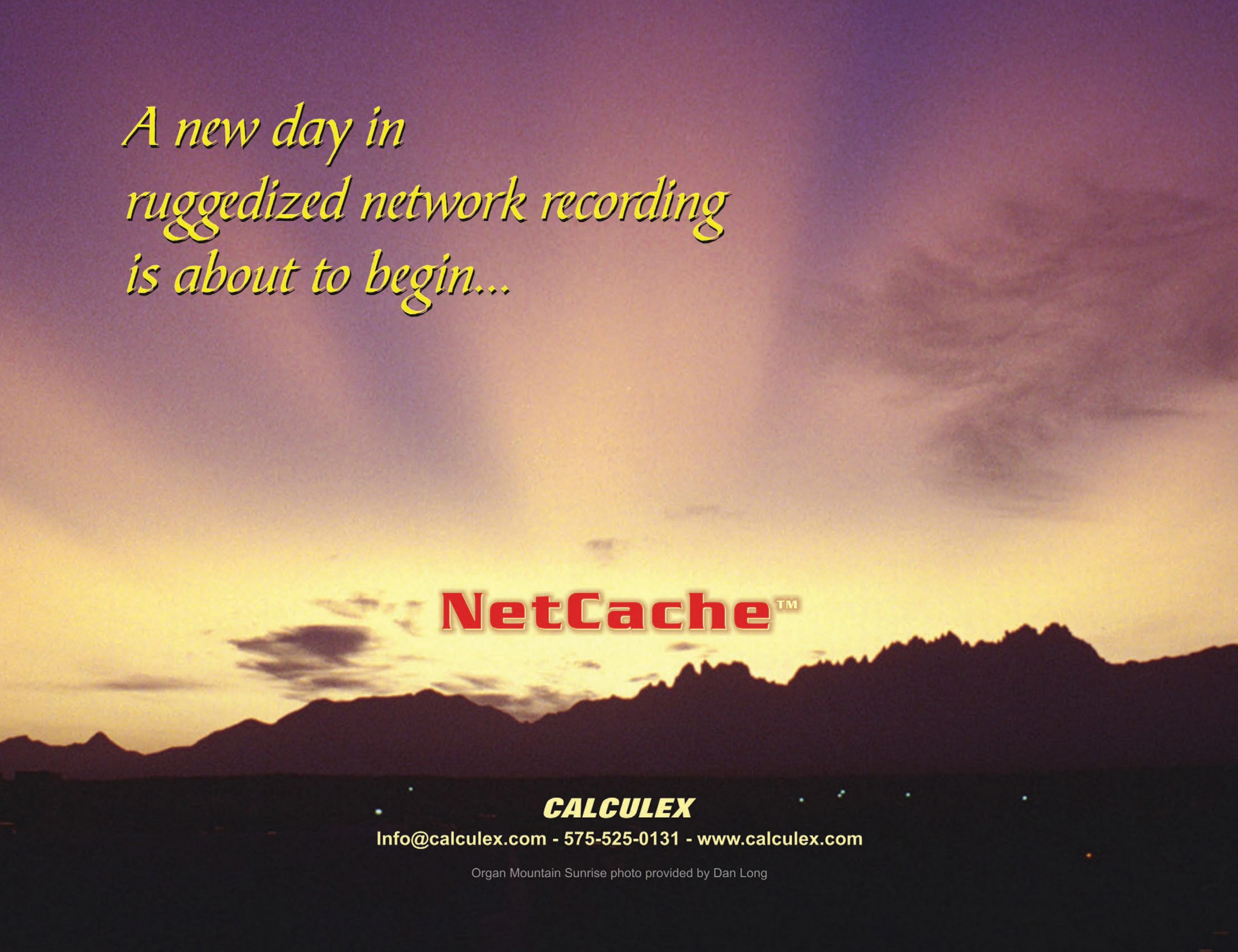
The ITEA Journal

December 2009
Volume 30, Number 4

Published quarterly
by the International
Test and Evaluation
Association

Air & Space


A new day in
ruggedized network recording
is about to begin...
NetCache™
CALCULEX
Info@calculex.com - 575-525-0131 - www.calculex.com
Organ Mountain Sunrise photo provided by Dan Long

CONTENTS

The ITEA Journal
2009
Volume 30, Number 4

TECHNICAL ARTICLES

DEPARTMENTS

ON THE COVER: One hundred and six years ago, December 17, 1903, Orville Wright piloted the first successful powered airplane while his brother, Wilbur, and a small crew watched. The flight lasted 12 seconds and covered 120 feet. Sixty-six years later, July 20, 1969, Neil Armstrong and Buzz Aldrin took the first steps on the moon while the world watched. The Apollo 11 mission lasted 195 hours, 18 minutes, and 35 seconds and covered somewhat more than 120 feet. Few things stimulate the imagination more than flight and space exploration and this issue celebrates both. The cover shows the first frame of the spectacular Earthrise sequence taken during Apollo 11 as Earth emerges over the lunar horizon. (Apollo 11 mission photograph courtesy of the National Aeronautics and Space Administration.)

■ ITEA Headquarters: 4400 Fair Lakes Court, Suite 104, Fairfax, Virginia 22033-3899; Tel: (703) 631-6220; Fax: (703) 631-6221, E-mail: *itea@itea.org;* Web site: *http://www.itea.org.*
■ ITEA is a not-for-profit international association founded in 1980 to further the development and exchange of technical information in the field of test and evaluation.
■ *The ITEA Journal* (ISSN 1054-0229) is published quarterly by the International Test and Evaluation Association at 4400 Lakes Court, Suite 104, Fairfax, Virginia 22033-3899. Single issue cover price for *The ITEA Journal* is $20. ITEA membership dues are $50 for individuals, $25 for full-time students, and $800 for corporations. Annual dues include a one-year subscription to *The ITEA Journal.* The annual subscription rate for libraries and other organizations providing timely reference material to groups is $60. All overseas mail (air mail or AOA) requires an additional $20. *The ITEA Journal* serves its readers as a forum for the presentation and discussion of issues related to test and evaluation. All articles reflect the individual views of the authors and not official points of view adopted by ITEA or the organizations with which the authors are affiliated.

POSTMASTER: Send address changes to: ITEA, 4400 Fair Lakes Court, Suite 104, Fairfax, Virginia 22033-3899.

ITEA Journal 2009; 30: 445–446

I was first exposed to ITEA in the fall of 1999 as the Atlanta chapter hosted the annual symposium. A fellow by the name of Jim Cofer, a Georgia Tech Research Institute coworker and ITEA board member, coaxed me into running the most demanding aspect of the annual event—the golf tournament! Of course, in my personal opinion, the Atlanta golf tournament still reigns as the best one yet. I met many of the forerunners of ITEA and was immediately hooked. Since I first became an ITEA member, the organization has blossomed. Individual and corporate membership has grown by twenty-five percent, corporate sponsors have doubled, the organization has added three chapters, and our revenue has increased by fifty percent from $862K to $1.3M.

I am thrilled to serve as your 17th President! Following my military career, I spent over 21 years immersed in the T&E community. My first T&E involvement was the development of a threat medium range acquisition radar system for the Army, which we incorporated as a test asset into the Electronic Combat Range (ECR) at China Lake. Also at ECR, I was involved in the hardware-in-the-loop surface-to-air missile simulation, "Missile on a Mountain", which is still actively employed for testing electronic counter measures. Later, I had the opportunity to develop an airborne threat simulator for the Air Force. Through these programs, I became very familiar with the ranges and how the services conduct T&E. My ITEA involvement started with the small symposium in Atlanta and my involvement with local events. After just a few short years, I was nominated and elected to serve on the ITEA Board of Directors (BoD) at the national level. Given that springboard, I was privileged to work with George Rumford, Scott Foisy, Mark Brown and many other dedicated volunteers as I chaired the Annual Technology Review Conference for four years. Now, as President, I commit to you that I will give this opportunity my concerted effort to continue to enhance this excellent T&E organization.

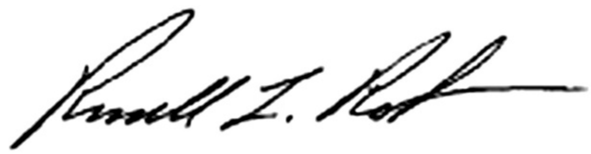
For those who do not know me, I am very much a people person. My Army experience taught me the value of caring for the troops and this has carried over in my civilian career. I have been blessed to work with so many highly motivated individuals within my organization (GTRI), within the T&E community and, in particular, within the ITEA constituency. That being said, I would like to focus my strengths during my tenure on the people aspects of ITEA: attracting new members and providing for the best possible environment for our volunteers and members. Allow me to mention a few of these highly motivated volunteers who are helping us to achieve our ITEA vision by concentrating on the "people". Mark Smith, ITEA BoD member, continues to accomplish new goals related to individual membership and increasing the "vibrancy" of ITEA local chapters. As Chairman of the Chapter and Individual Development Committee, Mark hosted a Chapter President meeting and an International Roundtable at the annual symposium in Baltimore that proved to be resounding successes. Most notably, he has successfully spearheaded the revitalization of the European Chapter. Along the same lines, recognition of our volunteers is paramount. I commend Denise De La Cruz for her stellar work chairing the Awards Committee and attracting nominations from all over the community, which culminated in the highly successful awards ceremony at our annual symposium. The list goes on, but allow me to mention one more volunteer—Jerry Sirmans. Jerry has gone way beyond just answering the volunteer bell. He has worked feverishly on both the Technology Review and the Annual Symposium. In addition, he reinvigorated his local chapter. Jerry said it best at the Annual Symposium as he accepted his well-deserved President's Award. He talked about the importance of volunteers in ITEA; he urged people to consider volunteering for the organization. What I will always remember is how he told the audience how rewarding it has been for him both personally and professionally. Jerry is the epitome of an ITEA volunteer.


Russell L. (Rusty) Roberts

As you read this edition of the journal, take note of the fact that the articles are from a diverse cross-section of the T&E community—the FAA, DHS, Korea, the Netherlands, as well as our core DoD constituency. Many of them are from our conferences (Live-Virtual-Constructive, Technology Review, and the Annual Symposium). This Journal content reflects our reaching out to other U.S. Government organizations as well as more active international participation, in addition to relative DoD information, which we rely on so heavily.

