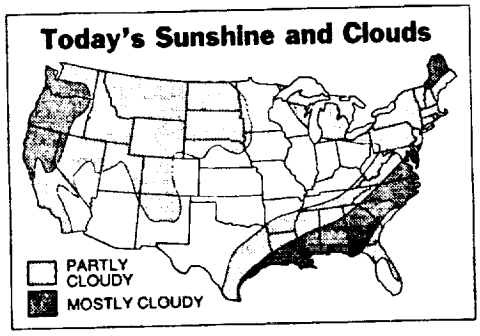
Typical Cloud Cover Patterns Across the United States

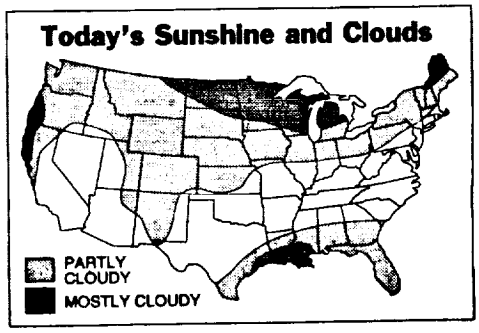


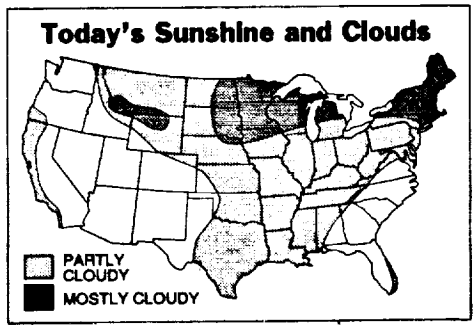












EVERY AIRCRAFT A WEATHER PROBE AND AIRBORNE SURVEILLANCE SYSTEM

Increased data bandwidth allows greatly expanded transfer of information about weather conditions and individual aircraft. Observational data from aircraft is integrated into a real-time four-dimensional weather map in ground-based computers. This information, in turn, becomes available to all aircraft in the system.

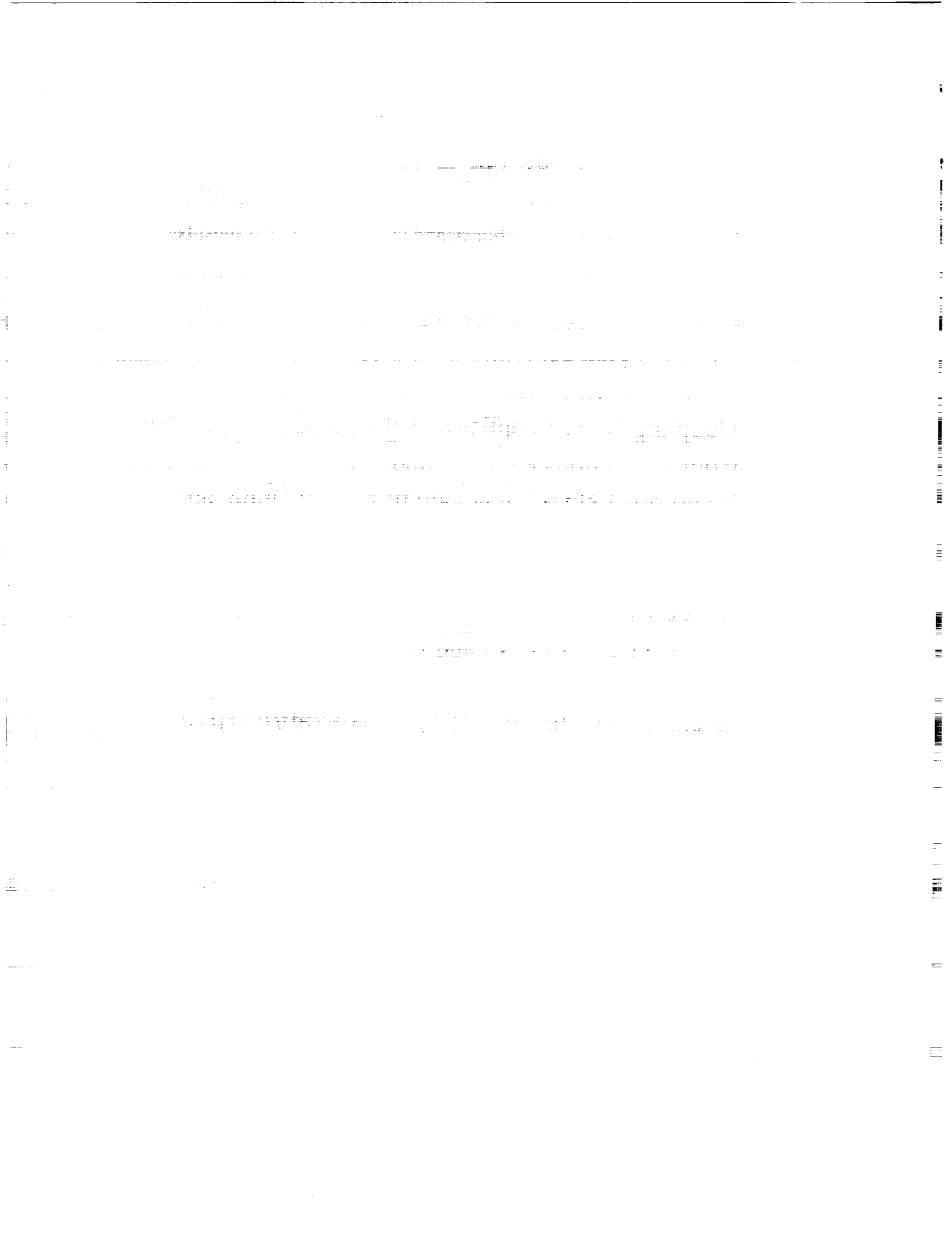
- DOWNLINK
  Own position and velocity vectors
  Own air temperature, pressure, and humidity
  Own wind velocity vector
  Own light intensity
  Own turbulence intensity
  Signal strengths from electrical activity and beacons
  Airborne hazard status monitoring and alerts
  Desired alternate flight plans

- UPLINK
  Air temperature, pressure, and humidity fields
  Wind and turbulence fields
  Cloud cover
  Traffic alerts
  Ground/satellite-based hazard status monitoring and alerts
  Arbitrated alternate flight plans

RESEARCH ISSUES

Numerous technical, operational, and institutional issues must be resolved before the suitability of optical communications for aircraft can be fully assessed. Many of these are topics for basic and applied research.

- Optical signal generation, transmission, and detection
- Coherence, filtering, power, multiplexing, and coding
- Coupling between optics and electronics
- Communication coverage modeling
- Telescope field of view, pointing, acquisition, and tracking
- Free-space/fiber-optic networking and data-relay protocols
- Architectures for CNS and ATM
- Interfaces with related systems
- Integration within an Intelligent Aircraft/Airspace System



Robustness of Solutions to a Benchmark Control Problem

Robert F. Stengel* and Christopher I. Marrison†
Princeton University, Princeton, New Jersey 08544

The robustness of 10 solutions to a benchmark control design problem presented at the 1990 American Control Conference has been evaluated. The 10 controllers have second- to eighth-order transfer functions and have been designed using several different methods, including H_∞ optimization, loop-transfer recovery, imaginary-axis shifting, constrained optimization, structured covariance, game theory, and the internal model principle. Stochastic robustness analysis quantifies the controllers' stability and performance robustness with structured uncertainties in up to six system parameters. The analysis provides insights into system response that are not readily derived from other robustness criteria and provides a common ground for judging controllers produced by alternative methods. One important conclusion is that gain and phase margins are not reliable indicators of the probability of instability. Furthermore, parameter variations actually may improve the likelihood of achieving selected performance metrics, as demonstrated by results for the probability of settling-time exceedance.

Introduction

CONTROL systems should be designed to maintain satisfactory stability and performance characteristics not only at nominal operating points but over a range of parameters that encompasses system uncertainty. These systems should be robust, but there is a limit. Unbounded robustness is no more attractive than inadequate robustness, because nominal performance and insensitivity to parameter variations tend to produce conflicting design requirements. Hence, the degree of robustness that must be furnished for satisfactory operation is related to the system variations that are most likely to occur.

Measures of robustness should be easily understood and should be directly connected to control design objectives. They should be consistent with what is known about the structure and parameters of the plant's dynamic model. These goals are best served when robustness is expressed in terms of the likelihood that commonly accepted properties fall within acceptable bounds and when parameter variations are expressed in terms of readily measured system specifications. A method of satisfying these evaluation criteria is presented here.

This paper demonstrates the application of stochastic robustness analysis (i.e., determining the probability of unsatisfactory stability or performance resulting from expected parameter uncertainty) to solutions of the 1990 American Control Conference Benchmark Control Problem.[1] Stochastic robustness is seen to provide a useful, unifying analytical framework that is intuitive and has a direct, physical meaning.

Description of the Problem

The benchmark plant is a dual-mass/single-spring system with noncollocated sensor and actuator[1]; its state-space model is

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\end{bmatrix} = \begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -k/m_1 & k/m_1 & 0 & 0\\ k/m_2 & -k/m_2 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix} + \begin{bmatrix}0\\ 0\\ 1/m_1\\ 0\end{bmatrix}u + \begin{bmatrix}0\\ 0\\ 0\\ 1/m_2\end{bmatrix}w \tag{1}$$

$$y = x_2 + v \tag{2}$$

$$z = x_2 \tag{3}$$

where x_1 and x_2 are the positions of the masses, x_3 and x_4 are their velocities, and u is a control force on m_1. The plant is disturbed by w on m_2, and the measurement of x_2 is corrupted by noise v in y. The corresponding actuator and disturbance input/output transfer functions are

$$\mathcal{TC}_{uy} = \frac{(k/m_1 m_2)}{s^2\left[s^2 + k(m_1+m_2)/m_1 m_2\right]} \tag{4}$$

$$\mathcal{TC}_{wy} = \frac{(1/m_2)(s^2 + k/m_1)}{s^2\left[s^2 + k(m_1+m_2)/m_1 m_2\right]} \tag{5}$$

The plant has eigenvalues at $(\pm j\sqrt{k(m_1+m_2)/m_1 m_2}, 0, 0)$ and is undamped. A single-input/single-output (SISO) controller must close its loop around $\mathcal{TC}_{uy}$, which has a pole-zero surplus of 4. The high-gain asymptote of at least one root lies in the right half plane for any SISO feedback compensator that has fewer than two surplus zeros. Because the open-loop roots are on the imaginary axis, the magnitudes of root departure angles must exceed 90 deg if marginal instability is to be avoided at low loop gain.

Three design problems are posed in Ref. 1. Benchmark problem 1 (BP-1) requires 1) a 15-s settling time for unit disturbance impulse and nominal mass-spring values ($m_1 = m_2 = k = 1$) and 2) closed-loop stability for fixed values of mass and $0.5 < k < 2$. The second problem, BP-2, replaces the unit-impulse disturbance by a sinusoidal disturbance with 0.5-rad/s frequency but unknown amplitude and phase. Asymptotic rejection of the signal should be achieved with a 20-s settling time for nominal masses and $0.5 < k < 2$. The third problem, BP-3, is like BP-1, except that m_1, m_2, and k are uncertain with mean values of 1 and unspecified bounds. A number of additional problem specifications are left to the discretion of the designer. For example, it is presumed that a noise model

Received May 14, 1991; revision received Oct. 4, 1991; accepted for publication Oct. 11, 1991.

*Professor, Department of Mechanical and Aerospace Engineering. Associate Fellow AIAA.

†Graduate Student, Department of Mechanical and Aerospace Engineering.

Research supported by government grant.

$v(t)$ would be considered, but details of the model are open. Subjective goals include achieving reasonable performance/stability robustness, minimizing controller effort, and minimizing controller complexity.

Design Solutions and Nominal Performance

Five papers containing design solutions appear in the American Control Conference Proceedings,[2-6] one paper became available after the conference,[7] and additional designs were obtained from the authors. The transfer functions for these controllers are presented in the Appendix. Fixed-order compensators achieving approximate loop-transfer recovery are motivated in Ref. 2, leading to designs A-C. An H_∞ plus $j\omega$-axis shifting approach is taken in Ref. 3, producing design D. Reference 4 uses nonlinear constrained optimization to produce design E. Structured covariance terms are added to the linear quadratic Gaussian (LQG) algebraic Riccati equations to generate design F in Ref. 5. Design G is a game-theoretic controller based on linear exponential Gaussian and H_∞ concepts and is discussed in Ref. 6. H_∞ controllers using the internal model principle are presented in Ref. 7 (designs H-J). G and J are designed to reject the sinusoidal disturbance (BP-2) rather than the unit impulse disturbance (BP-1). All but two of these designs (A and D) contain non-minimum-phase zeros. The benchmark criteria do not address command-input responses; hence, the initially reversed time response of systems with an odd number of non-minimum-phase zeros is not penalized. Design G has an even number of right-half-plane zeros, which would not produce reversed response.

The problem statement contains an ambiguity that could have affected the designers' interpretations of satisfactory response. Settling time is normally defined as an attribute of unit-step-function response. For example, Ogata[8] states that "The settling time is the time required for the response curve to reach and stay within a range about the final value of size specified by absolute percentage of the final value (usually 2% or 5%)." For a second-order system the 2% settling time can be precisely calculated as $4/\zeta\omega_n$, where ζ is the damping ratio and ω_n is the natural frequency of the oscillatory mode. However, Takahashi et al.[9] found that "Exact analytical expressions for ... settling time become prohibitively complicated for systems of order higher than two." The benchmark ambiguity is that the final value of a strictly stable impulse response (BP-1) is zero; hence, there is no steady-state value on which to base percentage response.

Nominal performance characteristics of the controllers are summarized in Table 1, which presents compensator numerator and denominator order (Num Ord and Den Ord), two definitions of settling time (T_S^* and T_S^{**}), maximum control usage (u_{max}) resulting from a unit w disturbance, gain margin (GM), phase margin (PM), output response to 0.5/rad/s sinusoidal disturbance (SR), and covariance of control response (U_{cov}) to measurement noise (v) with unit standard deviation. All compensators are proper (the number of zeros does not exceed the number of poles), but three (C, D, and E) are not strictly proper (the number of zeros equals the number of poles). Hence, designs A, B, and F-J can be classified as low-pass filters, whereas designs C-E do not roll off at high frequencies.

T_S^* portrays the settling time as the time for which x_2 is captured within a 0.1-unit envelope about its zero steady-state value, given an initial unit w disturbance impulse. T_S^{**} is based on the damping ratio and natural frequency of the dominant mode and is calculated as $4/\zeta\omega_n$. Neither of these definitions adheres to the conventional definition, but each has its merits. T_S^* is consistent with the BP-1 problem specification, in that it reflects a response to a unit w disturbance; however, it is amplitude dependent. T_S^{**} is independent of amplitude, but it is unrelated to the disturbance input and is not an accurate portrayal of the full system's settling time in response to a unit step input. Table 1 indicates that only three of the compensators satisfy a 15-s criterion by the first definition, whereas six compensators have settling times of ≤ 15.2 s by the second definition.