As I close my first article for the "corner", I would like to thank the outgoing President, John Smith. He has given his heart and soul to ITEA over the past two years as our leader. John will continue to serve on the BoD and we have tapped him to be the Chair for the Annual Symposium in Colorado Springs next year!

Feel free to contact me anytime with comments, questions, and/or suggestions; the BoD needs your feedback and I personally value your opinions. I look forward to seeing you at our many 2010 workshop offerings, beginning with the Live-Virtual-Constructive Conference in January, as well as the two ITEA national sponsored events: the Technology Review in Charleston, SC in July and the Annual Symposium in Colorado in September.



ITEA Journal 2009; 30: 447–453

Team of Strangers: Reflections on the Crew of Apollo 11

Michael H. Gorn, Ph.D.
NASA Dryden Flight Research Center, Edwards, California

As the world relives the events that culminated in the first human landing on the Moon in July 1969, the anniversary offers an opportunity to reflect on Apollo 11 with the perspective of the past 40 years. This article represents some provisional thoughts about Apollo 11 and its crew based on a portion of the author's present research on the early U.S. space program.

Far more often than not, the epic explorations in history—especially those entailing great risks—have involved intensive collaborations, depending not only on painstaking planning and the instruments of science but also on the shared sense of unity and purpose among the participants. A famous example of one such expedition is found in the perilous travels of Meriwether Lewis and William Clark (1803–1806). These explorers, appointed by President Thomas Jefferson, commanded the self-described Corps of Discovery to penetrate and catalog the interior of the North American continent. Another historic journey involved the Norwegian Roald Amundsen, whose company became the first to reach the South Pole (1910–1912). More recently, in 1953, Edmund Hilary and Tenzing Norgay led their party to the summit of Mount Everest, the first to ascend the highest peak on Earth. All three groups, and many besides, relied on the intangible but potent factor of personal cohesion to achieve their objectives.

Curiously, the Apollo 11 mission, as dangerous and ambitious as any in history, does not conform to this pattern. If anything, the extraordinary success of the first Moon landing has obscured the unusual atmosphere that pervaded the crew. Rather than the product of a band of brothers, like so many of the great explorations before it, the success of the Apollo 11 mission might be called the triumph of a team of strangers.

Despite outward appearances of harmony and similarity (for instance, the coincidental birth of all three Apollo 11 crew members in 1930), Neil Armstrong, Michael Collins, and Buzz Aldrin possessed perhaps the most dissimilar temperaments ever packed into a space capsule—or into any other place, for that matter. In the end, they functioned effectively; no one can question the result. Their actions minutely scripted from the ground and extravagantly rehearsed, the Apollo 11 astronauts succeeded in achieving their historic mission through an exquisitely crafted technological process—a process, nonetheless lacking in the fraternal spirit of so many of the earlier explorers.

In fact, even in contrast to the other Apollo crews, the Apollo 11 astronauts never really united. They arrived at work in separate cars, and left the same way. Each took lunch by himself *(Figure 1)*. Rather than experiencing collective exhilaration, the three men found almost no common ground on which to savor their historic adventure. At the supreme moment of touchdown on the lunar surface, one they had practiced for years, Aldrin and Armstrong merely patted each other on the shoulder and shook hands, and then went about their tasks without any outward signs of relief, joy, or emotion.[1]

The explanation for this seemingly counterintuitive behavior may be found, in part, in the lives of the Apollo 11 astronauts themselves.

Like most of the early astronauts, Neil Armstrong grew up in small towns—in his case, many small towns. From his birth on August 5, 1930, until the age of 14, he and his family moved 16 times, their changes of address representing a travelogue of flat and rural northwest Ohio, an area of productive agriculture and high achievers. The Armstrongs eventually settled in Wapakoneta. His father, Stephen, worked for the state of Ohio as an auditor, and as a civil servant, he remained employed throughout the Great Depression.

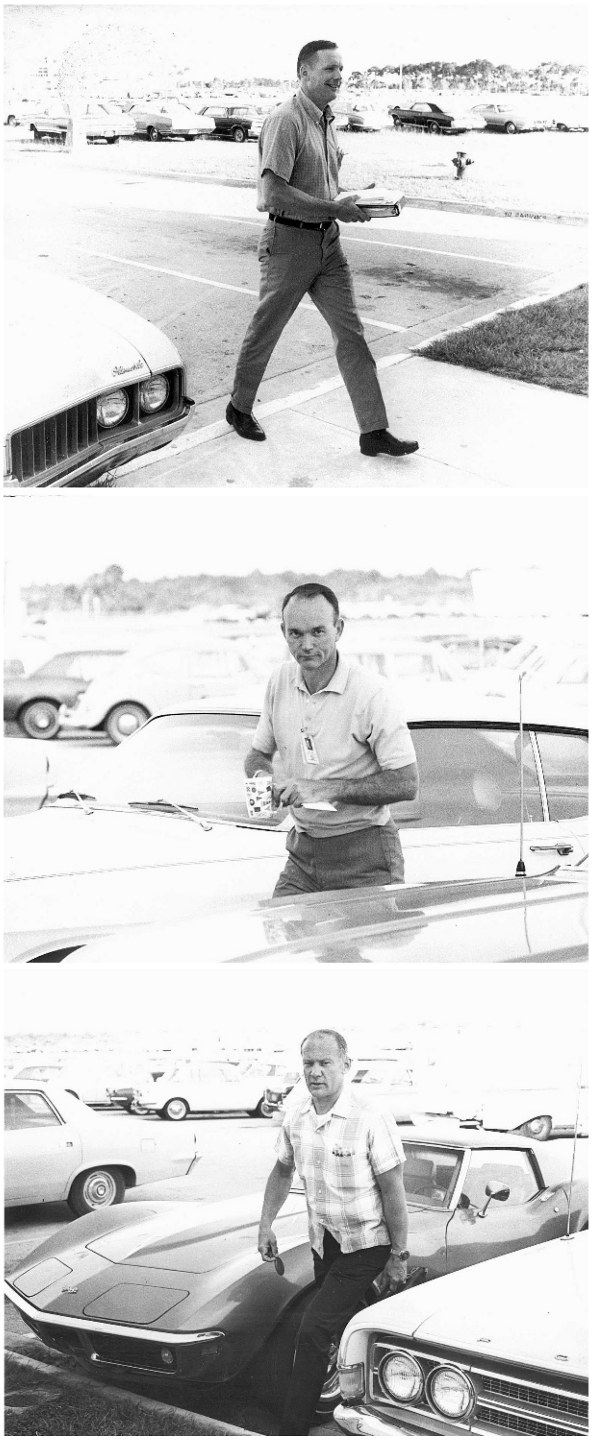


Figure 1. The three Apollo 11 astronauts photographed separately as they arrived at Kennedy Space Center's Flight Crew Training Building a week prior to the launch to the Moon. Top to bottom: Armstrong, Collins, and Aldrin. [NASA/Courtesy of nasaimages.org]

As a boy, Neil worked cutting grass at a cemetery and cleaning vats at a bakery. He became active in the Boy Scouts of America. He read voraciously and graduated from high school in the top 10 percent of his class, taking great pleasure from science and from playing horn in the school band. He seemed to be at once friendly *and* reserved, confident *and* modest, tolerant *and* stubborn. Most importantly, when faced with decisions, he based them on his own reasons and instincts, not on external advice or consultation. He never wavered from these inclinations. Even as a child, Armstrong's contradictory and paradoxical personal qualities caught the attention of his father; his pious mother, Viola (Engel); and his friends.[2]

Neil Armstrong followed the path of many young men born in the age of Charles Lindbergh's epic flight across the Atlantic Ocean in May 1927. He liked toy airplanes as a very small child, took his first plane ride at age six, became absorbed in model making around age eight, took flying lessons at age 16, and obtained his pilot's license before he could drive an automobile. He entered Purdue University's School of Engineering in 1947 on a U.S. Navy seven-year scholarship: two years of college, three years in the Service, and two more years of college to complete his degree. After his first segment at Purdue, he reported to Naval Air Station Pensacola, Florida, hoping to become a combat pilot. Armstrong passed the physical training, as well as the standard academic and flight school curriculum. He proved to be diligent, self-critical, and intent on improving his piloting techniques. As a result, in August 1950, Pensacola designated him a naval aviator.[3]

Armstrong's first assignment placed him with Fighter Squadron 51 aboard the carrier *Essex* in San Diego, California, where he qualified for carrier landings in June 1951. By August of that year the *Essex* had joined Navy Task Force 77 off Wonsan, Korea.

Respected for his precision as a pilot, but also known for his intellectual inclinations and bookish nature, he flew for the first time over North Korea on August 29. Five days later, while striking targets in hilly terrain, he barely avoided catastrophe when his F9F-2 Panther suffered damage from anti-aircraft fire. He lost elevator control, the aircraft dropped with the weight of its heavy ordnance, and his plane struck a power pole, chopping off 2 feet of the right wing. Although his life hung in the balance, Armstrong's cool demeanor did not fail him. He worked his controls deftly, flew to safety, ejected, and survived. He went on to complete 78 missions during his tour in Korea.[4]

Ensign Armstrong and Fighter Squadron 51 left the Far East for Hawaii in March 1952. He mustered out of the Navy the following August but remained in the Reserves until 1960. Meanwhile, he returned to Purdue University and completed his engineering studies in January 1955. Armstrong then considered a number of opportunities: graduate school, a career with the Navy, flying for Trans World Airlines, and test piloting at Douglas Aircraft. He chose none of these, and instead applied for a research pilot position with a federal agency familiar mainly to those in the aviation field, known as the National Advisory Committee for Aeronautics (NACA), predecessor of the National Aeronautics and Space Administration (NASA). All of the other prospects paid better salaries, but Armstrong preferred the combination of research, flying, and complex engineering projects offered by the NACA. Accordingly, he accepted an offer of employment at the Lewis Engine Propulsion Laboratory in Cleveland, Ohio.