Four compensators use measurably more control than the others in responding to the disturbance. Increasing gain margin generally is accompanied by increasing phase margin for these 10 designs (Fig. 1), although the relationship is not monotonic. With the exception of design D, stability margins are less than the 8-dB/30-deg rules of thumb (e.g., Ref. 10) often suggested as design goals for SISO systems. Sinusoidal disturbance rejection of most controllers is similar, although design D's response is an order of magnitude smaller. Designs G and J, specifically intended to reject a 0.5-rad/s sinusoid, eliminate the disturbance completely in the steady state. (The settling time in achieving this response was not evaluated.) The noise-response covariance of the control is generally proportional to its peak disturbance-impulse response for strictly proper compensators. The three non-strictly proper compensators have infinite control covariance in response to continuous white measurement noise v (with infinite bandwidth).

Stochastic Robustness Analysis

Stochastic robustness analysis (SRA) is based on a statistical portrayal of parameter variations and their effects. If parameters take a finite number of discrete values, each with known

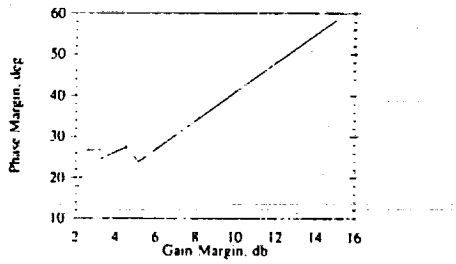


Fig. 1 Nominal gain and phase relationships of the 10 controllers.

Table 1 Nominal characteristics of 10 controllers

Design	Num Ord	Den Ord	T_S, s[a]	T_S, s[b]	u_{max}	GM, db	PM, deg	SR, db	U_{cov}
A	2	3	21.0	14.8	0.514	2.56	26.7	10.1	6.30
B	2	3	19.5	15.2	0.469	3.27	26.8	13.2	13.02
C	2	2	19.7	15.2	0.468	3.27	26.5	13.3	∞
D	4	4	9.9	8.8	297.8	15.10	58.7	1.47	∞
E	2	2	18.2	8.01	0.884	2.39	22.0	17.1	∞
F	3	4	13.7	22.0	2.397	5.15	23.8	13.4	6×10^4
G	5	8	31.3	35.7	1.458	3.61	25.4	$-\infty$	173.5
H	3	4	14.9	11.9	0.574	3.28	24.5	14.9	2.48
I	3	4	17.8	17.2	0.416	4.56	27.5	13.3	0.95
J	5	6	43.2	23.8	1.047	2.14	17.5	$-\infty$	77.42

[a] *Defined for 0.1-unit response envelope for unit-impulse w.
[b] **Defined by $4/\zeta\omega_n$ (provided by B. Wie).

or estimated probability, the analysis can be based on a finite number of function evaluations, and the probabilistic result is exact (within the accuracy and precision of problem modeling). If the parameters are continuous or the number of finite combinations is too large for practical computation, Monte Carlo evaluation can be used to estimate probabilities within arbitrarily small confidence intervals. If a binary judgment can be made of function values (e.g., satisfactory/unsatisfactory or stable/unstable), then the corresponding probability distribution is binomial, and confidence intervals are readily estimated from the number of function evaluations (e.g., Ref. 11). Further details of SRA can be found in Refs. 12–17.

Test Cases for the Benchmark Problem

Uncertain parameters are assumed to have continuous, bounded, uniform, and uncorrelated probability distributions for this analysis. (The original problem identifies uncertain parameters and their bounds, making no statement about distributions.[1]) Three increasingly demanding sets of parameter uncertainties are used to test the controllers. The first two are specified in Ref. 1, and the third is new.

Problem E-1: $0.5 < k < 2$, all other parameters take nominal values, as in BP-1 and BP-2.

Problem E-2: $0.5 < k < 2$, $0.5 < m_1 < 1.5$, and $0.5 < m_2 < 1.5$, as in BP-3. Reference 1 does not specify limits on m_1 and m_2; values of $\pm 50\%$ are adopted here.

Problem E-3: Same as E-2; in addition, $0 < c < 0.1$, $0.9 < f < 1.1$, and $0.001 < \tau < 0.4$ s, where c represents internal damping between the masses; f is loop-gain uncertainty due to multiplicative variation in observation, control gain, or actuator effectiveness; and τ is the time constant for a first-order lag between controller command and actuator response. Uncertainty in the damping ratio c increases open-loop damping, and the time lag is always greater than the nominal value of zero.

With all six parameters, the state-space model for E-3 becomes

$$\dot{x}' = F'x' + G'u_c + L'w \tag{6}$$

where x' is defined as $[x_1\ x_2\ x_3\ x_4\ u]^T$, and

$$F' = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -k/m_1 & k/m_1 & -c/m_1 & c/m_1 & f/m_1 \\ k/m_2 & -k/m_2 & c/m_2 & -c/m_2 & 0 \\ 0 & 0 & 0 & 0 & -1/\tau \end{bmatrix} \tag{7}$$

$$G' = [0\ \ 0\ \ 0\ \ 0\ \ 1/\tau]^T \tag{8}$$

$$L' = [0\ \ 0\ \ 0\ \ 1/m_2\ \ 0]^T \tag{9}$$

The compensators are modeled by

$$\dot{x}_c = Ax_c + By \tag{10}$$

$$u_c = Cx_c + Dy \tag{11}$$

where x_c is the compensator state; u_c is the actuator command; A, B, C, and D are the compensator matrices; and y is x_2.

Performance Metrics for the Benchmark Problem

Robustness is best characterized by problem-dependent metrics that have a direct bearing on the measurable stability and performance of the system. Here, they portray the likelihood that classical stability bbunds will be exceeded, that settling time will not be achieved, and that control usage will exceed acceptable values. For demonstration of SRA, parameter uncertainties are represented by uniform distributions within arbitrary (but reasonable) bounds. In practical application, the control-system designer would have similar, problem-specific specifications to meet.

Each of the following probabilistic performance metrics has a binomial distribution and is estimated using Monte Carlo evaluation. Uniform, bounded parameters are calculated by random-number generators according to the specifications of the previous section. The associated binomial confidence level depends on the number of evaluations and the value of the probability estimate.[13] Each estimate is the result of 20,000 evaluations; for a probability estimate of 0.1, the 95% confidence interval would be ± 0.004. The performance metrics are:

1) P_I: Probability of instability. This probability portrays the likelihood that parameter variations force at least one closed-loop root into the right half plane.

2) P_{T_s}: Probability of settling-time exceedance. This probability is derived from a time-history calculation with a unit-impulse w input (i.e., based on T_S^*) and estimates the likelihood that the actual response of z will fall outside a ± 0.1-unit envelope after 15 s.

3) P_u: Probability of control limit exceedance. This probability corresponds to the requirement in Ref. 1 to minimize controller effort. It is the probability that peak actuator displacement will exceed a saturation limit in response to a unit disturbance (w) impulse. The saturation limit was chosen to be one unit for this analysis.

4) P_f: Probability of unsatisfactory sinusoidal disturbance rejection. This probability involves the likelihood that the amplitude of steady-state z response exceeds one unit with a unit sinusoidal disturbance at 0.5 rad/s.

Computation times indicate that current workstations are fast enough to execute practical SRA, and massively parallel computers could provide interactive turnaround. For the typical closed-loop system considered here, roughly 900 sets of eigenvalues were generated per minute per million floating-point operations per sec (MFLOP). This is drawn from compiled Pascal code executed on a 0.9-MFLOP Silicon Graphics 4D/20 workstation. The complete evaluation was computed at a rate of 30 sets/min/MFLOP using MATLAB on a Macintosh IIx computer. At these rates, a 5000-MFLOP parallel computer (e.g., 64K CM-2 Connection Machine) would evaluate 20,000 sets of eigenvalues in 0.25 s, and the full evaluation would take about three times longer.

Results of the Analysis

The results of the SRA indicate a wide range of characteristics in the 10 controllers. This reflects varying emphasis in satisfying the problem specifications, as well as significant differences in compensator order and design philosophy. It should be emphasized that none of the controllers was designed for the express purpose of satisfying SRA criteria, and it is likely that each design approach could be fine-tuned to produce better results than those shown here. Using criteria that have high engineering significance, SRA provides a "level playing field" on which to judge the robustness of controllers that were designed by alternative methods. Tables 2–4 present results, with maximum probabilities for each evaluation problem indicated by bold letters and minimum values underlined.

Probability of Instability

For the least uncertain case (E-1), over half of the controllers are estimated to have zero probability of instability, whereas design A has a 16% likelihood of instability (Table 2). With increasing parameter uncertainty (E-2 and E-3), all controllers have nonzero P_I. The probability of design A is essentially unchanged, and design J becomes the controller most likely to be unstable.

It is interesting to compare the probabilities of instability on the bases of gain and phase margins, quantities often assumed to indicate the robustness of SISO systems. Figures 2 and 3 demonstrate that nominal values of GM and PM are *not* good predictors of P_I. (Note that these bar charts present results for the 10 compensators; hence, GM and PM are not evenly dis-

Table 2 Probability of instability

Design	E-1	E-2	E-3
A	0.160	0.159	0.165
B	0.023	0.042	0.039
C	0.021	0.040	0.041
D	0.000	0.004	0.059
E	0.000	0.097	0.125
F	0.000	0.119	0.224
G	0.000	0.203	0.232
H	0.000	0.046	0.099
I	0.000	0.013	0.029
J	0.039	0.237	0.245

Table 3 Probability of settling-time violation

Design	E-1	E-2	E-3
A	0.971	0.962	0.793
B	1.000	0.969	0.963
C	1.000	0.968	0.874
D	0.000	0.004	0.072
E	1.000	1.000	0.999
F	0.633	0.859	0.967
G	1.000	0.999	1.000
H	0.742	0.909	0.986
I	0.756	0.918	0.986
J	1.000	1.000	0.968

Table 4 Probability of control-limit exceedance

Design	E-1	E-2	E-3
A	0.160	0.159	0.165
B	0.023	0.043	0.047
C	0.021	0.041	0.041
D	1.000	1.000	1.000
E	0.000	0.391	0.409
F	1.000	1.000	1.000
G	1.000	0.886	0.889
H	0.000	0.133	0.162
I	0.000	0.023	0.030
J	0.857	0.542	0.527

tributed.) In most cases, increasing parameter uncertainty increases P_I, but there are no consistent trends with GM and PM. Parameter variations have complex effects on the *shape* of each controller's Nyquist plot, and these effects cannot be portrayed simply by changing loop gain or phase angle.

This result brings into question the utility of transfer-function/return-difference-matrix singular values as measures of the stability robustness of multi-input/multi-output (MIMO) systems. MIMO singular-value analysis is loosely equivalent to SISO gain-margin analysis (e.g., Ref. 18). Arbitrary, real parameter variations have complicated effects on the frequency distributions of MIMO singular values, changing their shapes as well as their magnitudes. Unless the frequency distributions of *nominal* MIMO norms retain their shapes under parameter variation (or follow some predictable pattern), the relationships of nominal maximum or minimum values to allowable bounds tells little about stability robustness. Norm bounds can be reliably evaluated only by considering the norms of *perturbed* systems.

A higher compensation order does not necessarily improve robustness (Tables 1 and 2). The compensators with the most stability robustness are fourth order, and the next most robust controllers are second and third order. Increased nominal control usage, either as a consequence of a disturbance impulse or measurement noise, generally corresponds to decreased stability robustness, although design D provides a significant exception.

Probability of Settling-Time Violation

All but three of the controllers (D, F, and H) exceed the 15-s settling-time objective (defined by T_S^*) in the nominal case (Table 1); hence, it is not surprising that the probability of settling-time violation with parameter uncertainty is high as well (Table 3). Design D provides a notable exception: Its nominal T_S^* is 9.9 s, and P_{T_s} is small for all three evaluation cases. For problem E-1, half of the controllers violate the goal all the time, but two of the controllers with nominal T_S^* above 15 s (H and I) have a considerable likelihood (25%) of satisfying the objective when the spring-constant uncertainty is considered. Further uncertainty (problems E-2 and E-3) reduces the probability of settling-time violation for more controllers, illustrating the counterintuitive result that the effects of uncertainty are not always unfavorable.

Probability of Control Limit Exceedance

The probability of excessive control response to disturbance impulse P_u is shown in Table 4. Over half of the nominal responses are within the u_{max} criterion chosen for this analysis (Table 1). Furthermore, there is an identifiable trend in the relationship between u_{max} and P_u (Fig. 4). Several controllers (E, H, and I) have zero probability of violating this criterion for problem E-1, and designs B, C, and I retain low values of P_u for all three problems. Designs D and F have 100% P_u in all three cases, which is traceable to very high nominal control usage. Once again, nominally marginal cases (G and J, the two controllers designed for rejection of the sinusoid) exhibit reduced probability of exceedance for problems E-2 and E-3. From Eq. (1), increased m_1 and m_2 decrease the effects of u and w, and added damping (c) and first-order lag (τ) reduce control peaks in some cases, reducing the probability of high control levels.

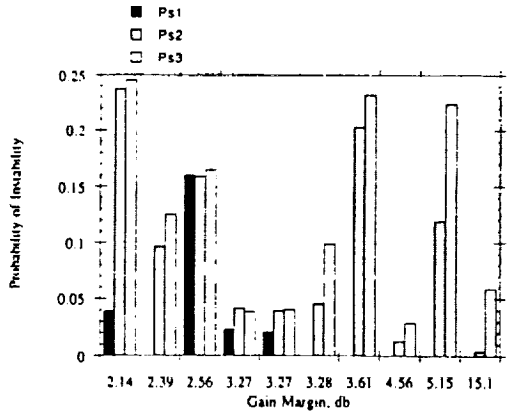


Fig. 2 Probability of instability vs gain margin for three evaluation problems.

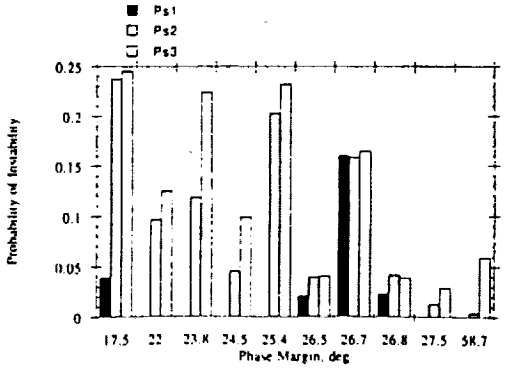


Fig. 3 Probability of instability vs phase margin for three evaluation problems.

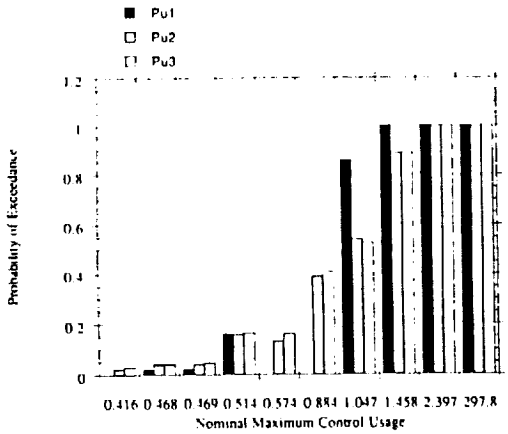


Fig. 4 Probability of control-limit exceedance vs nominal maximum control response to a disturbance impulse.

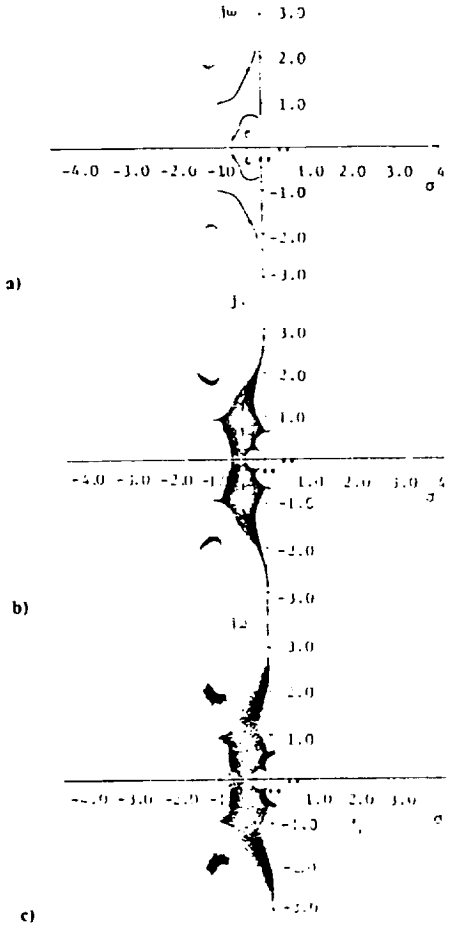


Fig. 5 Stochastic root locus of design H (scatter plot): a) problem E-1; b) problem E-2; c) problem E-3.

Sinusoidal Response Characteristics

When 0 dB is chosen as an upper response limit, the two controllers designed to reject the sinusoid (G and J) do so perfectly ($P_f = 0$), whereas all the others exceed the limit all the time. The transfer functions (Appendix) show that designs G and J effectively "notch" the 0.5-rad/s disturbance-input frequency to produce these results. Without notch filters the remaining controllers cannot give special attention to discrete-frequency inputs, and their frequency response of ~0.5 rad/s always exceeds 0 dB. If the frequency of the sinusoidal disturbance were uncertain, the notch filters could be less effective, but there would be little change in the response of the other controllers.

Stochastic Root Loci and Parametric Histograms

Graphical results give insight into the nature and causes of possible instability. The stochastic root locus is an s-plane plot of the eigenvalues that result from each Monte Carlo evaluation, expressed either as a two-dimensional scatter plot of closed-loop roots or an oblique three-dimensional view of the density of roots within subspaces of the s plane.[13] The former plot is easily generated from the calculations, and the latter has the advantage of showing the distribution along the real axis.[17] In addition, histograms of the parameters associated with instability can be related to origins of the problem.

Scatter plots for design H show the progression of eigenvalue uncertainty from problem E-1 to E-3 (Fig. 5). For prob-

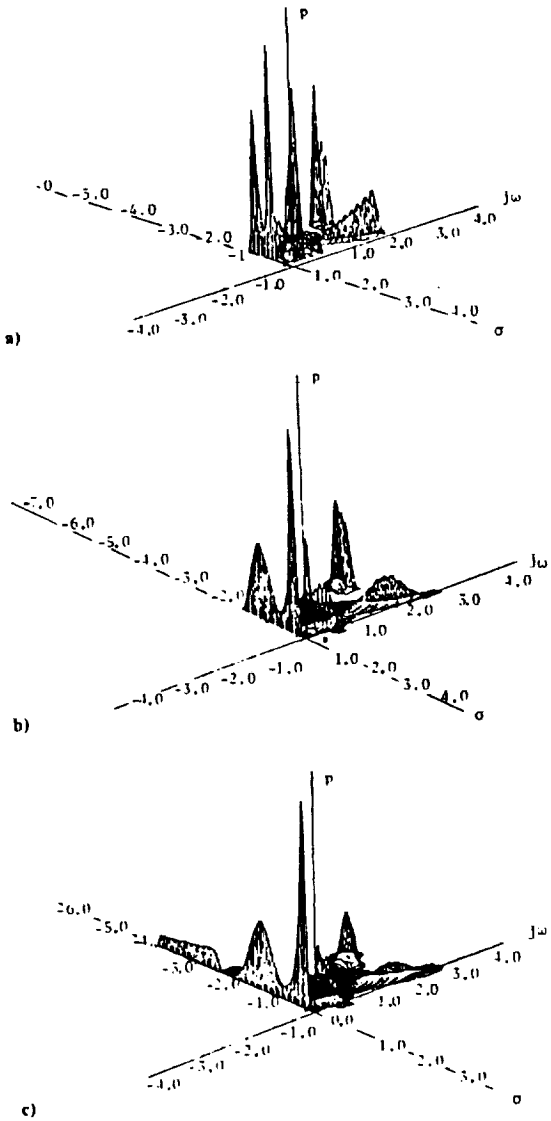


Fig. 6 Stochastic root locus of design H (three-dimensional view): a) problem E-1; b) problem E-2; c) problem E-3.

lem E-1 only a single parameter varies (the spring constant k). The distribution follows the conventional root locus (with nominal closed-loop locations indicated by ×), although the density of roots varies along the curves. The pairs of roots near the origin are most closely associated with the plant, whereas the higher-frequency roots are compensator modes. None of the root loci extend into the right half plane, and P_I is zero (Table 2). Three parameters vary in problem E-2, and the stochastic root locus becomes an areal distribution of roots, some of which extend into the right half plane (Fig. 5b). Because the parameter variations are bounded, there are crisp edges to the distributions. The unstable cusps at 0.6 and 2.6 rad/s can be associated with plant and controller modes. Further parametric uncertainty (problem E-3) broadens the distributions and increases the probability of instability.

The same information is presented in unsmoothed three-dimensional form in Fig. 6 (upper half plane only), which shows the distribution of real roots as well. The three-dimensional representation is especially effective when displayed on a graphics workstation that allows the viewpoint to "fly around" the distribution.

To see which parameter values are associated with instability, the values are recorded whenever the system is found to be unstable. These values are collected in intervals, the number of values in each interval is counted, and the resulting histogram provides an estimate of the conditional probability density function for each parameter. If a parameter has little effect on stability, then the histogram should show the same distribution as produced by the random number generator—in this case, a uniform distribution. If particular values of the parameter increase the probability of instability, the histogram has higher values in that region.

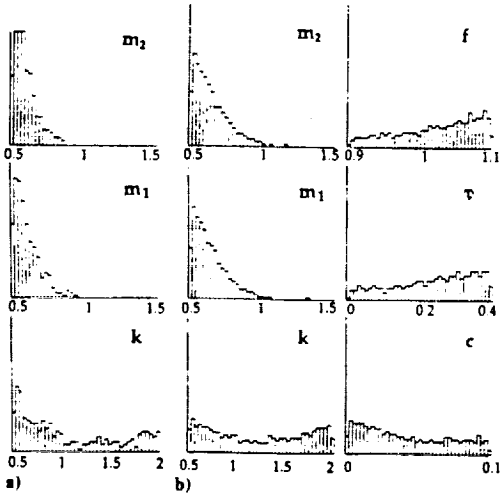


Fig. 7 Parameter histograms for all unstable cases, design H: a) problem E-1; b) problem E-3.

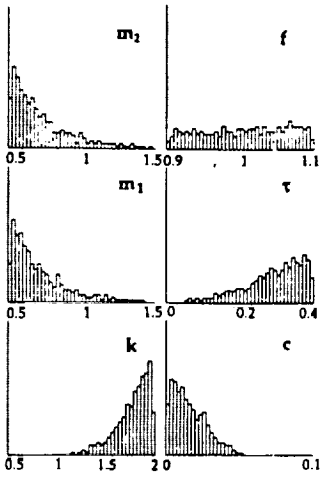


Fig. 8 Parameter histograms for high-frequency unstable cases, design H, problem E-3.

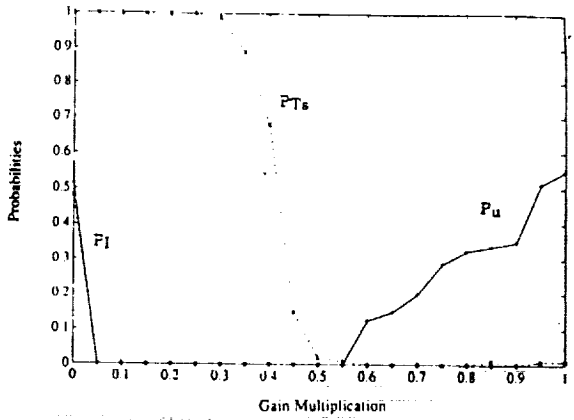


Fig. 9 Effect of reducing design D loop gain on P_I, P_{T_s}, and P_u.

For design H and problem E-2, instability often occurs when the masses have low values but never occurs with high values (Fig. 7a). Low mass values increased the probability of instability for all the designs. Extreme values of the spring constant also are associated with instability, low values having the edge in this example.

For problem E-3 (Fig. 7b), the distributions become less crisp, as otherwise unstable values of mass can be stabilized by damping and otherwise stable values of mass can be destabilized by increased loop gain or first-order lag. The spring constant shows a slight bimodal distribution due to the two modes of instability with roots of approximately 0.6 or 2.6 rad/s. This can be seen by recording the parameter values only when the system is found to be unstable *and* the unstable roots have a high frequency. The resulting histograms (Fig. 8) show that there are unstable high-frequency roots only if the spring constant is high and the damping is low. With increased damping, there is no high-frequency instability.

These results can be used in three ways. The probability of instability could be reduced if it were possible to ensure that the plant parameters did not move into the areas that are found to cause problems. This might be the result of improved quality assurance on the important parameters or by shifting the mean of the parameter variation. If it is not possible to affect the actual parameter variations, then the control system could be redesigned using the problematic values of parameters as nominal values. For example, the control system could be redesigned using nominal values of 0.7 for the masses.

A third use of the distributions can occur if one of the varying parameters represents a control design parameter. For instance, if the loop gain f were treated as a design variable, then it is clear that attenuating the gain would reduce the probability of instability. This alternative is demonstrated using design D. It has been seen that design D had generally good robustness but very high actuator use. Peak actuator usage can be reduced by reducing the loop gain, and the effect of gain attenuation on robustness subject to problem E-2 is shown in Fig. 9. For this analysis, only 100 Monte Carlo evaluations were carried out per design point, but the results show clear trends. As the gain is reduced, the probability of control saturation is reduced without significant increase in P_I or P_{T_s} until the attenuation reaches 0.6, when P_{T_s} begins to increase. Reducing the gain further produces a clear trade-

off between the probabilities of control saturation and settling-time violation. References 19 and 20 present similar methods of control system design based on search and statistical evaluation.

Conclusions

Stochastic robustness analysis of 10 controllers designed for the ACC Benchmark Control Problem provides useful quantification of stability and performance sensitivities to parameter variations. The SRA method is flexible and can be tailored to the design requirements and system specifications of particular control problems. Qualitative selection of the best controller depends on the relative importance of several metrics, which are readily described in a probabilistic framework.

Several conclusions can be drawn from this analysis. The analysis shows that gain and phase margins are not good predictors of the relative stability robustness of different SISO controllers, because robustness is tied closely to the actual plant uncertainties and their effects on (implied) Nyquist contours. This result implies that robustness analyses based on singular-value analysis of MIMO systems may have similar limitations. Nominal settling time did not give a good indication of the likelihood of exceeding settling-time limit, principally because most nominal values already exceeded the limit. Although this result may be an artifact of the settling-time definition (T_S^*), it reveals the counterintuitive result that uncertainty may improve the probability of remaining within a predefined limit. The relationship between maximum control response to a disturbance impulse and the probability of exceeding a control limit is more direct, as most nominal values were about half the limit value. Stochastic root loci and parameter histograms provide insight about the likely positions of the closed-loop roots and the parameter variations that lead to instability, and they suggest ways of improving plant and controller design.

Appendix: Transfer Functions of the Ten Compensators

Design A:

$$\frac{40.42(s+2.388)(s+0.350)}{(s+163.77)\left[s^2+2(0.501)(0.924)s+(0.924)^2\right]}$$

Design B:

$$\frac{42.78(s-1.306)(s+0.1988)}{(s+73.073)\left[s^2+2(0.502)(1.182)s+(1.182)^2\right]}$$

Design C:

$$\frac{0.599(s-1.253)(s+0.1988)}{\left[s^2+2(0.502)(1.182)s+(1.182)^2\right]}$$

Design D:

$$\frac{19881(s+100)(s+0.212)\left[s^2+2(0.173)(0.733)s+(0.733)^2\right]}{\left[s^2+2(0.997)(51.16)s+(51.16)^2\right]\left[s^2+2(0.838)(16.44)s+(16.44)^2\right]}$$

Design E:

$$\frac{5.369(s-0.348)(s+0.0929)}{\left[s^2+2(0.832)(2.21)s+(2.21)^2\right]}$$

Design F:

$$\frac{2246.3(s+0.237)\left[s^2-2(0.32)(1.064)s+(1.064)^2\right]}{(s+33.19)(s+11.79)\left[s^2+2(0.90)(2.75)s+(2.75)^2\right]}$$

Design G:

$$\frac{4430(s+0.08)(s-0.44)(s-2.83)\left[s^2-2(0.102)(0.49)s+(0.49)^2\right]}{\left\{\left[s^2+2(0.70)(11.17)s+(11.17)^2\right]\left[s^2+2(0.89)(3.67)s+(3.67)^2\right]\left[s^2+2(0.29)(3.11)s+(3.11)^2\right]\left[s^2+(0.5)^2\right]\right\}}$$

Design H:

$$\frac{2.13(s+0.145)(s-0.98)(s+3.43)}{\left[s^2+2(0.82)(1.59)s+(1.59)^2\right]\left[s^2+2(0.46)(2.24)s+(2.24)^2\right]}$$

Design I:

$$\frac{16.1(s+0.134)(s-1.174)(s+1.46)}{\left[s^2+2(0.82)(1.05)s+(1.05)^2\right]\left[s^2+2(0.5)(2.18)s+(2.18)^2\right]}$$

Design J:

$$\frac{51.47(s+0.06)(s-0.21)(s+5.41)\left[s^2-2(0.07)(0.51)s+(0.51)^2\right]}{\left[s^2+2(0.72)(2.05)s+(2.05)^2\right]\left[s^2+2(0.68)(5.21)s+(5.21)^2\right]\left[s^2+(0.5)^2\right]}$$

Stochastic Prediction Techniques for Wind Shear Hazard Assessment

D. Alexander Stratton* and Robert F. Stengel†
Princeton University, Princeton, New Jersey 08540

The threat of low-altitude wind shear has prompted development of aircraft-based sensors that measure winds directly on an aircraft's intended flight path. Measurements from these devices are subject to turbulence inputs and measurement error, as well as to the underlying wind profile. In this paper stochastic estimators are developed to process onboard Doppler sensor measurements, producing optimal estimates of the winds. A stochastic prediction technique determines the level of aircraft energy performance from the wind estimates. Aircraft performance degradation algorithms presented are based on optimal estimation techniques. The prediction algorithm must balance wind shear detection performance and turbulence rejection capability, as illustrated in simulations of microburst wind shear and severe turbulence environments.

Introduction

STRONG variable winds in the airport vicinity can cause unacceptable deviation of aircraft from their intended flight path. Known as low-altitude wind shear, this threat has caused at least 24 aviation accidents in the last 25 years.[1] Efforts to promote the avoidance of severe wind shear have focused on improving flight crew training programs,[2] understanding the meteorology of wind shear,[3-5] and developing technology to detect wind shear in the terminal area. Ground-based sensor systems to measure airport-vicinity winds are being developed and installed at major airports,[6,7] along with techniques to automatically identify a wind shear and predict its formation.[8-10] Sensors to detect wind-shear-induced flight-path deviations are being installed on aircraft,[11,12] and forward-looking sensors to detect wind shear in front of the aircraft also are under development.[13-15] Interpretation of this information in the cockpit is a topic of current research.