After working for five months at Lewis on a diverse portfolio that included aircraft icing and rocketry research, he transferred to the NACA's High-Speed Flight Station on Edwards Air Force Base, located in the Mojave Desert of Southern California. Here, eight years earlier, U.S. Air Force Captain Chuck Yeager had pushed the Bell X-1 over the sound barrier. Just 24 years old when he arrived in California, Armstrong joined an elite circle of pilots, becoming the youngest on the staff. In time, he distinguished himself not so much for his skill with stick and rudder as for his engineering acumen, probably unequaled among his peers.

During his seven years at Edwards Air Force Base, Armstrong made over 900 flights. He flew more than 100 times in the cockpits of the heavy bombers that the NACA used as launch platforms for its research aircraft. He flew the famed Century series fighters (F-100, F-101, F-102, F-104, F-105, and F-106) over 350 times in experimental flights. He flew exotic research aircraft, including the X-1B, the X-5, and the famous X-15 hypersonic rocket plane (called the "Black Bull" by fellow NACA research pilot Milt Thompson in recognition of its power and unpredictability). The X-15 gave Armstrong his first intimation of spaceflight. He piloted it seven times between November 1960 and July 1962, once to an altitude of 207,500 feet (almost 40 miles), a flight that also turned out to be of the longest duration (12 minutes and almost 30 seconds) of any X-15 mission. On a different flight, his X-15 attained speeds as high as Mach 5.74.[5]

Two days before Armstrong's 40-mile-high X-15 adventure in April 1962, NASA announced openings for its second astronaut class. He weighed his options. If he stayed at Edwards, he would continue to fly the challenging X-15. He had also been selected as a future pilot for the Air Force's Dyna-Soar spacecraft, which had yet to be developed. Deciding to pursue the astronaut training program, Armstrong applied to NASA. The application arrived past the deadline, but friends who had transferred earlier from the High-Speed Flight Station to the Manned Spacecraft Center in Houston, Texas, intervened, allowing his application to be accepted. Characteristically, he revealed his entry into the astronaut sweepstakes to no one at Edwards, not even those closest to him. Yet, when the media announced the selection of Astronaut Group 2 on September 17, 1962, his coworkers greeted the news of his successful entry to the program enthusiastically and felt sure that he had the "right stuff"—broad experience in the cockpit, ample engineering talent, and composure under duress.[6]

If Neil Armstrong's life began in small-town America, Mike Collins grew up in the U.S. Army tradition. His father, Major General James L. Collins, served as an aide to General John J. Pershing and received the Silver Star for service in World War I. Mike's uncle J. Lawton (Joe) Collins commanded one of General Dwight Eisenhower's Army corps during World War II and after the war became Army chief of staff. His mother, Virginia—a bibliophile who loved language and transmitted these passions to her son—had a decisive influence on his early life and preferences.

Collins matured in the shadow of his father's prestigious military postings, a distant world from the provincialism of small-town life. Born on October 31, 1930, near the Via Veneto in Rome, Collins lived subsequently on Governor's Island in New York Bay, where he became acquainted with Manhattan, and at Fort Hoyle, near Baltimore, Maryland, and the Chesapeake Bay. In San Juan, Puerto Rico, the family resided in a venerable old colonial home, where young Mike immersed himself in the study of the animals and plants of the tropics. Around that age, he said, "I discovered books…and it was just like somebody had turned on the light or opened the door." The family moved to Washington, D.C., during World War II, and he attended prestigious St. Albans Episcopal School. He then went to the U.S. Military Academy at West Point, and although he found the experience confining because of his sunny temperament and wide-ranging interests, Collins graduated in 1952, 185th among 527. He chose an Air Force career (in part to avoid any possible hint of nepotism in the Army) and in quick succession took advanced day fighter training in Nevada, where he flew simulated combat against MiG fighter planes; attended a nine-month aircraft maintenance officer course at Chanute Air Force Base in Illinois; and, after completing Test Pilots School at

Edwards Air Force Base, won a coveted assignment to Flight Operations at the Air Force Flight Test Center, also on Edwards Air Force Base.

Outwardly, Mike Collins may have seemed like many a rising Air Force star, but because of his catholic upbringing and intellectual tastes, he represented a different type than most of his contemporaries. Gregarious and witty, articulate, a skilled writer, and a graduate of the Test Pilots School besides, Collins exuded the easygoing charm of a natural leader. His future wife, Patricia Finnegan of Boston, saw this potential in the playful, mischievous, yet sophisticated young officer whom she met in France:

> *"He loved to eat and he understood French food and introduced me to a lot of things that I wouldn't have tried by myself. Some of the Americans were joining wine study clubs, but Mike didn't have to. He learned a lot about wines from his father, and he had studied up on vineyards so he could tell wonderful stories about them. In fact, he could talk about anything. He knew books; he knew poetry; he was interested in theater. He was bright about technical things, and he was lots and lots of fun. I couldn't get over all this combination in one man."*[7]

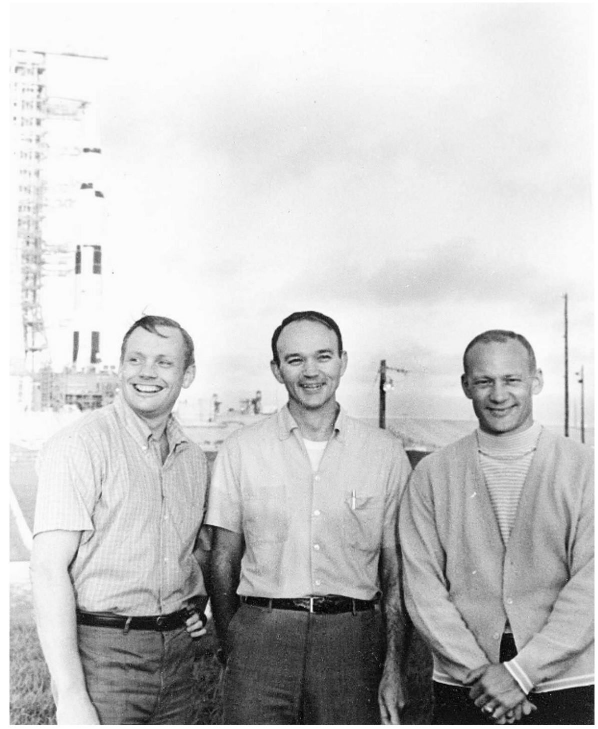


Figure 2. Before their mission, Armstrong, Collins, and Aldrin (left to right) posed for a casual photograph at Kennedy Space Center with a Saturn V launch vehicle looming in the background. [NASA/Courtesy of nasaimages.org]

Among the three Apollo 11 astronauts, only Edwin E. Aldrin Jr. could claim aeronautics as a birthright. His father, Edwin E. Aldrin Sr., saw action in World War I in the Army Air Service as a pilot and aide to the famed General Billy Mitchell. He eventually rose to the rank of colonel and earned a doctor of science degree from the Massachusetts Institute of Technology. The elder Aldrin also worked as a Standard Oil executive; later, he became an independent aviation consultant and knew many of the eminent figures of early flight, such as Jimmy Doolittle, Charles Lindbergh, and Howard Hughes. Edwin Sr.—almost impossible to please and unfailingly demanding—played a pivotal role in his son's life by establishing an all but unattainable standard of success *(Figure 2)*.

Like his father, Buzz Aldrin showed ample competitiveness and ambition. (Aldrin legally adopted the name Buzz in 1988. It originated with his sister Fay Ann, whose attempts as a child to pronounce "brother" sounded more like "buzzer.") In fact, Mike Collins described the adult Buzz as "a very intense, goal-oriented individual, accustomed to winning big—and not losing at all." Just two years after his birth on January 20, 1930, he got a taste of his father's love of aviation when the senior Aldrin flew him and one of his sisters to Florida aboard a Standard Oil Lockheed Vega. In his native Montclair, New Jersey, young Aldrin attended the local schools but did not distinguish himself in the early years. Then, from ninth grade on, and especially at Montclair High School, he excelled in math, science, and football. His choice of a college did not evolve haphazardly. Influenced in part by his father's military career, as well as by the example of his older sister's boyfriend (a West Point graduate), he set his heart on the Army's Service academy. He won admission and reported on July 1, 1947.[8]

Aldrin thrived under the regimentation and discipline of West Point, feeling at ease in a system of unambiguous scholastic and personal targets.

> *"I fit in well. To me, the beauty of that system was you knew exactly where you were at all times. You could measure your progress. You knew exactly what was expected of you from day to day."*

Although he graduated very high (third) in his class, he left with a feeling of bruised idealism after reporting a cheating incident in which he had good reason to think the perpetrator had received inadequate punishment. Sincere and purposeful, Aldrin found this episode disappointing, even deflating. Nonetheless, he decided—again with his powerful father's approval—to enter the U.S. Air Force. As a second lieutenant, he went to basic flight school in Bartow, Florida, in early

1951, scored well, and attended fighter school in Bryan, Texas. At Nellis Air Force Base near Las Vegas, Nevada, Aldrin learned to fly the Air Force's front-line F-86 Sabre fighter as well as the F-80 Shooting Star. His training culminated at the end of 1951 when he disembarked in Seoul, Korea, for combat, assigned to the 51st Fighter Wing. In the year and a half before the armistice ending the Korean War, he flew 66 combat missions in the F-86 and shot down two MiG-15s.

During the years immediately after the war, Buzz Aldrin's career progressed, but without high distinction. He taught aerial gunnery at Nellis Air Force Base and became a flight instructor in the F-100 Super Sabre fighter at Bitburg, Germany. But eventually he followed his father's example once more and applied for admission to the Massachusetts Institute of Technology in the doctorate of science program, where he studied from 1959 to 1962. This decision proved to be a pivotal one because he pursued a research subject that foreshadowed his career with NASA. His dissertation—which pioneered orbital approach and rendezvous techniques—had important practical applications to and benefits for the U.S. space program as a whole, for the success of the Gemini and Apollo space missions in particular, and for his own career as well.[9]

* * * * *

For four years, Neil Armstrong trained for his first foray into space, which turned out to be as command pilot aboard Gemini VIII. The mission proved in part momentous and in part ill omened for Armstrong. Begun on March 16, 1966, it paired him with copilot David Scott in an effort to achieve the first space docking of two separately launched vehicles. They succeeded in half of their mission, that of mating the Gemini capsule to the Agena target. But having joined the two pieces together, the combined spacecraft soon fell into a violent tumble and yaw. Even after Armstrong decided to disconnect the two halves, the Gemini portion continued to roll, pitch, and yaw even faster than before. Fearing the loss of consciousness if he and Scott continued to spin, Armstrong characteristically kept his head and made a life-saving choice: he disabled the orbital attitude maneuvering system and relied instead on the re-entry control system thrusters. This step ended the crisis, and Armstrong's judgment positioned him well for greater responsibilities.

Having shown his mettle in Gemini, he took his initial step into the lunar program on the Apollo 8 backup crew. This placement poised him for command of Apollo 11. But Deke Slayton, chief of the Astronaut Office, had made it clear earlier that the first Moon walk would happen on *Apollo 12*, not Apollo 11. Armstrong, therefore, seemed destined to be in charge of *an* Apollo mission, not *the* Apollo mission. Then chance and circumstance intervened. In the aftermath of the Apollo 1 fire in January 1967 and subsequent investigations, mission planners had no choice but to delay the timetable for the Moon landing. But they made up for ten lost months by compressing the flight schedule, adding a circumlunar mission to Apollo 8 in order to hasten the project. If this launch happened as planned, and Apollo 9 and 10 also flew without serious problems, then *Apollo 11*, not Apollo 12, would become the brass ring of the Moon program, giving Armstrong the opportunity of a lifetime.[10]

The fate of Apollo 11 came into sharper focus as the Apollo 8 circumnavigation mission approached the Moon on December 23, 1968, at which time Armstrong met with Slayton regarding the composition of his crew. Slayton suggested Buzz Aldrin and Mike Collins, adding that if Aldrin seemed problematic, James Lovell would also be available. Armstrong, as always, thought carefully, deciding in the end that despite Slayton's concerns about Aldrin's temperament—a concern Armstrong did not share, having worked with him when they both backed up Apollo 8—he would select Buzz. In Armstrong's mind, the more serious choice involved Collins and Lovell. Armstrong decided to by-pass Lovell (who had been commander of Gemini XII and command module pilot of Apollo 8), reasoning that his experiences had earned him a command of his own. So, the crew of Apollo 11 materialized: Commander Neil Armstrong, Command Module Pilot Mike Collins, and Lunar Module Pilot Buzz Aldrin.[11]

Thus, through the NASA selection process—an elusive cocktail of seniority, experience, personal capabilities, and (last, but far from least) luck—these three unlikely teammates found themselves bound together in one of the epic odysseys in world history *(Figure 3)*.

* * * * *

Of course, no one can argue with the outcome of Apollo 11. In the end, Armstrong, Aldrin, and Collins collaborated sufficiently to accomplish their objective *(Figure 4)*. Mike Collins correctly pinpointed one of the main sources of their success, although with exaggerated modesty. He called Armstrong and Aldrin

> *"smart as hell, both of them, competent and experienced, each in his own way. Neil was far and away the most experienced test pilot among the astronauts, and Buzz the most learned….But that didn't mean that Neil was all pragmatist or Buzz all theoretician. Among the dozen test pilots who had flown the X-15 rocket ship, Neil had been considered one of the weaker stick-and-rudder*

Figure 3. Above, the iconic official portrait of the Apollo 11 crew taken before their historic mission. Left to right: Armstrong, Collins, and Aldrin. NASA/Courtesy of nasaimages.org. Below, in a less heroic pose following a day of simulator training, the Apollo 11 crew stands in front of a lunar module mockup in the Kennedy Space Center Flight Crew Training Building. Collins is on the far left. [NASA/Courtesy of nasaimages.org]

> *men, but the very best when it came to understanding the machine's design and how it operated. Buzz, far from being an academic recluse, was an outstanding athlete, a pole-vaulter, and a Korean War MiG killer. As a pair, they brought a formidable array of talent with them. I considered myself damn fortunate to be joining them.*"[12]

Still, the Apollo 11 crew remained faithful to its strangely atomized personality. At least in part, the lack of fraternal unity resulted from their mismatched

Figure 4. Ticker-tape parade in New York City. About 4 million New Yorkers lined the Financial District, Broadway, and Park Avenue on September 13, 1969, to hail the three heroes of Apollo 11 and their accomplishment. [NASA/Courtesy of nasaimages.org]

temperaments. But in addition, the intense pressures associated uniquely with Apollo 11 may have contributed to the distance among Armstrong, Aldrin, and Collins *(Figure 5)*. Just as Project Apollo constitutes a high point in the annals of human exploration, the first touchdown on the Moon stands at the apex of the other lunar missions. Those that took place before, while indispensable, had less at risk. Those that came after flew with the certainty that President John Kennedy's quixotic vision had actually been achieved. Not surprisingly, all of the missions preceding and following Apollo 11 developed their own brand of camaraderie and esprit de corps. But the Apollo 11 astronauts—subjected to the cauldron of relentless public focus and fascination, the demand for success, and the sickening possibility of failure—never did jell. This difference went virtually unnoticed in the high drama of the event. Lacking the band-of-brothers mentality, the crew members of Apollo 11 relied instead on NASA's intensive technical planning and on elaborate practice drills, on their own commitment to the mission, and on their personal sense of duty. In the end, they succeeded *despite* the lack of a common emotional bond, the presence of which has sustained so many of the great expeditions of the past.[13] ❑

Michael H. Gorn, PhD, has worked as a public historian for more than 30 years in the National Aeronautics and Space Administration, the Department of the Air Force, and the Environmental Protection Agency. He is the author of many books about the history of aeronautics and spaceflight, most recently NASA: The Complete Illustrat-

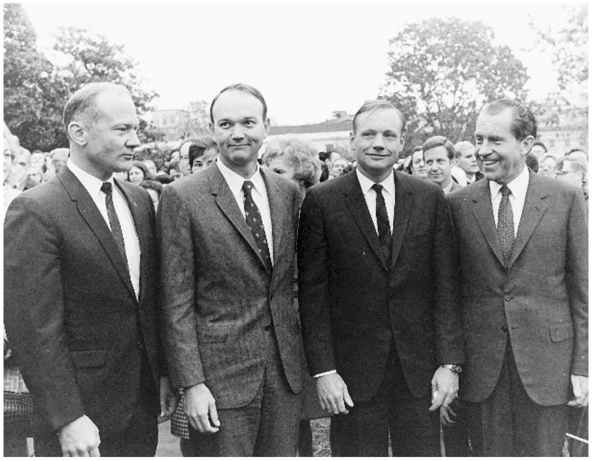


Figure 5. After an exhausting 45-day tour to 24 countries, Aldrin, Collins, and Armstrong (left to right) pose awkwardly with President Richard M. Nixon (far right) on the White House lawn. [NASA/Courtesy of nasaimages.org]

ed History *(Merrell, 2008) and* Superstructures in Space: From Satellites to Space Stations, A Guide to What's Out There *(Merrell, 2008). He is also the author of* Expanding the Envelope: Flight Research at NACA and NASA *(University Press of Kentucky, 2001), winner of the 2004 Gardner-Lasser Aerospace History Literature Award, presented by the American Institute of Aeronautics and Astronautics. E-mail: (care of) rjanssen@allenpress.com*

Endnotes

[1]For the title of this article, "Team of Strangers," I am indebted to Doris Kearns Goodwin's *Team of Rivals: The Political Genius of Abraham Lincoln* (New York: Simon and Schuster, 2005). I also owe a debt to James R. Hansen's remarkable biography of Neil Armstrong, entitled *First Man: The Life of Neil A. Armstrong* (New York: Simon and Schuster, 2005). In particular, chapter 24, "Amiable Strangers" (pp. 335–59) informed this article (see especially pp. 358–59, related to the lack of team camaraderie among the Apollo 11 crew, and pp. 474–75 for the reaction of Aldrin and Armstrong to the lunar landing). Also, see Buzz Aldrin's recollection of the moment of landing in Edwin E. Aldrin and Wayne Warga's *Return to Earth* (New York: Random House, 1973), p. 231–232. The story of the landing is re-told more recently, with slight revision, in Buzz Aldrin with Ken Abraham, *Magnificent Desolation: The Long Journey Home from the Moon* (New York: Harmony Books, 2009), pp. 21–22.

[2]Hansen, *First Man*, pp. 15, 29, 30–31, 33–34, 37–38, 40, 41.

[3]Hansen, *First Man*, pp. 45–46, 48, 50, 54, 67–79.

[4]Hansen, *First Man*, pp. 79–114.

[5]Hansen, *First Man*, pp. 117–59; Richard P. Hallion and Michael H. Gorn, *On the Frontier: Experimental Flight at NASA Dryden*, Appendix M: X-15 Program Flight Chronology, 1959–1968 (Washington, DC: Smithsonian Books, 2003), pp. 407–8.

[6]Hansen, *First Man*, pp. 168–203.

[7]Michael Collins, *Carrying the Fire: An Astronaut's Journey* (New York: Cooper Square Press, 2001), pp. 6–17; Hansen, *First Man*, pp. 344–46 (quoted sentence, p. 345; quoted paragraph, p. 346).

[8]I am indebted to *Return to Earth* for this section about Aldrin's career prior to his selection as an astronaut. See chapters 4 and 5. Quoted sentence: Collins, *Carrying the Fire*, p. 460.

[9]Hansen, *First Man*, pp. 348–50; quoted paragraph: Aldrin, *Return to Earth*, p. 111.

[10]Hansen, *First Man*, pp. 242–63, 312–13, 335–37.

[11]Hansen, *First Man*, pp. 230–31, 338–39.

[12]Quoted paragraph: Collins, *Carrying the Fire*, p. 314. For the vagaries of the astronaut selection process, see Hansen, *First Man*, pp. 230–31.

[13]For a comparison of the Apollo 11 crew with the other crews in the Moon program, see Hansen, *First Man*, pp. 358–59. The prevailing atmosphere of Apollo 11 is epitomized by the curious lack of affect in Neil Armstrong's remark when he and Aldrin touched down on the Moon: "So far so good. OK, let's get on with it." Hansen, *First Man*, p. 475.