As the amount of available information grows, accurate interpretation of the information by flight crews becomes more challenging, particularly during periods of high workload. Artificial intelligence technology provides a basis for a cockpit aid to assist flight crews in avoiding low-altitude wind shear. An expert system, the Wind Shear Safety Advisor,[16] depicted schematically in Fig. 1, will operate in real time, accepting evidence from onboard and ground-based sources, perhaps facilitated by a direct data link (represented by a dotted line in Fig. 1). The goal of this system is to increase flight crew situation awareness and decision reliability by summarizing information from a variety of information sources.

In the absence of direct measurements of the winds, a decision to avoid wind shear must be based on discrete alerts from wind shear detection systems and meteorological evidence. Various levels of reliability associated with this indirect evidence complicate the risk assessment process. A probabilistic model of this process has been developed that incorporates statistics from meteorological studies and reliability statistics for wind-shear-alerting systems.[17] The model can manage the uncertainty associated with indirect evidence, providing meaningful estimates of risk.

If onboard measurements of the winds were available, a hazardous level of wind shear could be identified by determining whether the level of some hazard metric, based on the wind measurements, exceeds a threshold. Hazard metrics considered previously include maximum horizontal winds[3] and F-factor,[14] which relates wind shear to aircraft performance. Computation of the hazard level is complicated by uncertainty surrounding the wind measurements, including turbulence and measurement errors. In this paper Kalman filters are developed to produce optimal wind estimates from onboard wind sensors, based on a stochastic wind model. These algorithms are demonstrated in a simulated microburst wind shear environment.

From the wind estimates, predictions of the aircraft's performance degradation can be made using stochastic prediction techniques.[18,19] In addition to the predictions themselves, these techniques produce measures of the possible error in the predictions due to turbulence and limitations of the measurement devices. In this paper a Kalman-filter-based prediction technique to predict F-factor and aircraft performance degradation is demonstrated in simulated microburst wind shear encounter. The response characteristics of the prediction technique must provide significant response to severe wind shear and limited response to turbulence. In this paper stochastic prediction techniques with different design parameters are demonstrated in a simulated microburst wind shear and severe turbulence environments.

Probabilistic Reasoning in Artificial Intelligence

The power of an intelligent system rests in its ability to produce meaningful conclusions by reasoning, i.e., by applying knowledge stored in the system to available evidence. In probabilistic models of reasoning, knowledge is stored in the form of probabilities, and Bayes's rule[20] and the axioms of probability[21] are used to condition these probabilities on evidence. When several pieces of evidence are supplied, the application of Bayes's rule is complicated by dependencies between pieces of evidence. A structure to these dependencies must be provided for efficient reasoning. In Bayesian network representation[22] a graphical representation provides this structure, such as the one for wind shear avoidance graphed in Fig. 2. Nodes in the diagram represent discrete random variables, and the links between them represent sets of conditional probabilities used during reasoning. The network representation enables efficient probabilistic reasoning because all of the dependencies between variables are specified by the links.

The network of Fig. 2 was developed using guidelines for wind shear avoidance presented in the FAA's Windshear Training Aid document,[2] which was written by a team from

Received Feb. 12, 1991; revision received Sept. 12, 1991; accepted for publication Oct. 6, 1991.

*Research Assistant, Department of Mechanical and Aerospace Engineering. Member AIAA.

†Professor, Department of Mechanical and Aerospace Engineering. Associate Fellow AIAA.

Research supported by government grant.

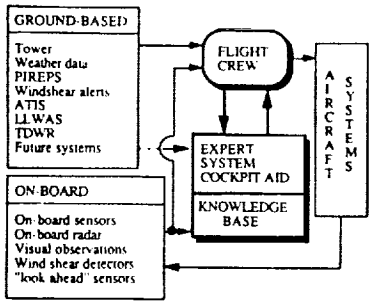


Fig. 1 Wind shear safety advisor schematic diagram.

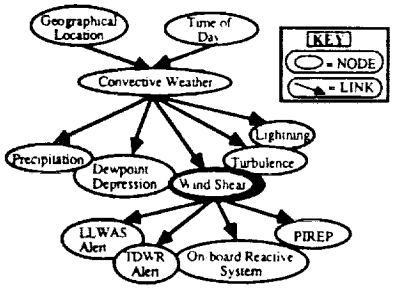


Fig. 2 Graphical representation of a Bayesian network for wind shear avoidance.

the airframe industry with the support of airlines, the government, and academia. The network model incorporates statistical results from the NIMROD,[3] JAWS,[3,4] and FLOWS[5] studies and for the enhanced Low-Level Windshear Alert System (LLWAS) evaluation.[7] Demonstrations of the network[17] show that it can approximate the subjective judgments required to establish the possible presence of wind shear.

A probabilistic model establishes a scientific basis for the Windshear Training Aid avoidance guidelines. Since the completion of the Windshear Training Aid, a variety of new ground-based and airborne wind shear detection systems are being devleoped, such as the Terminal Doppler Weather Radar (TDWR) system. The probabilistic model can be expanded to include statistics from new detection systems established during their evaluation. New knowledge gained from meteorological studies, such as geographical variation of wind shear frequency, can also be included.

Kalman Filter Development for Doppler Wind Measurements

Airborne sensor technology with the capability to detect wind shear in front of the aircraft is currently under development, including Doppler radar,[13] Doppler lidar,[14] and infrared[15] technology. Doppler devices measure a shift in frequency of radar or light waves emitted along a radial line, measuring the component of wind velocity parallel to that line. Operational devices could provide measurements of head winds or tail winds at a series of locations along the aircraft's intended approach or takeoff path. For example, airborne Doppler radars could provide measurements spaced at ~500-ft intervals over a range of 3-5 miles, spanning 50-100 s of flight at approach speed.[13] This sequence of measurements contains the effect of turbulence and is corrupted by measurement noise as well. A bank of Kalman filters can improve the accuracy of hazard estimates based on successive measurement sequences, minimizing measurement noise and accounting for correlation in the wind field using a stochastic model.

As the aircraft travels down the flight path measurements in successive sequences are offset by a distance d (Fig. 3), which is assumed to be small relative to the distance between adjacent range gates L. At a given time, a sequence of measurements is obtained. Each member of this sequence represents the average value of the radial wind component in an interval of length L at that time.

A first-order Markov model for the turbulent winds can be based on the Dryden power spectrum for horizontal turbulence, given by Ref. 23 as

$$\Phi_u(\omega) = \left(\frac{2L_u\sigma_u^2}{\pi}\right)\frac{1}{\left[1+(L_u\omega)^2\right]} \tag{1}$$

Parameters of this model include the turbulence scale length L_u and the root-mean-square turbulence amplitude σ_u. The corresponding discrete Markov sequence is

$$w_{t_k} = \exp(-d_u)\,w_{t_{k-1}} + \sqrt{1-\exp(-2d_u)}\,\eta_{k-1} \tag{2}$$

where d_u is the ratio of d to L_u. The η is a normally-distributed white noise sequence with mean and variance:

$$E[\eta_k] = 0 \tag{3}$$

$$E[\eta_k^2] = \sigma_u^2/\pi \tag{4}$$

This model uses the discrete white noise sequence η to approximate the integrated effect of continuous white noise. Figure 4 presents the autocovariance function associated with Eq. (1), along with the autocovariance function of the sequence of Eq. (2), indicating the agreement of the turbulence models.

With the assumption that measurement noise is superimposed on the radial wind components, the measurement at range gate j during measurement sequence k, z_{jk} can be related

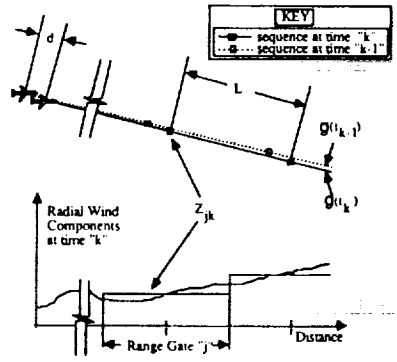


Fig. 3 Forward-look sensor measurement process.

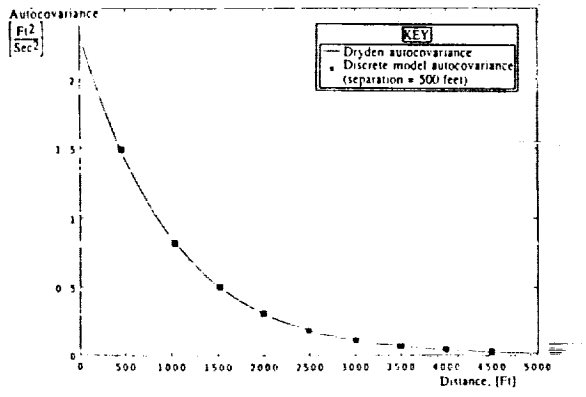


Fig. 4 Comparison of Dryden turbulence spectrum autocovariance function and autocovariance function of discrete turbulence model.

to the corresponding scalar radial wind component by the relationship

$$z_{jk} = w_{r_{jk}} + n_{jk} + V_{I_k} \tag{5}$$

This can be rewritten as

$$\tilde{z}_{jk} = w_{r_{jk}} + \tilde{n}_{jk} \tag{6}$$

where V_I is the aircraft's inertial speed at the time of measurement sequence k, and $\tilde{z}_{jk}$ has this bias subtracted out. Error in the inertial speed estimate, n_{V_I}, which is made from onboard measurements, is added to n_{jk} to produce $\tilde{n}_{jk}$:

$$\tilde{n}_{jk} = n_{jk} + n_{V_I} \tag{7}$$

The measurement error $\tilde{n}$ is assumed to be a zero-mean, normally distributed white noise sequence, with a known constant standard deviation σ_n.

With the aforementioned assumptions, an estimator dedicated to each range gate can be constructed in the form of a Kalman filter. From the measurement $\tilde{z}_{jk}$, each Kalman filter constructs an estimate $\hat{w}_{r_{jk}}(+)$ and a variance $p_{jk}(+)$, which is a measure of the uncertainty in $\hat{w}_{r_{jk}}(+)$, in three steps. First, the state estimate and variance form the previous measurement sequence, $\hat{w}_{r_{jk-1}}(+)$ and $p_{jk-1}(+)$, are extrapolated according to

$$\hat{w}_{r_{jk}}(-) = \exp(-d_u)\hat{w}_{r_{jk-1}}(+) \tag{8}$$

$$p_{jk}(-) = \exp(-2d_u)p_{jk-1}(+) + \left[1 - \exp(-2d_u)\right]\sigma_u^2/\pi \tag{9}$$

Equation (8) is obtained by taking the expected value of Eq. (2). Note that Eq. (9) is an approximation to the integrated effects of continuous white noise. Next, the extrapolated variance $p_{jk}(-)$ is used to compute a gain K_{jk}:

$$K_{jk} = \frac{p_{jk}(-)}{p_{jk}(-) + \sigma_n^2} \tag{10}$$

Finally, the post-update wind estimate and variance are computed:

$$\hat{w}_{r_{jk}}(+) = \hat{w}_{r_{jk}}(-) + K_{jk}\left[\tilde{z}_{jk} - \hat{w}_{r_{jk}}(-)\right] \tag{11}$$

$$p_{jk}(+) = \left[p_{jk}(-) + \sigma_n^2\right] / \left\{p_{jk}(-)\sigma_n^2\right\} \tag{12}$$

The Kalman filters compute a weighted average of the wind measurements obtained at each range gate, compensating for the movement of the sensor platform by making an assumption of frozen Dryden turbulence in the interval between the measurements. Wind shear estimates are updated at each measurement step, compensating for turbulence and weighing current and prior information according to its relative uncertainty. Because each range gate's state estimator is decoupled from the others, the computation could be performed on a set of identical processors running in parallel. This decoupling is achieved as a consequence of the Markov property of the wind model: the probability distribution at a given wind state $w_{r_{jk}}$ is conditionally independent of $w_{r_{j-1,k}}$ given the closer state $w_{r_{jk-1}}$. This assumption could be relaxed, coupling adjacent states or larger groups of states together with a corresponding increase in computational complexity.

Prior state estimates and variances are required to initialize each filter. This may be accomplished by applying a separate initialization Kalman filter to the first sequence of wind measurements. This filter is initialized with an onboard wind estimate and variance at the aircraft's location, perhaps from a Kalman filter processing onboard sensor measurements. An initial sequence of wind measurements from the forward-looking sensors is then processed to initialize the state and variance of each Kalman filter. The initialization Kalman filter takes the same form as Eqs. (8-12), except that the distance between range gates L is used as the distance between measurements d.

Hazard Metrics and Stochastic Prediction

The detection of the presence of a wind shear can be based on the output of the stochastic estimators. A reasonable approach to detecting wind shear is to predict whether the level of some hazard metric based on the wind estimates will exceed a threshold. The F-factor hazard metric relates wind shear to aircraft air-referenced specific energy rate, which is defined by

$$\frac{dE_s}{dt}(t) = \left(\frac{V_a}{g}\right)\frac{dV_a}{dt} + \frac{dh}{dt} \tag{13}$$

where V_a is the airspeed, h is aircraft altitude, and g is the gravitational constant. Using longitudinal aircraft equations of motion and assuming small flight-path angles, it can be shown[14] that

$$\frac{dE_s}{dt}(t) = \frac{(T - D)V_a}{W} - \mathcal{F}(t)V_a \tag{14}$$

where T is thrust, D is drag, and W is aircraft weight. $\mathcal{F}(t)$ is the F-factor, defined as

$$\mathcal{F}(t) = \left(\frac{1}{g}\right)\frac{dw_x}{dt}(t) - \frac{w_h(t)}{V_a} \tag{15}$$

where $w_x(t)$ is the wind component in the inertial horizontal direction, and $w_h(t)$ is the vertical wind component. For small flight-path angles, the radial wind components are approximately the same as the longitudinal horizontal wind components. Wind shear effects enter Eq. (14) in three ways: 1) by changing the airspeed, 2) by altering the drag, and 3) directly through $\mathcal{F}(t)$. For conditions typical of jet transport flight through severe wind shear, only the direct impact of $\mathcal{F}(t)$ is significant. Prediction of aircraft specific energy along the intended trajectory appears to involve the prediction of airspeed, but using a constant nominal value of airspeed in Eq. (15) introduces a small, conservative error.

The first component of $\mathcal{F}$ in Eq. (15) is proportional to the rate of change of the horizontal wind component. If the wind field is assumed stationary, prediction of $\mathcal{F}$ along the intended trajectory could be made by differencing adjacent wind estimates:

$$\hat{\mathcal{F}}_j = 1/L\left(\hat{w}_{r_{jk}} - \hat{w}_{r_{j-1,k}}\right) \tag{16}$$

This would amplify high-frequency noise, resulting in excessive prediction error. Alternatively, predicted energy deviation and $\mathcal{F}$ can be computed by a Kalman filter algorithm using the wind estimates as inputs. $\mathcal{F}$ is obtained through a weighted sum of the radial wind estimates, with the weights selected by definition and minimization of a suitable cost function.