Acknowledgments

The author wishes to acknowledge Dr. Von Hardesty of the National Air and Space Museum for his contributions to this article.

ITEA Journal 2009; 30: 455–460

A Fiber Optic Seismic Sensor for Unattended Ground Sensing Applications

Joseph Dorleus, Ph.D.
U.S. Army Program Executive Office for Simulation, Training, and Instrumentation, Orlando, Florida
Yan Zhang, Ph.D., Jing Ning, Ph.D., Thomas Koscica, Ph.D., Hongbin Li, Ph.D., and H. L. Cui, Ph.D.
Stevens Institute of Technology, Hoboken, New Jersey

Fiber optic seismic sensors have been increasingly recognized as promising technologies for many applications, such as intruder detection and perimeter defense systems. Among these, a military seismic sensor is especially challenging because it requires a robust, compact, reliable, easily installable and operated product. This article reports on our recent experimental investigations of a military fiber optic seismic sensor. In particular, an improved sensor design and new signal processing techniques are illustrated in this article, which include: (a) a scanning laser wavelength-based demodulation system; (b) a digital lock-in amplifier and field programmable gate array techniques for weak signal detection and processing; and (c) overall improved seismic sensitivity based on carbon fiber composite cantilever and fiber-Bragg-grating sensing technology. The experimental results show that the fiber-Bragg-grating seismic sensor frequency response band is between 10 and 300 Hz and that it has a higher sensitivity than the conventional electromagnetic geophone within the frequency response band most important for battlefield monitoring. The fiber-Bragg-grating seismic sensor is totally immune to electromagnetic interference and has a working temperature between −40°C and +50°C for military harsh environments. Comparisons are also presented of the fiber-Bragg-grating seismic sensor with the commercial REMBASS sensor in a series of field tests.

Key words: Fiber-Bragg-grating seismic sensor; unattended sensor; intruder detection; perimeter defense; battlefield monitoring; sensors; signal detection; electromagnetic geophone; laser-based optical demodulation.

Fiber optic sensor technology has been and is being increasingly exploited by the research community because of its relatively simple design, low power consumption, low cost, relatively low maintenance cost, and the flexibility it offers for both commercial and military applications. In particular, fiber optic seismic sensors have been recognized as promising technologies for numerous applications, which include intruder detection and perimeter defense systems for military applications. However, seismic military sensors are required to be robust, reliable, compact, and easy to install and operate to be effective in the battlefield environment. This article addresses the challenges presented by the design of military seismic sensors and reports on experimental investigations of fiber optic seismic sensors. Specifically, we address three types of sensor technologies that provide improved design and novel signal processing techniques: (a) a wavelength scanning, pulsed-laser–based demodulation system, (b) digital lock-in amplifier and field programmable gate array (techniques) for weak signal detection and processing, and (c) improved seismic sensitivity based on carbon fiber optic composite cantilever and fiber-Bragg-grating (FBG). We present our experimental results on the FBG seismic sensor frequency response along with comparisons of those with the REMBASS II sensor (Remotely Monitored Battlefield Sensor System II) resulting from field tests.

We have developed a new scheme of laser-based optical demodulation with excellent results. At 100-Hz scanning speed we were able to demodulate the wavelength down to 1.1 pm. The signal is completely digital, clear, with excellent signal-to-noise ratio. At lower scanning speeds we can achieve subpicometer wavelength resolution. The dynamic following range is about 10 nm with a dynamic range of 120 dB.

Principle and design

The design of the seismic sensor is illustrated in *Figure 1*. The detection is implemented by the FBG dynamic strain sensor, which is attached on a spring-mass system (Pastore, 2006). The acceleration of ground motion is transformed into strain variation on the FBG sensor through this mechanical design and, after the optical demodulation, generates the analog voltage output proportional to the strain changes. By adjusting the mechanical parameter of the spring–mass configuration, we can mechanically tune the natural response frequency of the system within a certain range in adapting to the different frequencies of seismic wave sources (signals of personnel and vehicles). This sensor head has a compact size and a mass of only a few grams. It has the advantage that the sensor can be easily embedded and hidden in the battlefield without any radio frequency emission or thermal signature to the environment. The optical fiber–based sensor itself is resistant to corrosion, high temperature, and fatigue, and so is suitable for deployment in the harsh environment of battlefield. A damping mechanism is incorporated into the design with critical damping provided so that the mass–spring system will return quickly to its ready state after detecting a signal. This damping mechanism is shown in *Figure 1*, where a Faraday induction loop and a permanent magnet provide the damping. The small induced current in the Faraday loop is properly sealed so that no electromagnetic signal will go in or out of the sensor head. By using the carbon fiber composite cantilever, the overall sensor performance is improved, which leads to a higher sensitivity, better linearity, and smaller weight on the sensor head.

Considerations for field applications

For rough environments, such as the battlefield environment, two things are of serious concern. First, the sensor needs to be strictly waterproof, dustproof, and adaptable to vast temperature changes in excess of −40°C to +60°C. Two of the advantages of a fiber optic device are that the device is intrinsically waterproof and dustproof, but it is also very sensitive to temperature changes. Our sensor uses a pair of matched FBGs; one is the sensing element and the other is the demodulator. A thermostat and a

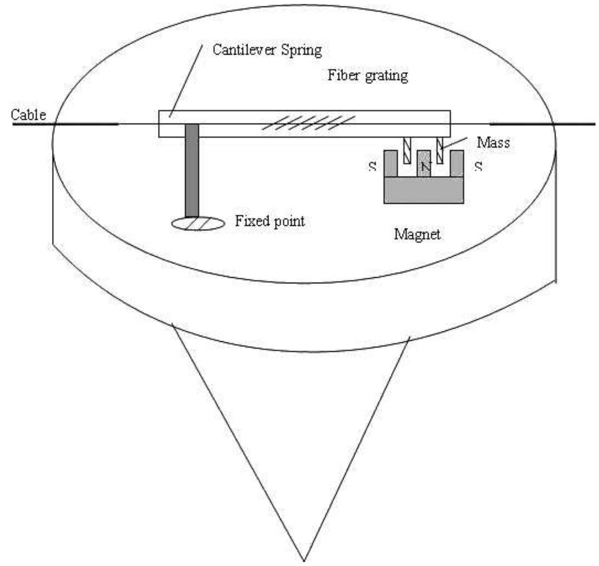


Figure 1. Basic structure of the FBG sensor head.

temperature control make sure that the two FBGs are always kept at the same temperature.

Power consumption is another consideration for unattended system design. We employ a solar cell on the top cover of the sensor and a high-density battery inside the sensor. A digital control block permits cycling of wake–sleep configuration. Sleep time is set as long as possible based on average current consumption, and the ratio of wake–sleep is set depending on the required reliability of detection, alarm level, and the states of neighboring sensors.

Improvement of signal detection

One of the most profound challenges in the design of the seismic sensor is acquiring and separating weak seismic signals from strong background noise (for example, wind); another is distinguishing signals from different sources. Basic requirements for a practical unattended seismic sensor are that it has to be very sensitive to small vibrations but also have a proper strategy to decrease the false alarm rate (FAR). To solve these problems, we designed an original algorithm. We illustrate the workings of this algorithm by using the example of detecting the signals of a human walking. First, the system records the seismic response of a person walking, saving the basic digital model on the sensor. Second, the system compares the sensor's subsequently acquired seismic signal with the basic digital model using a correlation operation. Third, the system estimates the degree of correlation. If the degree of correlation is higher than a preset level, it is counted as a signal impulse. Finally, the system sums the number of impulses in a given time interval. If this sum matches the rate of previous signals, the sensor

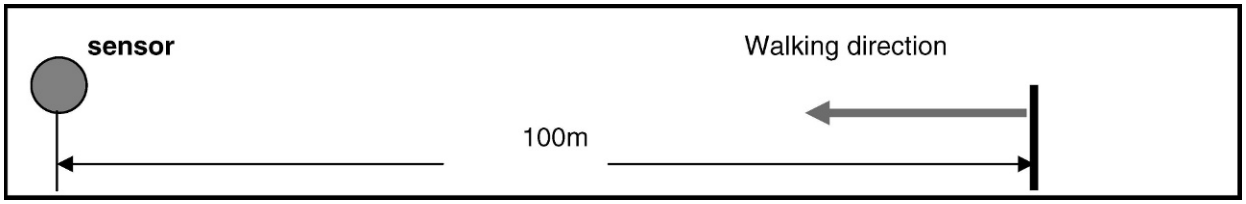


Figure 2. Experimental setup.

initiates an alarm. Both improved signal detecting ability and decreasing false alarm rate are achieved with this algorithm.

We have carried out a preliminary experimental study to investigate the applicability of our algorithm. Both a fiber-optic Bragg-grating geophone and an electromagnetic geophone were used for this work. *Figure 2* shows the experimental setup with a person walking directly to the geophone from 100 m away. *Figure 3* (top) shows the geophone response without any signal processing, where we could only see a few high-quality signal cycles. *Figure 3* (bottom) shows that more cycles of the walking signal were detected by our processing method of signal correlation even though the signal was buried under noise.

Experiment results

The sensitivity test of the FBG seismic sensor was carried out on a vibration stage that is driven by a series of sinusoidal waves. We use the conventional electromagnetic geophone for comparison. *Figure 4* is the result of the FBG seismic sensor and the geophone response to the vibration at 30 Hz. Comparison tests at different frequencies showed that the FBG seismic sensor has a higher sensitivity than the conventional electromagnetic geophone—between 10 and 70 Hz. The frequency analysis results showed that the FBG seismic sensor frequency response band is between 10 and 300 Hz, which covers the major frequency band of the military seismic signals.