An important limitation of Doppler wind measurement devices is their inability to measure winds perpendicular to the direction of the Doppler pulse. As a consequence, the second component of $\mathcal{F}$ in Eq. (15), due to vertical winds, is not measured by the device. In downburst wind shears, head-tail wind shear is produced by vertically descending winds that flow outward as they near the ground. These downdraft winds pose a hazard to the aircraft that the Doppler sensors cannot directly measure. Current research is attempting to model the vertical wind as a function of the horizontal wind for hazard estimation.[23] In the simple downburst model of Ref. 23, the correlation between horizontal and vertical winds depends on the size of the downdraft, the altitude, and the distance from the downburst core. In a well-measured and well-studied microburst, four major downdraft regions were found.[24] As the relationship between horizontal and vertical winds remains to

be established, the present study is based on radial wind alone. If a consistent correlation between vertical wind and radial-wind measurement is found, vertical wind could be added to the stochastic model.

To predict the wind-shear-induced energy deviation E_{sw}, Eq. (14) can be integrated across a typical range gate j, resulting in the recursive form

$$E_{sw_j} = E_{sw_{j-1}} - \left(\frac{V_a L}{V_t}\right)\mathcal{F}_{j-1_{av}} \tag{17}$$

where V_t is average inertial speed of the aircraft. $\mathcal{F}_{av}$ is modeled as a stationary process driven by a discrete random sequence:

$$\mathcal{F}_{j_{av}} = \mathcal{F}_{j-1_{av}} + \eta_{j-1} \tag{18}$$

where η is a normally distributed white noise sequence with zero mean and standard deviation σ_η. This standard deviation is a design parameter that alters the response characteristics of the prediction filter, as demonstrated by simulation. Equations (17) and (18) may be written in vector-matrix form:

$$x_j = \begin{bmatrix} 1 & -\left(\frac{V_a L}{V_t}\right) \\ 0 & 1 \end{bmatrix} x_{j-1} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \eta_{j-1} \tag{19}$$

where

$$x_j = \begin{bmatrix} E_{sw_j} & \mathcal{F}_{j_{av}} \end{bmatrix}^T \tag{20}$$

The relationship between prediction and estimation is obtained by substitution of Eq. (15) into Eq. (14) and integration from the aircraft (denoted with subscript 0) to a typical range gate j. This results in the equation

$$w_{x_j} - w_{x_0} = -\left(\frac{g}{V_a}\right)(E_{sw_j} - E_{sw_0}) + \left(\frac{L}{V_t}\right) w_{h_{av}} \tag{21}$$

If the prediction is initialized with the condition

$$E_{sw_0} = -(V_a/g) w_{x_0} \tag{22}$$

then Eq. (21) may be rewritten as

$$w_{x_j} = \begin{bmatrix} -\left(\frac{g}{V_a}\right) & 0 \end{bmatrix} x_j + \left(\frac{gL}{V_a V_t}\right) w_{h_{av}} \tag{23}$$

In this paper vertical wind is modeled as a normally distributed white random sequence, uncorrelated with the radial winds, with mean and variance

$$E\{w_{h_j}\} = 0 \tag{24}$$

Table 1 Simulation parameters

Aircraft initial conditions	
Airspeed, V_a	160 Kt
Altitude, h	2000 ft
Inertial flight-path angle, γ_t	−3 deg
Distance to microburst core	20,100 ft
Doppler sensor	
Range gate separation, L	500 ft
Distance between sequences, d	27 ft
Noise standard deviation, σ_n	1 ft/s
Distance to aircraft	20,000 ft
Turbulence	
rms turbulence intensity, σ_u	2.7 ft/s
Turbulence scale length, L_u	1000 ft
Microburst	
Downdraft radius	2070 ft
Maximum horizontal winds	58.4 ft/s
Height of boundary layer	131 ft

and

$$E\{w_{h_j}^2\} = \sigma_{w_h}^2 \tag{25}$$

With the previously given model, prediction of the hazard level can be made from the output of the estimation Kalman filters after each measurement sequence. The wind estimates are processed using a recursive procedure based on the Kalman filter.[18,19] The prediction is initialized with onboard estimates of w_{x_0} and $\mathcal{F}_0$. Predictions of E_{sw} and $\mathcal{F}_{av}$, denoted $\hat{E}_{sw}$ and $\hat{\mathcal{F}}$, are made for each range gate using the recursive equations

$$\hat{E}_{sw_j} = \hat{E}_{sw_{j-1}} - \frac{V_t}{gL}\hat{\mathcal{F}}_{j-1} + K_{E_j}\left[\hat{w}_{r_{j_k}} - \frac{g}{V_t}\hat{E}_{sw_{j-1}} - \frac{V_t}{gL}\hat{\mathcal{F}}_{j-1}\right] \tag{26}$$

$$\hat{\mathcal{F}}_j = \hat{\mathcal{F}}_{j-1} + K_{\mathcal{F}_j}\left[\hat{w}_{r_{j_k}} - \frac{g}{V_t}\hat{E}_{sw_{j-1}} - \frac{V_t}{gL}\hat{\mathcal{F}}_{j-1}\right] \tag{27}$$

These equations involve two gains, K_{E_j} and $K_{\mathcal{F}_j}$, that are computed at each step based on the covariance propagation and filter gain computations of the Kalman filter.[18,19] The design parameter σ_η influences the size of these gains, influencing the response characteristics of the prediction filters.

Simulation of Stochastic Prediction Techniques

The stochastic estimation and prediction algorithms are demonstrated using a batch simulation of aircraft encounters with downburst wind shear and with severe turbulence. For each simulation, two different predictions are made, based on different choices of the design parameter σ_η. The wind shear is modeled by the Oseguera-Bowles stagnation-point-flow downburst model,[25] and severe turbulence is modeled using the Dryden spectrum as presented in Ref. 26. A twin-jet transport aircraft is represented by a point-mass longitudinal model,[27] trimmed along an approach path at a constant airspeed of 160 Kts. Normally distributed white noise is superimposed on measurements to simulate Doppler sensor error. Table 1 lists the parameters of the simulation.

The wind shear simulation is initiated with the microburst just out of the sensor's detection range. Figure 5 depicts the

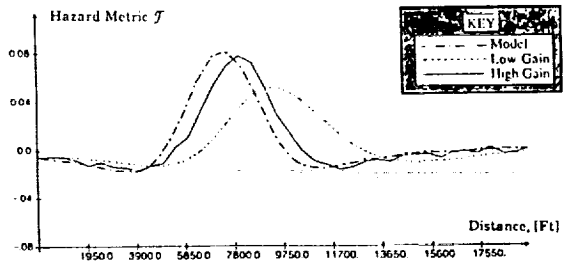


Fig. 5 Comparison of microburst model headwind-tailwind component of F-factor with predicted F-factor.

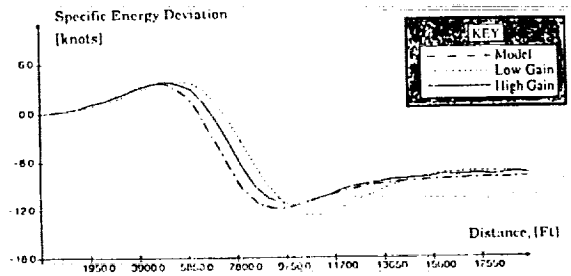


Fig. 6 Comparison of aircraft energy deviation due to headwind-tailwind shear and predicted energy deviation.

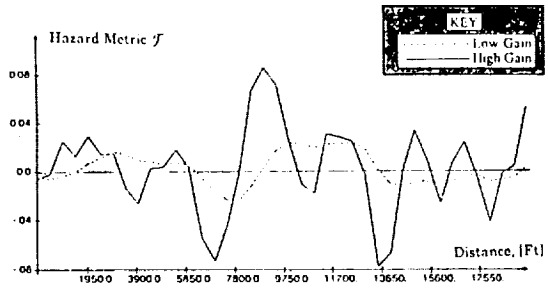


Fig. 7 Comparison of F-factor predictions in severe Dryden turbulence.

situation 10 s later, comparing the predicted hazard metric $\mathcal{F}_a$, along the flight path with the model's $\mathcal{F}_a$, component due to the headwind/tailwind shear alone. The predictions agree well with the model's head/tail wind component of $\mathcal{F}_{a_x}$, but the peak magnitude of the prediction is attenuated due to the finite bandwidth of the prediction algorithm. In addition, the distance between the aircraft and the wind shear is overpredicted due to phase shifting. With a lower value of σ_η, the estimators have lower gains, and these effects are more pronounced. If a wind shear warning were issued each time a critical value of $\mathcal{F}_a$, was exceeded, the algorithm with higher σ_η would have a greater chance of positively identifying severe wind shear.

For the same simulation, Fig. 6 compares the predicted energy deviation, normalized as an airspeed deviation, and the energy deviation due to the component of the wind shear. Although the error in prediction of distance to the microburst is greater for the lower value of σ_η, both predictions perform favorably in predicting peak energy loss. However, the total energy loss to the aircraft is greater than either prediction, due to the effect of the unobserved downdraft winds.

Figure 7 compares the predicted hazard metric $\mathcal{F}_a$, for each of the prediction designs in severe Dryden turbulence. The higher choice of σ_η results in greater response to turbulence. If wind shear warnings were issued each time a critical value of $\mathcal{F}_a$, was predicted, the algorithm with higher σ_η would issue more frequent false alarms. The optimization of a prediction algorithm must take into account both detection performance and false alarm prevention. Wavelengths corresponding to severe wind shear should be passed, but short wavelength disturbances that do not affect the flight path should be eliminated.

Conclusions

Doppler wind sensors can provide advance warning of a wind shear threat, but wind measurements are influenced by turbulence and measurement error. Optimal estimation provides a framework for minimizing the error of wind estimates given a hypothesis of the wind field structure. The estimation procedures presented here assume a structure to the local wind field at each range gate of the Doppler sensor, resulting in a bank of parallel Kalman filters. A first-order Markov turbulence model accounts for spatial correlation in the wind field due to turbulence. Measures of uncertainty are produced during the optimal estimation process. Stochastic prediction techniques are used to predict the impact of estimated winds on the energy performance of the aircraft. These techniques extend naturally to multiple Doppler sensors and could be expanded to predict other quantities such as altitude deviation error and touchdown dispersion error, given a nominal model of pilot compensation.

If wind shear warning is based on a critical threshold value of a hazard prediction, the detection reliability depends on the design of the prediction algorithm. Kalman-filter-based designs may be band limited, identifying areas with a sustained level of substantial wind shear. To further refine the algorithm, a comparative analysis of prediction algorithm designs can be conducted, using an ensemble of representative severe wind shear models. The potential for false warning in severe turbulence also can be compared. Both threshold and design bandwidth may be chosen to further optimize detection reliability.

Hazard prediction from Doppler sensors can provide the sole basis for a wind shear alert, but the lack of vertical wind estimates limits the alert's reliability. Other sources of information could improve the reliability of Doppler-based stochastic predictions through adaptive prediction techniques. Moreover, threshold exceedance of a hazard prediction could be viewed as uncertain evidence supporting a hypothesis of severe wind shear in the Bayesian network. With the reliability of threshold exceedance as evidence established through statistical analysis, hazard prediction can be incorporated into a probability-based expert system for wind shear avoidance.

Acknowledgment

This research has been sponsored by the NASA Langley Research Center under Grant NAG-1-834.

References

[1]Townsend, J., (ed.), *Low-Altitude Wind Shear and Its Hazard to Aviation*, National Academy Press, Washington, DC, 1983.

[2]*Windshear Training Aid*, U.S. Department of Transportation, Federal Aviation Administration, Associate Administrator for Development and Logistics, Washington, DC, Feb. 1987.

[3]Fujita, T. T., "The Downburst: Microburst and Macroburst," Satellite and Mesometeorology Research Project, University of Chicago, Chicago, IL, 1985.

[4]McCarthy, J., Roberts, R., and Schreiber, W., "JAWS Data Collection, Analysis Highlights, and Microburst Statistics," *Preprints, 21st Conference on Radar Meteorology*, American Meteorological Society, Boston, MA, 1983, pp. 596-601.

[5]Rinehart, R. E., and Isaminger, M. A., "Radar Characteristics of Microbursts in the Mid-South," *Preprints, 23rd Joint Conference on Radar Meteorology*, American Meteorological Society, Boston, MA, 1986, pp. J116-J119.

[6]Turnbull, D., McCarthy, J., Evans, J., and Zrnić, D., "The FAA Terminal Doppler Weather Radar (TDWR) Program," *Preprints, 3rd International Conference on the Aviation Weather System*, American Meteorological Society, Boston, MA, 1989, pp. 414-419.

[7]Barab, J. D., Page, R. D., Rosenburg, B. L., Zurinskas, T. E., and Smythe, G. R., "Evaluation of Enhancements to the Low Level Windshear Alert System (LLWAS) at Stapleton International Airport," Final Rept., DOT/FAA/PS-88/14, July 1987-March 1988.

[8]Campbell, S. D., and Olson, S., "Recognizing Low-Altitude Wind Shear Hazards from Doppler Weather Radar: An Artificial Intelligence Approach," *Journal of Atmospheric and Oceanic Technology*, Vol. 4, No. 1, March 1987, pp. 5-18.

[9]Campbell, S. D., "Microburst Precursor Recognition Using an Expert System Approach," *Preprints, Fourth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, American Meteorological Society, Boston, MA, 1988.

[10]Roberts, R. D., and Wilson, J. D., "A Proposed Microburst Nowcasting Procedure Using Single-Doppler Radar," *Journal of Applied Meteorology*, Vol. 28, No. 4, April 1989, pp. 285-303.

[11]Saint, S., "The Missing Element in Wind Shear Protection," *Business Aircraft Meeting and Exposition*, Society of Automotive Engineers Rept. SAE 830715, April 1983.

[12]Zweifel, T., "Sensor Consideration in the Design of a Windshear Detection and Guidance System," *Aerospace Technology Conference and Exposition*, Society of Automotive Engineers Rept. SAE 881417, Oct. 1988.

[13]Bracalente, E. M., and Jones, W. R., "Airborne Doppler Radar Detection of Low Altitude Windshear," *Journal of Aircraft*, Vol. 27, No. 2, 1990, pp. 151-157.

[14]Targ, R., and Bowles, R. L., "Airborne LIDAR for Avoidance of Windshear Hazards," *Proceedings of the Second Combined Manufacturer's and Technology Airborne Windshear Review Meeting* (Williamsburg, VA), Vol. 1, Oct. 1988, pp. 369-377.

[15]Scott, W. B., "Researchers Develop Airborne Flir with Ability to Pinpoint Microbursts," *Aviation Week and Space Technology*, Feb. 17, 1990, pp. 69-71.

[16]Stengel, R. F., and Stratton, D. A., "An Expert System for Wind

Shear Avoidance," *Engineering Applications of Artificial Intelligence*, Vol. 2, No. 3, Sept. 1989, pp. 190-197.

[17]Stratton, D. A., and Stengel, R. F., "Probabilistic Reasoning for Intelligent Wind Shear Avoidance," *Proceedings of the 1990 AIAA Guidance, Navigation, and Control Conference*, AIAA, Washington, DC, 1990, pp. 1099-1107.

[18]Stengel, R. F., *Stochastic Optimal Control*, Wiley, New York, 1986.

[19]Anderson, B. D. O., and Moore, J. B., *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, NJ, 1979.

[20]Bayes, T., "An Essay Towards Solving a Problem in the Doctrine of Chances," *Two Papers by Bayes*, Hafner, New York, 1963.

[21]Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1984.

[22]Pearl, J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, CA, 1988.

[23]Byrd, G. P., Proctor, F. H., and Bowles, R. L., "Evaluation of a Technique to Quantify Microburst Windshear Hazard Potential to Aircraft," *Proceedings of the 29th Conference on Decision and Control* (Honolulu, HI), Vol. 2, 1990, pp. 689-694.

[24]Proctor, F., "Model Comparison of July 7, 1990 Microburst," *Proceedings of the Third Combined Manufacturer's and Technology Airborne Windshear Review Meeting* (Williamsburg, VA), Vol. 1, 1990, pp. 81-103.

[25]Oseguera, R., and Bowles, R. L., "A Simple Analytic, 3-Dimensional Downburst Model Based on Boundary Layer Stagnation Flow," NASA TM-100632, July 1988.

[26]"Flying Qualities of Piloted Airplanes," Military Specification 8785-C, Wright-Patterson Air Force Base, OH, Nov. 1980.

[27]Hinton, D., "Flight Management Strategies for Escape from Microburst Encounters," NASA TM-4057, Aug. 1988.

Stochastic Measures of Performance Robustness in Aircraft Control Systems

SEE ALSO
93A 14595

Laura Ryan Ray*
Clemson University, Clemson, South Carolina 29634
and
Robert F. Stengel†
Princeton University, Princeton, New Jersey 08544

Stochastic robustness, a simple technique used to estimate the robustness of linear, time-invariant systems, is applied to a twin-jet transport aircraft control system. Concepts behind stochastic *stability* robustness are extended to stochastic *performance* robustness. Stochastic performance robustness measures based on classical design specifications and measures specific to aircraft handling qualities are introduced. Confidence intervals for comparing two control system designs are presented. The application of stochastic performance robustness, the use of confidence intervals, and tradeoffs between performance objectives are demonstrated by means of the twin-jet aircraft example.

Introduction

STANDARD linear control system design techniques rely on accurate models of the system to be controlled. Because models are never perfect, robustness analysis is necessary to determine the possibility of instability or inadequate performance in the face of uncertainty. Robustness to these uncertainties, parametric or unstructured, is normally treated deterministically and often without regard for possible physical variations in the system. Consequently, overconservative control system designs or designs that are insufficiently robust in the face of real-world uncertainties are a danger.

Stochastic robustness analysis (SRA), a simple technique to determine the robustness of linear, time-invariant systems by Monte Carlo methods, was introduced in Ref. 1 and presented in detail in Refs. 2 and 3. These references described stochastic stability robustness analysis and introduced the probability of instability as a scalar measure of stability robustness. Confidence intervals for the scalar probability of instability were presented, and the stochastic root locus, or probability density of the closed-loop eigenvalues, graphically portrayed robustness properties. Because it uses knowledge of the statistics of parameter variations directly, SRA provides an inherently precise yet simple characterization of robustness. The physical meaning behind the probability of instability is apparent, and overconservative or insufficiently robust designs can be avoided. Applications of SRA to full-state feedback aircraft control systems were described in Ref. 4. The results presented there illustrated the use of stochastic stability robustness techniques in comparing control system designs and in including finite-dimensional uncertain dynamics.

Concepts behind stochastic stability robustness can be extended to provide insight about control system design for performance. Design specifications such as rise time, overshoot, settling time, dead time, and steady-state error normally are used as indicators of adequate performance and lend themselves to the same kind of analysis as already described. Concepts of stochastic stability robustness analysis can be applied to these criteria giving probabilistic bounds on scalar performance criteria. Metrics resulting from SRA can be related to controller design parameters, thus providing a foundation for design tradeoffs and optimization. Extensions and uses of stochastic performance robustness in aircraft control system design and analysis are described in the following, and they are illustrated by means of an example.

Stochastic Performance Robustness

Stochastic stability robustness analysis is based on Monte Carlo analysis of the probability of instability P, and associated confidence intervals, given a statistical description of parameter uncertainty.[2-4] Because the stability test is binomial (i.e., the outcome of each Monte Carlo evaluation takes one of two values: stable or unstable), lower L and upper U confidence bounds are calculated using the binomial test.[5] While stability is an important element of robustness, performance robustness analysis is vital to determining whether important design specifications are met. Adequate performance, such as initial condition response, command response, control authority, and rejection of disturbances, is difficult to describe by a single scalar metric. Nevertheless, elements of stochastic stability robustness analysis apply for binomial performance metrics.

Numerous criteria stemming from classical control concepts exist as measures of adequate performance. Appealing to these, one can begin a smooth transition from stability robustness analysis to performance robustness analysis simply by analyzing the degree of stability or instability rather than strict stability. As described in Ref. 2, one method of doing this is to shift the vertical discriminant line from zero to $\Sigma <$ (or $>$) 0. Histograms and cumulative distributions for varying degrees of stability are readily given by the Monte Carlo estimate of the probability of any eigenvalue real-part exceeding Σ. Binomial confidence intervals are applicable to each point of the cumulative distribution as there are just two values of interest, e.g., satisfactory or unsatisfactory. P is a special case where $\Sigma = 0$. The robustness metric resulting from the cumulative probability distribution is directly related to classical concepts of rates of decay (growth) of first- and second-order closed-loop responses, time-to-half, and time-to-double. Taking degree-of-stability analysis further, rather than a vertical discriminant line, one can confine the closed-loop roots to sectors in the complex plane bounded by lines of constant damping and arcs of constant natural frequency.[6] Systems with roots

Presented as Paper 90-3410 at the AIAA Guidance, Navigation, and Control Conference, Portland, OR, Aug. 20-22, 1990; received Feb. 25, 1991; revision received Dec. 16, 1991; accepted for publication Jan. 2, 1992.

*Assistant Professor, Department of Mechanical Engineering. Member AIAA.

†Professor, Department of Mechanical and Aerospace Engineering. Associate Fellow AIAA.

Research supported by government grant.

confined to these regions would be expected to display a certain transient response speed. Again, the probability of roots lying within a sector follows a binomial distribution, and binomial confidence intervals apply.

Performance specifications for aircraft flying qualities are detailed in Ref. 7 in terms of longitudinal and lateral-directional criteria at three levels of performance for each flight phase. Many flying-qualities criteria require little computation above and beyond eigenvalue computation, making performance robustness as easy to characterize as stability robustness. For example, the short-period response can be characterized by its damping ratio and natural frequency vs normal acceleration sensitivity to angle-of-attack n_α. The latter is illustrated in Ref. 7 by plotting the short-period undamped natural frequency vs n_α, as shown in Fig. 1. n_α is simply a function of the dynamic pressure $\bar{q}$ and vehicle parameters

$$n_\alpha = \frac{\bar{q} S_{ref}}{mg} C_{L_\alpha} \tag{1}$$

C_{L_α} is the lift-curve slope, S_{ref} the wing reference area, m the mass, and g the gravitational constant. Short-period-mode requirement levels for each flight phase are characterized by calculating the closed-loop eigenvalues and evaluating Eq. 1. Repeated evaluations using Monte Carlo analysis give a distribution that can be shown pictorially on Fig. 1; the resulting measure of performance robustness is the probability of remaining within level 1, 2, or 3 criteria.[7] Binomial confidence interval computations can be applied to the scalar probability estimate.

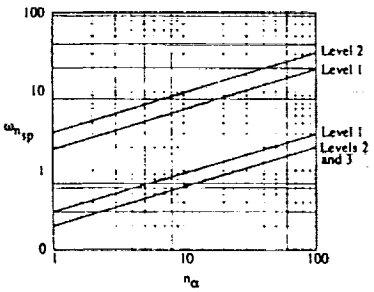


Fig. 1 Short-period response as characterized by n_α vs $\omega_{n_{sp}}$ for category B flight phase (climb, cruise, descent) and all aircraft classes.[7]

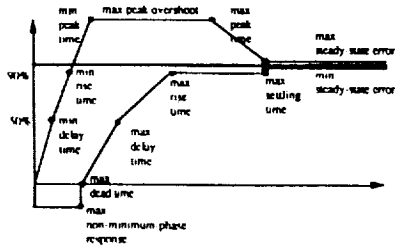


Fig. 2 Example of step response bounds formed by scalar performance characteristics.

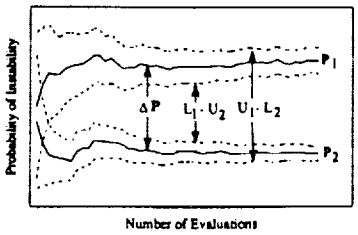


Fig. 3 Confidence interval calculation on the difference ΔP between two probabilities P_1 and P_2.

Table 1 Longitudinal parameters of the twin-jet aircraft

Uniform variation[a]	Description
15	Mass, slugs
15	Moment of inertia about the y axis, slug-ft^2
2	Wing reference area, ft
2	Aerodynamic chord, ft
2	Wing span, ft
30	Center-of-gravity location as a percent of mean aerodynamic chord
25	Lift-curve slope
25	Lift-curve intercept
40	Deviation of the basic lift coefficient due to Mach effects on lift-curve intercept
40	Deviation of the basic lift coefficient due to Mach effects on lift-curve slope
5	Variation in lift coefficient with rate of change of nondimensional α
7.5	Variation in lift coefficient with rate of change of nondimensional q
10	Variation in lift coefficient with change in elevator angle
50	Basic low-speed drag coefficient
25	Moment-curve slope
25	Moment-curve intercept
25	Deviation of the basic moment coefficient due to Mach effects on moment-curve intercept
10	Deviation in the basic moment coefficient due to Mach effects on moment-curve slope
10	Variation in moment coefficient with rate of change of nondimensional α
10	Variation in lift coefficient with rate of change of nondimensional q
15	Variation in moment coefficient with change in elevator angle
10	Center-of-gravity variation factor

[a] ± percent of nominal parameter value

Time responses provide the most clear-cut means of evaluating performance. Stochastic performance robustness can be portrayed as a distribution of possible trajectories around a nominal or desired trajectory. After defining "envelopes" around the nominal trajectory (Fig. 2), the probability of violating the envelopes can be computed using Monte Carlo evaluation. The envelope chosen around the nominal trajectory encompasses scalar performance measures; the trajectories in Fig. 2 are examples of bounds defined by minimum and/or maximum allowable dead time, delay time, rise time, time-to-peak overshoot, peak overshoot, settling time, and steady-state error.[6] Although it is simple to conclude that a response violates an envelope, individual responses within the envelope may not be acceptable. In such cases, the derivative of a response and envelopes around the derivative also can be used as performance criteria.[3]

The criteria defining envelopes that bound an acceptable time response are not unique; the segmented envelopes in Fig. 2 can be smoothed, or other scalars can be used to define points on the envelope. However, once an envelope is defined, time response distributions due to a command input, disturbance, initial condition, or some combination can be computed by Monte Carlo methods. For each evaluation, the trajectory is a binomial variable; it either stays within the envelope or violates the envelope, and binomial confidence intervals apply. Although individual time responses require more computation time than do individual sets of eigenvalues, such analysis is well within the capability of existing workstations.

Confidence intervals for the difference between two probabilities are useful when comparing two control system designs. A statistic on the difference decides whether one controller is more robust than another, either as part of an iterative design process or as imbedded in an optimization technique. The

Table 2 Scalar performance criteria defining command response envelope

Scalar metric	Value
Maximum dead time	2.5 s
Maximum nonminimum-phase response	−0.1 of desired steady-state value
Minimum and maximum delay time	1.0 s and 7.5 s
Minimum and maximum rise time	2.0 s and 15.0 s
Minimum and maximum peak time	3.0 s and 18.0 s
Maximum peak overshoot	1.25 of desired steady-state value
Maximum settling time	22.0 s
Minimum and maximum steady-state error	±0.025 of desired steady-state value

Table 3 Setpoint for individual velocity and flight-path-angle commands

Command	δT, %	δE, deg	V, fps	γ, deg	q, rad/s	α, deg
V = 15 fps	1.1	15.3	15	0	0	−0.25
γ = 4 deg	24.1	0.6	0	4	0	−0.01

statistics literature gives several methods of computing the confidence interval for the difference between two binomial variables. Reference 8 presents a method based solely on individual confidence intervals. Given individual intervals based on independent Monte Carlo trials,

$$\Pr(L_1 \le P_1 \le U_1) = 1 - \alpha_1 \tag{2}$$

$$\Pr(L_2 \le P_2 \le U_2) = 1 - \alpha_2 \tag{3}$$

the confidence interval around $\Delta P \triangleq P_1 - P_2$ is given by [8]

$$\Pr[(L_1 - U_2) \le \Delta P \le (U_1 - L_2)]$$
$$\ge 1 - \alpha_1 - \alpha_2 + \alpha_1\alpha_2 \tag{4}$$

When identical parameter sets are used to generate individual intervals, the right-hand side of Eq. (4) is $1 - \alpha_1 - \alpha_2$. Since (L_1, U_1) and (L_2, U_2) are computed using the binomial test and represent exact intervals for the individual estimates, Eq. (4) is not an approximation. Confidence interval comparisons are illustrated schematically in Fig. 3. The interpretation of the confidence interval for the difference is straightforward; the probability that the true difference lies within $[(L_1 - U_2), (U_1 - L_2)]$ is at least $1 - \alpha_1 - \alpha_2 + \alpha_1\alpha_2$. If the interval on ΔP contains zero (i.e., if the individual intervals overlap as they do initially in Fig. 3), then the difference in robustness between the two systems is not proven significant at that number of evaluations. If the true difference ΔP is small, a larger number of evaluations may result in an interval that does not contain zero, as in Fig. 3.

A given ΔP can result from many combinations of individual probability estimates, and it is difficult to generalize the number of evaluations necessary to detect a difference of a certain magnitude. Nevertheless, the number of evaluations required for an individual confidence interval can be used to foretell the number of evaluations necessary to detect a difference between two estimates. Figure 4 gives the required number of evaluations J for each individual confidence interval, for the special case, $\alpha_1 = \alpha_2 = 0.05$. Using the difference $P_2 - P_1$ as the ordinate and P_1 as the abscissa, the curves show the minimum number of evaluations required to establish a significant difference. For example, if the probability estimates (denoted $\hat{P}$) are $\hat{P}_2 = 0.45$ and $\hat{P}_1 = 0.4$, Fig. 4 shows that a statistically significant difference (i.e., nonoverlapping confidence intervals) can be determined using approximately 1500 Monte Carlo evaluations. Individual estimates of $\hat{P}_2 = 0.15$ and $\hat{P}_1 = 0.1$ result in the same difference, but fewer than 750 evaluations are required to detect the difference. Figure 4 is based on individual confidence interval calculations, as presented in Ref. 3.