The electromagnetic interference test of the FBG seismic sensor was carried out and compared with the electromagnetic geophone *(Figure 5)*. Both geophones were initially still and placed near an AC current supply (power: 80W; frequency: 60 Hz) at equal distances of 20 cm. The conventional electromagnetic geophone was severely disturbed by the output signal at 60 Hz. The FBG seismic sensor did not exhibit

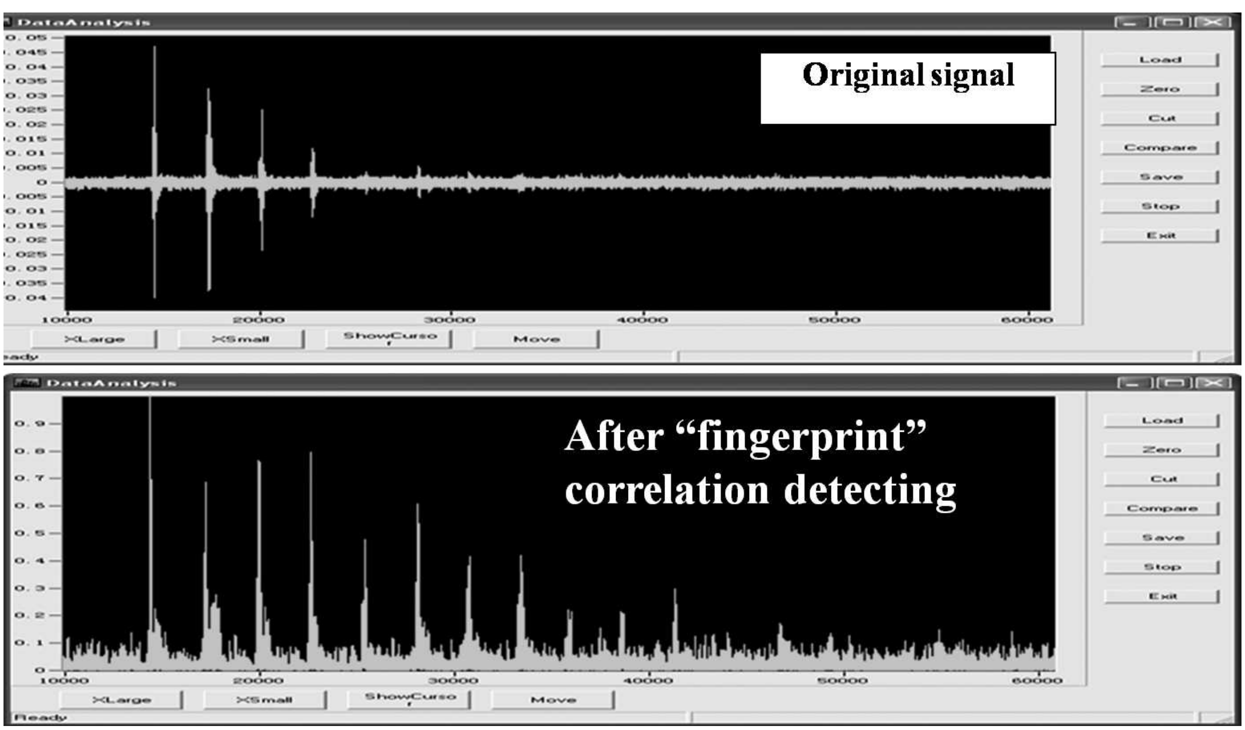


Figure 3. (top) Geophone response without any signal processing; (bottom) geophone response with signal processing.

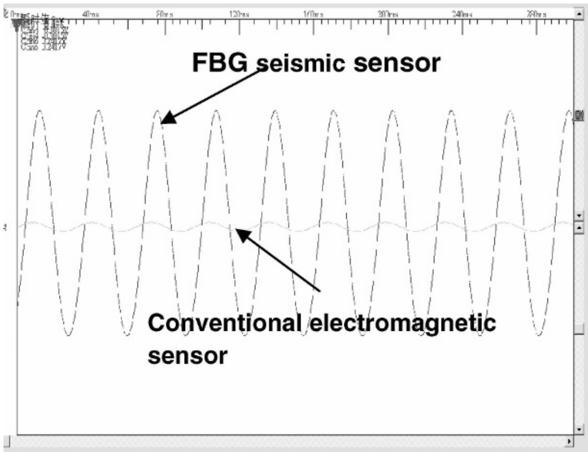


Figure 4. FBG sensor vs. electromagnetic sensor sensitivity comparison.

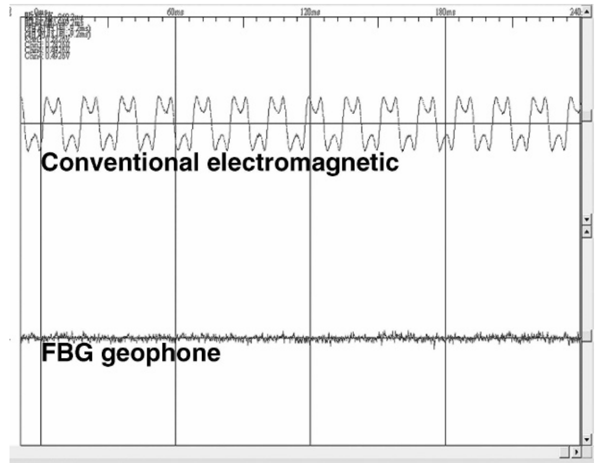


Figure 5. FBG sensor vs. electromagnetic sensor electromagnetic interference.

electromagnetic interference. This demonstrates the advantage of a FBG unattended seismic sensor in a military environment, an attribute that is one of the intrinsic properties of the fiber optic sensor.

The military field test was carried out at one of the U.S. Army shooting ranges. The REMBASS-II S/A sensor was used as a benchmark. *Figure 6* is the sample test result of the FBG seismic sensor response to military vehicle generated signals. The detection distance was calculated by an on-board global positioning system receiver on each vehicle, and it can be interpreted as the maximum distance over which the sensor can respond to the target. In these tests we used the first impulse of the seismic signal as the trigger signal to discriminate the sensor response instantaneously. A series of field tests showed that the FBG sensor equaled or outperformed the REMBASS-II system with both personnel signals and detection of almost every wheeled vehicle signal.

Estimated delays with time of arrival for two pair of sensors

For a pair of signals from two sensors,

$$S_1(t) = a_1 x(t - t_1) + n_1$$
$$S_2(t) = a_2 x(t - t_2) + n_2,$$

where $x(t)$ is the common signal received at both sensors, n_1 and n_2 are additive interferences, and a_1 and a_2 are the attenuation factors at the two sensors.

The relative time of arrival can be computed as

$$dt = \arg\max S_1(t - \tau) \cdot S_2(t).$$

We know the delay of time of arrival can be translated into distance difference because the transmission speed is constant. That is, if we know the transmission speed Sp, we have the distance difference, dl, between two sensors

$$dl = \mathrm{Sp}\ dt.$$

For a given dl, the object of interest will be on a hyperbolic curve where every point has the same distance difference to the two sensors.

If we have one more pair of sensors, we have one more distance difference on another hyperbolic curve. For two hyperbolic curves, the intersection point indicates the exact object location.

The time difference plot in *Figure 7* shows that for both the sensor pair 2 and 13 and the sensor pair 2 and 14 the time difference gets smaller as time progresses, which corresponds to the one approaching from sensor 2 to sensor 4. At sensor 2 the distance difference is largest, and therefore the time difference is also largest. When the person stops at sensor 4, the time difference between two sensors, corresponding to the distance

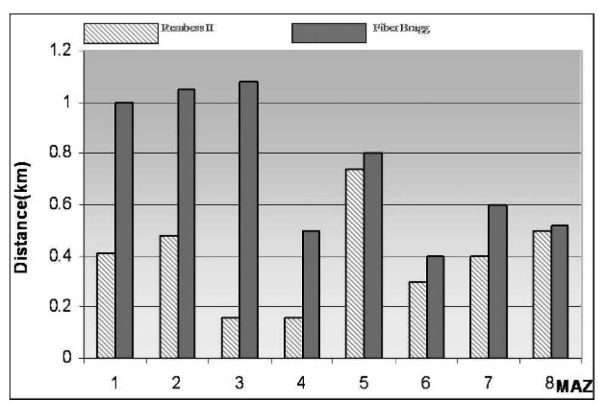


Figure 6. Sample of field test results of FBG response to military vehicles signals (in comparison with REMBASS II sensor).

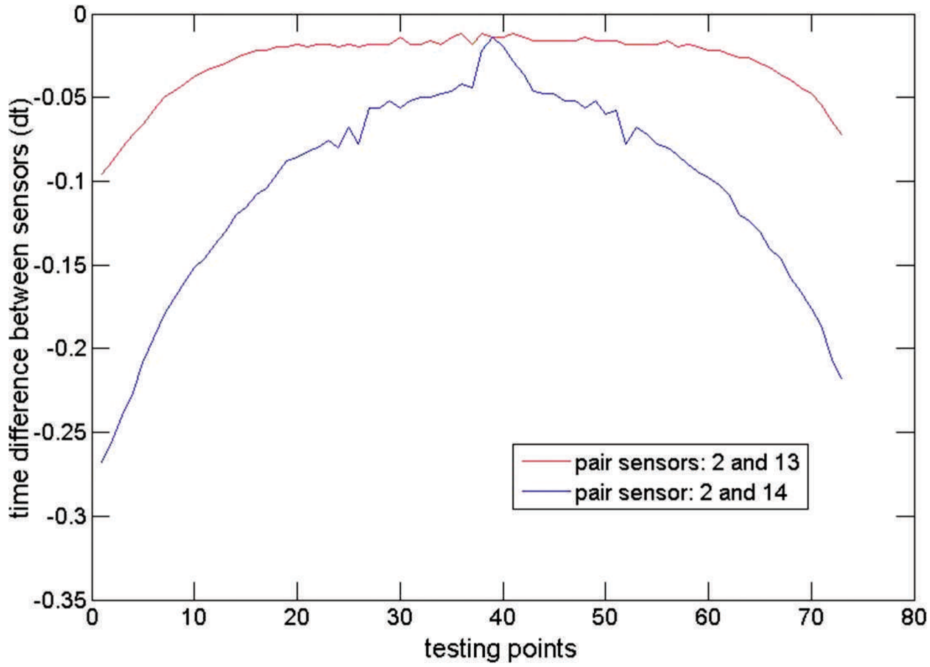


Figure 7. Corresponding location estimates from pair of time delays.

difference, has the smallest value. Then the person turned around and walked backward. One can see that the time difference gap gets larger and larger as sensor 2 is approached. Moreover, we find that two pairs of sensors have two different time difference values because the distances from the person to the sensor pairs are different. The sensor pair 2 and 14 has a larger time difference value.