Performance Robustness of Longitudinal Controllers For a Jet Transport

SRA is applied to a twin-jet transport aircraft, with the goal of characterizing the performance robustness of longitudinal

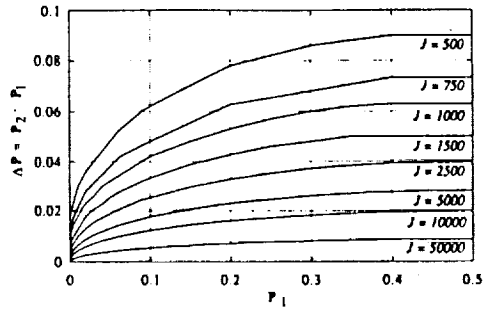
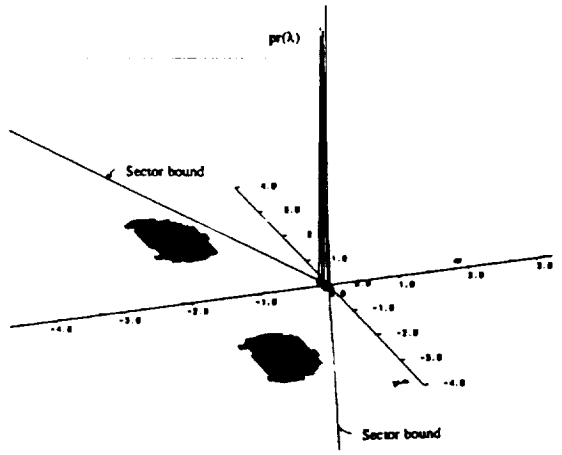
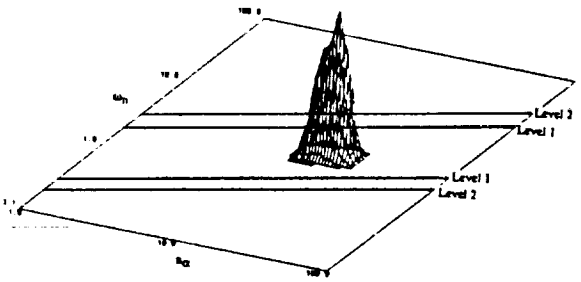


Fig. 4 Number of evaluations establishing significant differences between two probabilities for 95% confidence intervals and equal numbers of evaluations for individual probabilities.



a) Stochastic root locus with sector bounds defined by minimum level 1 short-period damping for cruise flight



b) Short-period frequency vs acceleration sensitivity distribution

Fig. 5 Stochastic robustness evaluation of the open-loop short-period dynamics of the twin-jet aircraft, based on 10,000 Monte Carlo evaluations.

command responses. The rigid-body nonlinear longitudinal equations are

$$\begin{bmatrix} \dot{V} \\ \dot{\gamma} \\ \dot{q} \\ \dot{\alpha} \end{bmatrix} = \begin{bmatrix} \dfrac{-D + T\cos(\alpha)}{m} - g\sin(\gamma) \\ \dfrac{L + T\cos(\alpha)}{mV} - \dfrac{g\cos(\gamma)}{V} \\ \dfrac{M}{I_{yy}} \\ q - \dot{\gamma} \end{bmatrix} \tag{5}$$

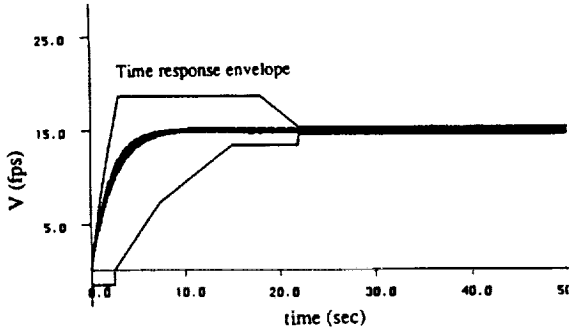
where $[V, \gamma, q, \alpha]$ represent velocity, flight-path-angle, pitch rate, and angle-of-attack, $[L, D, M]$ are aerodynamic lift, drag, and pitching moment, T is the thrust, and g is the gravitational constant. Equation (5) depends on a number of parameters given in Table 1. Mean parameter values of the stability derivatives in Table 1 are functions of Mach number and altitude; they are interpolated from aerodynamic data curves for the aircraft at a given trim condition.[9] The aerodynamic model used to compute L, D, and M is a simplified version of that given in Ref. 9, modified to use only two longitudinal controls (thrust and elevator). In this example, each Monte Carlo evaluation begins with the nonlinear equations of motion and associated parameters. The nonlinear equations are evaluated using appropriately distributed random parameters and are then linearized around the nominal trim condition. The closed-loop eigenvalues and performance metrics are evaluated from the linearized system.

The parameters are assumed to have uniform variations of the magnitudes given in Table 1. For the wing parameters (S_{ref}, chord, span), these variations are representative of loose manufacturing tolerances. The mass and moment-of-inertia variations are based on the maximum and minimum possible values of these parameters given in Ref. 9. The remaining parameter-variation estimates are based on interpolation accuracy and possible flight condition variations around the nominal value.

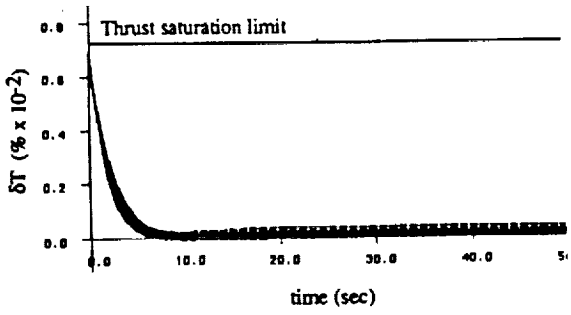
Trim conditions for a flight condition of $V = 425$ fps (130 m/s) at an altitude of 5000 ft (1524 m) are as follows: thrust = 27.3%, elevator = −0.65 deg, and angle-of-attack = 2.15 deg. The open-loop eigenvalues for the state matrix resulting from linearizing Eq. (5) around trim are $\lambda = -1.32 \pm 2.44j$, $-0.0053 \pm 0.0962j$. Stochastic robustness evaluation using the short-period Mil-spec requirements[7] shows an acceptable open-loop short-period mode for the uniform parameter variations given in Table 1. Figure 5a shows the stochastic root locus with sectors defined by minimum level 1 short-period damping ratio for cruise or climb (category B flight phase); for 10,000 evaluations, the short-period eigenvalues never violate the level 1 damping restriction. Figure 5b characterizes the short-period frequency vs acceleration sensitivity, which also remains within level 1 constraints for 10,000 evaluations. The probability estimate of violating level 1 short-period specifications is 0, with 95% confidence intervals of (0, 3.69E − 4).

Design of Longitudinal Controllers

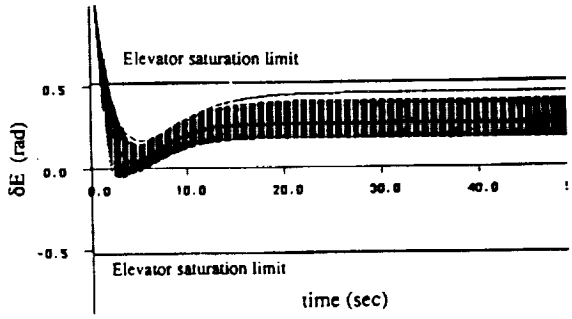
A command response that stays within the envelope described by scalar criteria in Table 2 serves as the performance requirement for designing linear regulators for velocity and flight-path-angle commands. In addition, elevator deflections are limited to ±30 deg, and thrust commands must remain between 0 and 100%. The desired commands $y^* = V^*$ or $y^* = \gamma^*$ and corresponding setpoints $\mathbf{x}^* = [V\ \gamma\ q\ \alpha]^T$, $\mathbf{u}^* = [\delta T\ \delta E]$ are given in Table 3. The open-loop responses to individual velocity and flight-path-angle commands are inadequate because of the slow, lightly damped phugoid mode. Numerical values of the results that follow depend heavily on the performance criteria chosen. The envelopes defined in Table 2 reflect tolerable variations around an acceptable nominal response. The control limits are typical of those for a jet transport. Changing the time response envelopes or control authority limits would give different numerical results. The emphasis in this example is not on the specific criteria chosen, but on how SRA characterizes performance given a control system design and performance specifications.



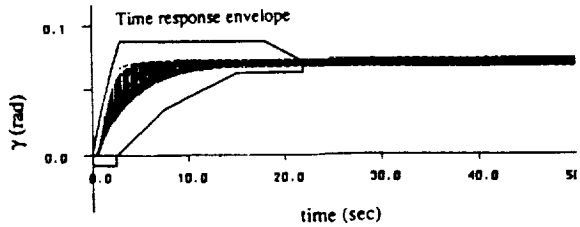
a) 15 fps velocity command: velocity response



b) 15 fps velocity command: thrust response



c) 15 fps velocity command: elevator response



d) 4-deg flight-path-angle command: flight-path-angle response

Fig. 6 Closed-loop command responses using IMF controller, 500 Monte Carlo evaluations. Nominal response is indicated by the solid line.

Structured linear-quadratic regulators[10] offer a simple means of designing a linear control system with desirable performance and robustness characteristics. Specifications of the linear-quadratic performance index and subsequent control gains using implicit-model-following (IMF) minimizes the dynamic response error between the closed-loop system and an ideal model.[10] State, control, and cross-weighting matrices (Q, R, M) are based on a quadratic cost function that weights the difference between the actual state rate ($\dot{x}$) and that of an ideal model ($\dot{x}_M$), where

$$\dot{x}_M = F_M x_M \tag{6}$$

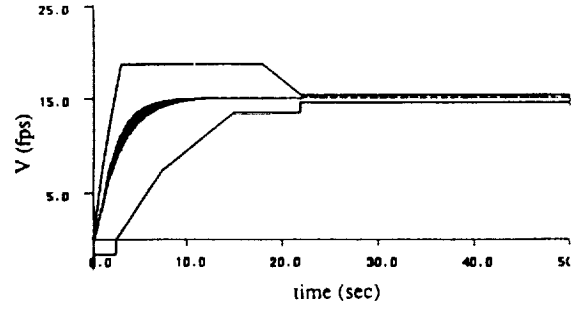
IMF offers a straightforward way of designing controllers that approximate desired dynamic characteristics. For this example, the ideal model was chosen to increase the natural frequency and damping of the phugoid mode, while maintaining acceptable short period response:

$$F_M = \begin{bmatrix} -0.3 & -32.17 & -0.0104 & -23.34 \\ 0.00381 & -0.1949 & 0.0006 & 1.356 \\ 0.0 & -0.0 & -1.273 & -5.981 \\ -0.0038 & 0.1949 & 0.999 & -1.356 \end{bmatrix} \tag{7}$$

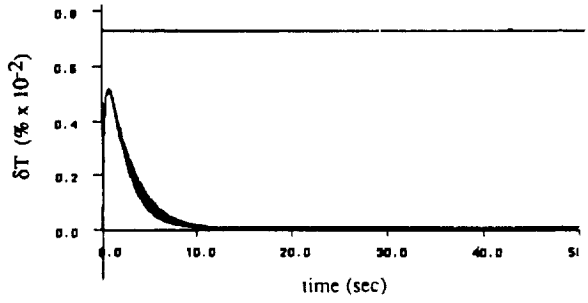
$$\gamma_M = \begin{cases} -1.35 \pm 2.39j \\ -0.213 \pm 0.314j \end{cases} \tag{8}$$

Stochastic performance robustness analysis is based on the probability of violating the desired time response envelopes ($\hat{P}_V$ and $\hat{P}_\gamma$) and the probability of control saturation ($\hat{P}_{\delta T}$ and $\hat{P}_{\delta E}$).

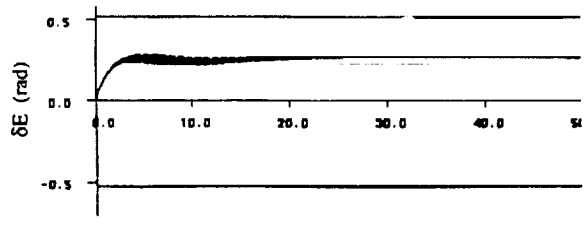
The IMF controller gives a nominal closed-loop command response to separate velocity (Figs. 6a–c) and flight-path-angle (Fig. 6d) commands that is within the acceptable time-response envelope. Figure 6 also shows 500 Monte Carlo evaluations of the command response; the nominal steady-state control inputs and state are given in Table 3, and the nominal response in Fig. 6 is indicated by a solid line. The response and associated envelopes in Fig. 6 are shown for the commanded variable only; the remaining state elements do not require performance constraints in this example. Thrust and elevator time histories are shown for the velocity command response only. Parameter uncertainty effects appear as variations around the nominal response, indicated by the dark distribution and associated outliers. Parameter uncertainty results in a distribution of transient responses that stays within the envelope, and nonzero steady-state errors that violate the envelope for both velocity (Fig. 6a) and flight-path-angle (Fig. 6d) commands. Based on 500 Monte Carlo time response evaluations, the estimate $\hat{P}_V$ is 0.002 with 95% confidence intervals (5.1E − 5, 0.0111) and the estimate $\hat{P}_\gamma$ is 0.368 (0.326, 0.412). The nominal elevator response violates control limits for both command responses, and in each case, the probability of elevator saturation is $\hat{P}_{\delta E} = 1.0$. Note that the control saturation limits in Figs. 6b–c are adjusted to reflect the remaining control authority after considering trim requirements.



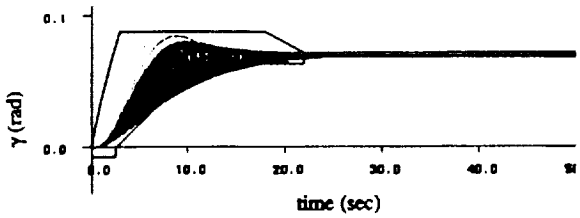
a) 15 fps velocity command: velocity response



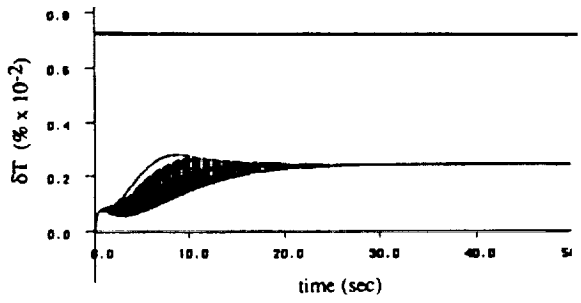
b) 15 fps velocity command: thrust response



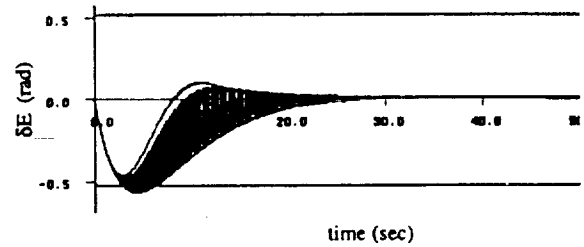
c) 15 fps velocity command: elevator response



d) 4-deg flight-path-angle command: flight-path-angle response



e) 4-deg flight-path-angle command: thrust response



f) 4-deg flight-path-angle command: elevator response

Fig. 7 Closed-loop command response using PFIMF controller, with filter control weighting R_F = diag(10, 50), 500 Monte Carlo evaluations. Nominal response is indicated by the solid line.

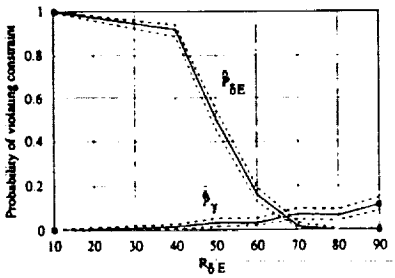


Fig. 8 Stochastic performance robustness evaluation with PFIMF: Probability of violating flight-path-angle command response $\hat{P}_{\gamma}$ and probability of violating elevator saturation limits $\hat{P}_{\delta E}$ vs filter weight $R_{\delta E}$. Solid lines give probability estimates, dashed lines give confidence intervals.