Summary

We have demonstrated and discussed the feasibility of an unattended FBG seismic sensor in the battlefield monitoring and intrusion detection application. Based on a novel yet practical design, the FBG seismic sensor meets most of the military requirements such as being light weight, of compact size, easy to use, reliable, all-weather, and consuming little power, etc. The performance of the FBG seismic sensor in laboratory tests and field tests exceeds several commercial seismic sensors. We could expect FBG unattended seismic sensors to have the potential capability of detecting time-critical targets (personnel and vehicles) in battlefield applications.

More systematic research by our team is currently focused on combining FBG magnetic sensors with the FBG geophone and improving the algorithm to obtain both signal enhancement and reduced FAR. Also, further engineering measures of sealing and packaging are underway to ensure that the sensor can be deployed in the harsh military environment for extended periods. ❑

JOSEPH DORLEUS is currently a lead telecommunication–systems engineer at PEO STRI, Orlando, Florida. He has worked and held both technical and managerial positions in the private sector as well as in the government. He holds a bachelor's of science degree and a master's of science degree in electrical engineering from Polytechnic University (formerly Polytechnic Institute of New York), Brooklyn, New York, and a doctor of philosophy degree in electrical engineering from Stevens Institute of Technology, Hoboken, New Jersey. His research interests include optical networks, all-optical network management and monitoring, and modeling and simulation of wireless networks. He is a member of the International Test and Evaluation Association (ITEA), the Institute of Electrical, Electronics Engineering (IEEE), the Defense Technical Information Center (DTIC), the Army Acquisition Corps (AAC), and the International Society for Optical Engineering (SPIE). He has authored, coauthored, and presented numerous technical papers that are published in technical journals, conferences, and proceedings such as IEEE, SPIE, ITSEC, and ITEA. Dr. Dorleus is the recipient of the Army Achievement Medal for Exceptional Civilian Service. He was also the Army Materiel Command's nominee for Black Engineer of Year Award in 2001. E-mail: Joseph.dorleus@us.army.mil

YAN ZHANG is a senior engineer at L.C. Pegasus. She obtained her bachelor's of science and master's of science degrees in optical engineering from Beijing Institute of Technology and her doctor of philosophy degree in applied physics from Stevens Institute of Technology. Her current research interest includes research and development of fiber

optic devices and components in optical sensing and optical communication areas. She has several patents and published extensively in the areas of fiber optic sensors and laser spectroscopy.

JING NING received a bachelor's of science degree in electric engineering in 1984 and a doctor of philosophy degree in geophysics in 2005. Now, his professional interests include new sensing technology, which has application to geological disaster warning and mineral exploration.

THOMAS KOSCICA is presently a principal engineer with L.C. Pegasus, where he leads the effort to design and implement fiber optic and microwave sensors in structural health monitoring, methane gas detection, and distributed temperature and pressure monitoring in oil wells, as well as THz sensors for chemical and biological warfare agent detection. He served as a consulting designer to Smith and Nephew for next-generation therapeutic ultrasound products that leverage embedded software to maximize product flexibility and minimize costs during the entire product lifecycle. From 1987–1996, while working at the U.S. Army Research Laboratory in Ft. Monmouth, New Jersey, Dr. Koscica was involved in advanced electronics engineering development for high-performance military systems; during the last 2 years there he served as team leader. For his doctoral thesis, completed at Rutgers University's Electrical Engineering Department in 1997, he explored novel semiconductor device development to utilize the physics of "real space transfer" within a semiconductor to express compound device behavior from a single device. Dr. Koscica holds more than 40 U.S. patents and has numerous publications and presentations.

HONGBIN LI received the bachelor's of science and master's of science degrees in electrical engineering from the University of Electronic Science and Technology of China, Chengdu, in 1991 and 1994, respectively, and a doctor of philosophy degree in electrical engineering from the University of Florida, Gainesville, Florida, in 1999. From July 1996 to May 1999, he was a research assistant in the Department of Electrical and Computer Engineering at the University of Florida. He was a summer visiting faculty member at the Air Force Research Laboratory, Rome, New York, in the summers of 2003 and 2004. Since July 1999, he has been with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, New Jersey, where he is an associate professor. His current research interests include statistical signal processing, wireless communications, and radars. Dr. Li is a member of Tau Beta Pi and Phi Kappa Phi. He received the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for excellence in research in 2001 from Stevens Institute of Technology, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He is a member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society. He is or has been an editor or associate editor for the IEEE Transactions on Wireless Communications, IEEE Signal Processing Letters, and IEEE Transactions on Signal Processing, and he served as a guest editor for EURASIP Journal on Applied Signal Processing, Special Issue on Distributed Signal Processing Techniques for Wireless Sensor Networks. E-mail: Hongbin.Li@stevens.edu

HONG-LIANG CUI is a professor of physics at Stevens Institute of Technology, where he directs the Applied Electronics Laboratory. He received his undergraduate education in applied physics with a concentration in laser optics from the Changchun Institute of Optics and Fine Mechanics in Changchun, China, with a bachelor's of engineering degree. In 1981 he came to the United States for graduate study as one of the first groups of Chinese physics students in the CUSPEA program, obtaining a doctor of philosophy degree in theoretical condensed matter physics in 1987, from Stevens Institute of Technology, where he has been on the faculty ever since. His research efforts have been concentrated in the areas of solid-state electronics and nanoelectronics, optical communications and sensing, electromagnetic wave propagation and interaction with matters such as chemical and bio-agents, and high-performance computing approach to modeling of physical devices and phenomena. His work has been funded by NSF, ARO, ONR, and DARPA. He has published more than 190 research papers in peer-reviewed scientific journals, holds nine U.S. patents, and has guided more than 30 Ph.D. dissertations to completion. He holds membership in the American Physical Society, the Institute of Electrical and Electronics Engineers, the Optical Society of America, and Sigma Xi. E-mail: Hong-Liang.Cui@stevens.edu

References

Robert Pastore, Jr, John Kosinski, Hong-Liang Cui, Yan Zhang, Zhifan Yin, and Bingquan Chen. 2006. Seismic activity monitor based on optical fiber Bragg gratings U.S. patent number: 7,122,783, filed October 17, 2006.

ITEA Journal 2009; 30: 461–467

A Big Fish in a Big Pond: Theodore von Kármán and Southern California's Aviation Empire[1]

Michael H. Gorn, Ph.D.
NASA Dryden Flight Research Center, Edwards Air Force Base, California

After 4 years of prodding and persuading, Robert Andrews Millikan, chairman of the California Institute of Technology (Caltech) finally got his wish. In December 1929 the renowned Hungarian scientist Theodore von Kármán—accompanied by his sister, Josephine, and his mother, Helen—boarded an ocean liner for New York, following which they would travel by train to Pasadena, California, and his new academic home. Unlike the two previous visits to Millikan's school, Kármán came not as a guest, nor as a consultant; he had made the momentous decision to shed his European post as director of the Aachen Aerodynamics Institute, which he had raised to international prominence, to become director of the Guggenheim Aeronautical Laboratory at the California Institute of Technology (GALCIT).

Millikan considered Kármán's arrival the capstone of his drive to mold Caltech into one of the world's great centers of scientific inquiry. He based his strategy not on the classical concept of the university as a solitary temple of learning, but rather on the idea of the university as the center of the community's intellectual life and, as such, the incubator of local industries. He had a straightforward goal: to attract the eminent Kármán to this quiet suburb and to enable him to replicate what he achieved so effectively at Aachen—weave research, teaching, and industrial consulting into a single, seamless fabric. Millikan sought to harness the inherent advantages of Southern California aviation—good year-round weather, vast and inexpensive real estate for hangars and test facilities, and low-cost labor to build it all—to the technical brainpower of Caltech. Ultimately, he enlisted Kármán as the indispensable medium and catalyzing agent for aeronautical manufacturing, transforming it into an economic powerhouse for the Golden State, and in the process earning a place of honor for Millikan's small but ambitious school.

In one sense, Theodore von Kármán seemed an unlikely candidate for this transformative objective. Soft-spoken, slight, and not well versed in English, he hardly embodied the dynamic scientist-salesman that Millikan seemed to want. But Millikan, a realist, knew the facts. As Harry Guggenheim—son of philanthropist Daniel Guggenheim, the founder of the Guggenheim Fund for the Promotion of Aeronautics and financier of GALCIT—once observed, "Of all the good scholars, only a few stand out, and those few . . . have wisdom as well as intellectual capacity. If it interested him, [Kármán] could make a great ambassador, or a great . . . anything. I think that is true of the few great men with whom I've come in contact. They are very few."[2] Accurate though Guggenheim's observation may have been, Kármán certainly had some peculiarities, not least of which was a nomadic life, lived almost in chaos. But he also possessed great personal strengths, including humor, charm, the capacity to win people to his causes, and an effortless ability to forge friendships.

Born on May 11, 1881, Theodore von Kármán grew up in a solidly middle-class Hungarian family in a cluster of comfortable apartments on the Buda side of the Danube River. Together with Pest on the other bank, the city of Budapest rivaled Vienna as the capitol of the Austro-Hungarian Empire. On his mother's side, his ancestors had been scholars, scientists, and prominent rabbis, including Yehuda Loew Ben Bezalel, an associate of the Danish astronomer Tycho Brahe. In contrast, his grandfather—born into the artisan class like much of the rest of the family—worked as a tailor for the Hungarian aristocracy. Theodore's own father, Maurice, embodied the exception. Endowed with drive and full of ambition, he had become a philosophy and education professor at the University of Budapest, ultimately earning a title of nobility for his reforms of Hungarian education, which he modeled on the German *gymnasia*.