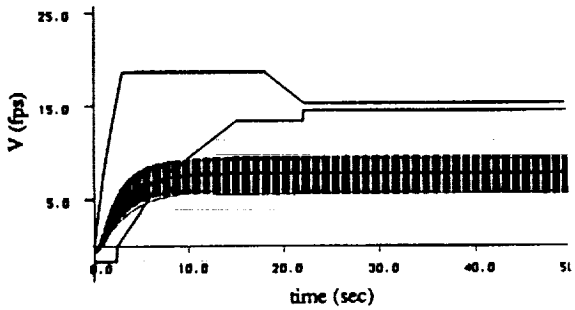


Fig. 9 Closed-loop command response using PFIMF controller, with filter control weighting $\mathbf{R}_F$ = diag(10, 50), 500 Monte Carlo evaluations: 15 fps velocity command subject to constant disturbance w_v = 40 fps. Nominal response is indicated by the solid line.

Implicit model following modified by state augmentation[10] can help meet control authority constraints. Proportional-filter (PF) compensation adds integrators to restrict the control rates, thus preventing instantaneous control changes and reducing the maximum control effort. The control vector is appended to the state vector

$$\begin{bmatrix} \dot{\tilde{\mathbf{x}}} \\ \dot{\tilde{\mathbf{u}}} \end{bmatrix} = \begin{bmatrix} \mathbf{F} & \mathbf{G} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{u}} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} v(t) \tag{9}$$

where $\mathbf{F}$ and $\mathbf{G}$ are the nominal dynamic and control effect matrices, $\tilde{\mathbf{x}} = \mathbf{x}(t) - \mathbf{x}^*$, $\tilde{\mathbf{u}} = \mathbf{u}(t) - \mathbf{u}^*$, and $v(t)$ is a commanded control rate. The PFIMF state weighting matrix is

$$\mathbf{Q}_F = \begin{bmatrix} \mathbf{Q} & \mathbf{M} \\ \mathbf{M}^T & \mathbf{R} \end{bmatrix} \tag{10}$$

where $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{M}$ are the original (IMF) weighting matrices. A weighting matrix, $\mathbf{R}_F$, constrains the control rates. Elements of $\mathbf{R}_F$ affect the bandwidth of each control; the larger the weight, the more the control rate is restricted.

The IMF regulator is augmented to include low-pass filtering of the control command, with a diagonal control-rate weighting matrix $\mathbf{R}_F$ = diag[10, 50]. Figure 7 shows 500 stochastic state and control histories to individual velocity and flight-path-angle commands using the PFIMF controller and a stream of random numbers independent from the IMF case. The (1, 1) element of $\mathbf{R}_F$ ($R_{\delta T}$) determines the amount of filtering on thrust rate, and the (2, 2) element ($R_{\delta E}$) controls elevator rate. With filter elements, the control rates are no longer unlimited, and the mean control responses remain unsaturated. Steady-state error due to parameter uncertainty remains within the desired state history envelope for the velocity command response (Fig. 7a). Steady-state error for the γ command improves, although the variation in the γ transient response is much greater than that of the IMF regulator alone, as seen by comparing Figs. 6d and 7d. $\hat{P}_V$ and $\hat{P}_{\gamma}$ estimates corresponding to Fig. 7 are 0.0 (0.0, 0.0074) and 0.034 (0.0199, 0.0539), respectively. For 500 evaluations, the PFIMF flight-path-angle command response improvement over the IMF case alone proves significant by application of confidence intervals on the difference ($P_{\gamma \text{IMF}} - P_{\gamma \text{PIMF}}$). Applying Eq. 4,

$$\Pr[0.2721 \le (P_{\gamma \text{IMF}} - P_{\gamma \text{PFIMF}}) \le 0.3921] \ge 0.9025 \tag{11}$$

Equation 11 states that with PF augmentation between 27 and 39%, more of the flight-path-angle responses lie within the envelope, with a confidence coefficient of at least 0.9025. The mean elevator response for the flight-path-angle command dips just to saturation limits, and the probability of elevator saturation is $\hat{P}_{\delta E_{\text{PFIMF}}} = 0.502$ (0.457, 0.547).

Stochastic robustness analysis shows that PF augmentation improves performance objectives by reducing control rates and steady-state error due to uncertainty. The state and control response to the velocity command prove acceptable ($\hat{P}_V$, $P_{\delta E}$, and $\hat{P}_{\delta T}$ all equal 0), and the improved responses to flight-path-angle command are statistically significant. For the flight-path-angle command, SRA demonstrates the tradeoff between the two performance objectives; increasing the (2, 2) element ($R_{\delta E}$) of $\mathbf{R}_F$ will further reduce elevator command authority at the expense of the γ time response. Figure 8 illustrates this tradeoff by showing $\hat{P}_{\gamma}$, $\hat{P}_{\delta E}$, and their confi-

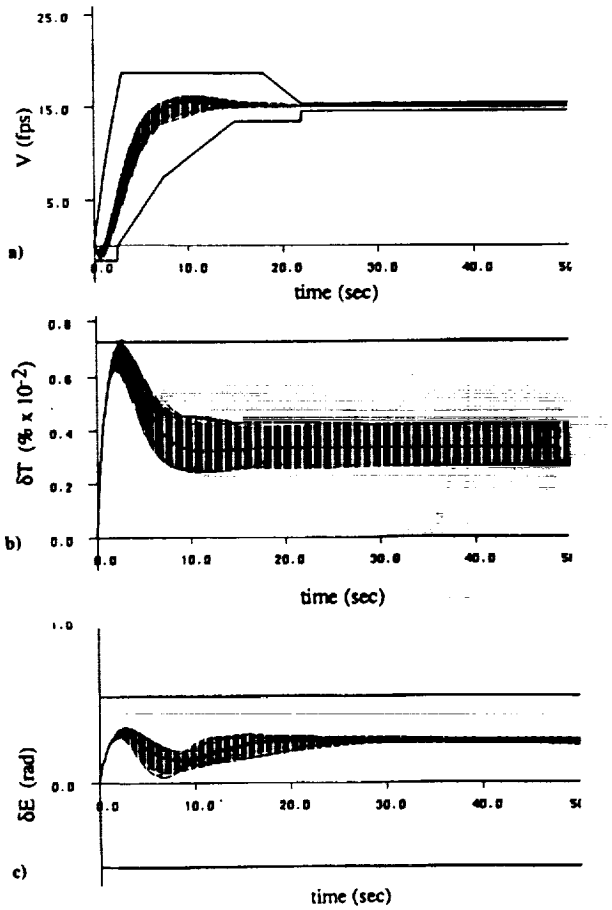


Fig. 10 Closed-loop command response using PIFIMF controller, with filter control weighting $\mathbf{R}_F$ = diag(200, 50), and integral state weighting $\mathbf{Q}_I$ = diag(0.1, 100), 500 Monte Carlo evaluations: 15 fps velocity command subject to constant disturbance w_v = 40 fps. Nominal response is indicated by the solid line.

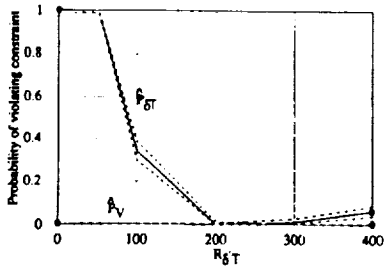


Fig. 11 Stochastic performance robustness evaluation with PIFIMF: Probability of violating velocity command response envelope $\hat{P}_V$ and probability of violating thrust saturation limits $\hat{P}_{\delta T}$ vs filter weight $R_{\delta T}$. Solid lines give probability estimates, dashed lines give confidence intervals.

dence intervals as functions of the design parameter $R_{\delta E}$. A plot like Fig. 8 can be used to choose the filter weight that gives the smallest probabilities of envelope violation while adhering as well as possible to the control authority restrictions. In this case, it is not possible to simultaneously reduce $\hat{P}_\gamma$ and $\hat{P}_{\delta E}$ to zero by varying $R_{\delta E}$. Nevertheless, stochastic robustness analysis offers a simple, understandable means of relating design parameters to performance objectives and of choosing the best control gains to meet those objectives.

Design of a Longitudinal Controller for Disturbance Rejection

As a final example, the preceding analysis is extended to encompass a performance constraint on disturbance rejection. The equations of motion are modified to include a vertical wind disturbance w_v

$$\begin{bmatrix} \dot{V} \\ \dot{\gamma} \end{bmatrix} = \begin{bmatrix} \dfrac{-D\sin(\alpha-\alpha_a)+T\cos(\alpha-\alpha_a)}{m} - g\sin(\gamma) \\ \dfrac{L\sin(\alpha-\alpha_a)+T\cos(\alpha-\alpha_a)}{mV} - \dfrac{g\cos(\gamma)}{V} \end{bmatrix} \tag{12}$$

where

$$\alpha_a = \alpha + \gamma - \tan^{-1}\frac{V\sin(\gamma)+w_v}{V\cos(\gamma)} \tag{13}$$

With the disturbance present, the state components represent inertial velocity, flight-path-angle, pitch, and angle-of-attack, and the disturbance enters through the expression for air-relative angle-of-attack α_a. A disturbance input matrix is defined for robustness analysis by numerical linearization of the nonlinear equations with respect to w_v, around the nominal condition $w_v = 0$. Velocity command response subject to a constant 40-fps vertical velocity disturbance using the PFIMF controller is shown in Fig. 9. The mean response shows a nonzero steady-state error that violates the command response envelope, and uncertainty causes a larger spread around the nominal response than that of the system without the disturbance (Fig. 7). Also, the steady-state flight-path-angle (not shown) is less than zero due to the disturbance.

Proportional-integral (PI) compensation introduces a command-error integral for each commanded state element, zeroing steady-state error and improving disturbance rejection characteristics. The perturbation equations for the nominal system are

$$\begin{bmatrix} \dot{\tilde{\mathbf{x}}} \\ \dot{\xi} \end{bmatrix} = \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}(t) \\ \xi(t) \end{bmatrix} + \begin{bmatrix} \mathbf{G} \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{u}}(t) \tag{14}$$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{15}$$

where

$$\xi(t) = \xi(0) + \int_0^t \tilde{\mathbf{y}}(\tau)\,d\tau \tag{16}$$

and $\tilde{\mathbf{y}}(t) = \mathbf{y}(t) - \mathbf{y}^*$. Here, $\mathbf{y}^* = [V\ \gamma]^T$, and a (2×2) weighting matrix $\mathbf{Q}_I$ is appended to the original state weighting matrix. Diagonal elements of $\mathbf{Q}_I$ affect the rate at which the command error integrals approach zero. The diagonal components are chosen to keep the velocity command within the desired envelope and to zero the flight-path-angle response. Command error integrals are added to the existing PFIMF controller, and for the resulting PIFIMF system with $\mathbf{Q}_I = \text{diag}[0.01, 100]$ and $\mathbf{R}_F = \text{diag}[200, 50]$, Fig. 10 shows an improved velocity command response $\mathbf{y}^* = [V^*\ 0]^T$. The 500-evaluation probability estimates and 95% confidence intervals are $\hat{P}_V = 0$ (0.0, 7.4E-3) and $\hat{P}_{\delta T} = 0.002$ (5.1E-5, 0.0111). The (1, 1) component of $\mathbf{R}_F$ is increased to restrain thrust as the command error integrals are introduced. Figure 11 shows analysis of the tradeoff between $\hat{P}_V$ and $\hat{P}_{\delta T}$ as a function of design parameter $R_{\delta T}$ comparable to that presented for the flight-path-angle response in Fig. 8. Again, Fig. 11 can be used to choose control system design parameters that best meet performance objectives.

Conclusion

Stochastic robustness analysis offers a rigorous yet straightforward alternative to other robustness metrics that is simple to compute and is unfettered by normally difficult problem statements, such as non-Gaussian statistics, products of parameter variations, and structured uncertainty. The analysis embraces both stability and performance metrics, handling qualities requirements, and more general responses. Binomial confidence intervals provide statistical bounds on the probability of instability and on performance metrics. Statistical comparisons of control system robustness also are rendered through confidence intervals. Both stability and performance metrics resulting from stochastic robustness analysis provide details relating system specifications intrinsic to a given application and control system design parameters. Stochastic robustness analysis has a significant role to play in computer-aided control system design.

Acknowledgments

This research has been sponsored by the Federal Aviation Administration and the NASA Langley Research Center under Grant No. NGL 31-001-252 and by the Army Research Office under Grant No. DAAL03-89-K-0092.

References

[1]Stengel, R. F., "Some Effects of Parameter Variations on the Lateral-Directional Stability of Aircraft," *Journal of Guidance and Control*, Vol. 3, No. 2, 1980, pp. 124-131.

[2]Stengel, R. F., and Ray, L. R., "Stochastic Robustness of Linear Time-Invariant Control Systems," *IEEE Transactions on Automatic Control*, Vol. 36, No. 1, 1991, pp. 82-87.

[3]Ray, L. R., *Stochastic Robustness of Linear Multivariable Control Systems: Towards Comprehensive Robustness Analysis*, Ph.D. Dissertation, Dept. of Mechanical and Aerospace Engineering, Princeton Univ., MAE-1902-T, Princeton, NJ, 1991.

[4]Ray, L. R., and Stengel, R. F., "Application of Stochastic Robustness to Aircraft Control Systems," *Journal of Guidance, Control, and Dynamics*, Vol. 14, No. 6, 1991, pp. 1251-1259.

[5]Conover, W. J., *Practical Non-Parametric Statistics*, Wiley, New York, 1980.

[6]Franklin, G. F., et al., *Feedback Control of Dynamic Systems*, Addison-Wesley, Reading, MA, 1991.

[7]*Military Specification Flying Qualities of Piloted Airplanes*, U.S. Air Force, MIL-F-8785C, Wright-Patterson AFB, OH, Nov. 1980.

[8]Lavenberg, S. S., (ed.), *Computer Performance Modeling Handbook*, Academic, New York, 1983.

[9]*TVC Simulation Engineering Manual*, Boeing Co., Jan. 1982.

[10]Stengel, R. F., *Stochastic Optimal Control: Theory and Application*, Wiley, New York, 1986, Chap. 6.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 1994	Conference Publication

4. TITLE AND SUBTITLE
FAA/NASA Joint University Program for Air Transportation Research 1992–1993

5. FUNDING NUMBERS
505-64-52-01

6. AUTHOR(S)
Frederick R. Morrell, Compiler

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, VA 23681-0001

8. PERFORMING ORGANIZATION REPORT NUMBER
L-17366

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Federal Aviation Administration
Washington, DC 20546 and
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA CP-3246
DOT/FAA/CT-94/03

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified–Unlimited

Subject Category-01

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report summarizes the research conducted during the academic year 1992–1993 under the FAA/NASA sponsored Joint University Program for Air Transportation Research. The year end review was held at Ohio University, Athens, Ohio, June 17–18, 1993. The Joint University Program is a coordinated set of three grants sponsored by the Federal Aviation Adminstration and NASA Langley Research Center, one each with the Massachusetts Institute of Technology (NGL-22-009-640), Ohio University (NGR-36-009-017), and Princeton University (NGL-31-001-252). Completed works, status reports, and annotated bibliographies are presented for research topics, which include navigation, guidance, and control theory and practice, aircraft performance, human factors and air traffic management. An overview of the year's activities for each university is also presented.

14. SUBJECT TERMS
Aircraft guidance; Avionics; Navigation and control; Human factors; Air traffic management

15. NUMBER OF PAGES
173

16. PRICE CODE
A08

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