Young Theodore's gifts, especially in mathematics, became evident early. By the age of 6 he could multiply 5- and 6-digit figures in his head. Yet, rather than spur

Figure 1. Theodore von Kármán (center, with wide stance and no hat) amid the leaders of the Langley Memorial Aeronautical Laboratory, Hampton, Virginia, in 1926, the same year in which he made his initial visit to Caltech. The Langley operation reported to a small federal organization known as the National Advisory Committee for Aeronautics, the predecessor of NASA. [NASA/courtesy of nasaimages.org]

his son to pursue the life of a prodigy, Maurice von Kármán insisted instead that he obtain a rounded education, infused with history, literature, and the arts, and defer the intense study of the sciences until college. Unfortunately, Maurice suffered a mental breakdown just as his son prepared for his university career, which narrowed Theodore's options. Until then, the younger Kármán showed such promise that, under normal circumstances, he might have expected to study in London, Paris, or Berlin. Instead, he attended the Royal Joseph Technical University in Budapest, where he apprenticed himself to engineering professor Donat Banki. Banki influenced Kármán profoundly, training him to think not just theoretically, but also practically, like a scientist *and* like an engineer. After Theodore's graduation, his father's health stabilized, and he enrolled at Göttingen University in Germany to pursue a doctorate in fluid mechanics under the master of the boundary layer, Professor Ludwig Prandtl. During this time, Kármán encountered the study of turbulence, which became a lifelong source of intellectual curiosity, and in its pursuit during his college days he identified a curious string of parallel vortices, later called the Kármán vortex street in his honor.

This discovery won him celebrity and ultimately resulted in his appointment to the directorship of the Aachen Aerodynamics Institute (at the Aachen Technical University) at the tender age of 31. Unlike Göttingen, whose academic rules and rigidity rankled Kármán's passion for disorder, he basked in the cosmopolitan air of Aachen, a German city adjacent to Holland, Belgium, and France. Here Kármán fashioned a little-known department into a famous research organization. He accomplished this feat by earning the personal trust, the technical admiration, and the consulting fees of the leaders of such manufacturing firms as Junkers, Zeppelin, and Fokker. He reconstructed Aachen's open-ended wind tunnel, converting it into a more efficient closed circuit model. He hired able faculty and threw himself heart and soul into teaching, the approach to which he based on one of Donat Banki's principles: use simple examples from life to illustrate complex physical phenomena. A popular professor and lecturer, Kármán realized that the classroom represented only one kind of pedagogy; parties at his home and informal sessions in the cafés of Aachen offered alternative and more agreeable places to nurture his students. Finally, he burnished his reputation at the institute by scoring three new technical achievements: a moveable trailing edge for airfoils, a restatement of Prandtl's boundary layer theory, and a reformulation of the boundary layer equations (enabling comparisons between theory and experiment).

Robert Millikan knew as early as 1926 that in the brilliant and still relatively young Kármán he had found his first and only real candidate to guide the new aeronautics institute at Caltech *(Figure 1)*. Consequently, just after he persuaded the Guggenheims to fund GALCIT (with an initial investment of $305,000), Millikan also persuaded Kármán to make a consulting visit to Pasadena, in the fall of 1926. Almost from the hour of his arrival on campus, he joined Clark Millikan (Robert's son, who held a doctorate in fluid mechanics) and Arthur Klein (a Caltech experimental physicist) in designing a new wind tunnel.[3] Wasting no time, Kármán scrapped the Millikan-Klein proposal, an open-ended model still in vogue in the United Kingdom and elsewhere. He advised them to supplant it with a closed system (like that at Aachen), which offered greater efficiency and more space for adjoining offices, classrooms, laboratories, and the like. Additionally, the Hungarian suggested a daring modification by situating the tunnel's motor *inside* the airstream, close to the propeller, in order to shorten the driveshaft and reduce the number of bearings. Despite objections raised by his son and Klein, the elder Millikan ruled in favor of Kármán's concept.

Two years after finishing his initial consulting at Caltech, the Aachen director returned to Pasadena in September 1928 for the official opening of the Guggenheim Laboratory and the first semester of classes. The building itself—an imposing 5 stories tall by 160 feet by 55 feet—housed the tunnel, outfitted with a 750-horsepower motor and a 15-foot propeller, capable of producing airflows of 200 miles per hour. At one end of the machine, a chamber 4 stories high by 50 feet by 20 feet had been reserved for mounting model aircraft, which could measure up to 30 feet in length. In the basement, a water channel 140 feet long, 10 feet wide, and 10 feet deep awaited researchers. The first and second floors held a machine shop, a motor-testing room, six small laboratories for engine research, and a big woodworking area for the fabrication of tunnel models. The third floor contained five offices, a seminar room, a drafting facility, and a library.

Perhaps even more than the physical milestones in the laboratory's construction, the awakening of GALCIT operations began to enmesh Kármán in Southern California life. He found himself corresponding frequently with Clark Millikan during the building's construction (1927 to 1928), the substance of which read suspiciously like a boss instructing a subordinate. In addition, the elder Millikan had engineered a rotating teaching exchange that began in fall 1927 with Caltech mathematician Paul Epstein decamped to Aachen, reciprocated by Kármán's visit to Caltech in 1928, and so on in future years. During this stay, Kármán not only worked out the last details of the tunnel but also participated in the curriculum development of 2-year postgraduate studies in aeronautical engineering. Finally, by this time, Kármán actually bore an official title: GALCIT research associate.

Before he left for Aachen in December 1928, Robert Millikan decided to ask Kármán directly: would he accept the position of director of the Guggenheim Laboratory? Still hesitant, and hesitant also to subject his mother and sister to the trauma of moving not just to another country but to another continent, Kármán deferred an answer. Then the heavy guns came out. Harry Guggenheim wrote an emotional appeal to him, reminding him that at Caltech he would be free of the mandatory pressures of fund raising and consulting, both integral parts of his life at Aachen. Millikan upped the ante, promising him a princely salary of $10,000 per year (three times his Aachen pay and third among all Caltech employees) and an annual operating budget of at least $50,000, under Kármán's personal control. Millikan also got lucky. The first signs of worldwide economic depression started to be felt in Germany, which rendered a solid offer of this kind too good to pass by. Kármán wired Millikan in October 1929 advising that he would assume the GALCIT post in April 1930, more than 4 years after his initial visit to Caltech.

For the third time in 4 years, Theodore von Kármán rode the rails from New York to Los Angeles. This time, though, he faced the prospect of setting up a household. Fortunately, he soon found a place that appealed to him, as well as to his sister and his mother. Just 2 miles from campus on 1501 South Marengo Avenue, the home straddled the borders of Pasadena, San Marino, and Los Angeles. Kármán needed a residence like the one he inhabited when he worked at Aachen, in which he could entertain faculty, businessmen, students, and other varied visitors. He found it on this spot. Situated on 2 acres, it looked like a Spanish villa, dominated by a big square house with a heavy tiled roof and low stucco walls. At the front, a rolling lawn and high gate offered privacy; at the rear, serpentine walkways crossed a handsomely laid-out garden.

Buying a home, helping his mother and sister adapt, and meeting many new faces all at once contributed to Kármán's growing involvement in Pasadena. But these factors paled in comparison to his absorption in the life of GALCIT. Millikan reminded him just what Caltech expected from the Guggenheim Laboratory: to act as a magnet to attract the American aviation industry to Southern California, and to lure the regional aviation industry—as well as national recognition—to Caltech. These expectations came into high relief with the publication in May 1930 of a GALCIT brochure entitled, "Announcement of the

Figure 2. A profile of Kármán taken at a public event in 1939, the year in which the Caltech rocketry program under his supervision became organized and funded as the Air Corps Jet Propulsion Research Project. At this stage, Caltech rocketry consisted in part of designing, fabricating, and flight testing rocket canisters to boost aircraft performance; in time, it (and other aspects of the rocket program) would metamorphose into the Jet Propulsion Laboratory. [Image courtesy of NASA]

Figure 3. Before trial flights of the GALCIT rocket canisters at March Field, California, in 1941, Kármán makes calculations on a wing as Clark Millikan (far left), student Frank J. Malina (second from right), and Captain Homer Boushey (far right) look on. [NASA/courtesy of nasaimages.org]

Graduate School of Aeronautics." Issued the month after Kármán's permanent relocation, it offered the first clear sign of the direction of his leadership. It announced a faculty not too different from that present when GALCIT opened in 1928 (with the major exception of Kármán himself), consisting of Clark Millikan, Harry Bateman, and Arthur Klein. But it also noted the additions of Arthur Raymond (assistant chief engineer at Douglas Aircraft); doctoral candidate Ernest Sechler; and (on a part-time basis) faculty members Paul Epstein, fellow mathematician Eric Temple Bell, and geophysicist Beno Gutenberg. The brochure also gave Kármán the opportunity to define the school's research emphases, with two general fields dominating its curriculum: aeroelastic structures/elasticity and aerodynamics/fluid mechanics.

Not surprisingly, these areas closely paralleled Kármán's personal specialties. They also reflected his preference for applying theoretical research to practical problems *(Figures 2 and 3)*. The structures program represented a conscious effort to provide aircraft companies with reliable data on metal skins and supports based upon "sound theory . . . substantiated by experimental evidence."[2] Raymond and Klein taught the aeroelasticity courses while Kármán, whose dissertation and earlier works concentrated on columnar strength, developed new theories on the buckling properties of sheet metals. The aerodynamics program, on the other hand, concentrated on skin friction and boundary layer phenomena, exploring fundamental concepts of drag both on test aircraft and in the GALCIT wind tunnel. Kármán, Klein, and Clark Millikan took the leading role in this department, pursuing wing theory and new frontiers in fluid mechanics.

Very quickly, the results flowing from the Guggenheim Laboratory influenced the design of the latest aircraft produced by such Pacific Coast firms as Douglas, Lockheed, Consolidated-Vultee, Northrop, North American, Hughes, and Boeing. In fact, a hint of the success to come occurred just as Kármán took control of the lab. Originally, the staff had assumed that tunnel usage would be divided about evenly among GALCIT researchers and the regional aircraft industries. But because the new Caltech machine outperformed almost every other wind tunnel then operating, the proportion shifted heavily toward the commercial firms. Indeed, it became so popular that despite 17-hour-per-day operations, the laboratory's faculty and students could only find a few weeks out of a year for their own work.

The heavy booking may be explained in part by the labor-intensive nature of the experiments, most of which took from 1 to 4 weeks to complete. The process began with the delivery to the lab of large hardwood prototypes—some with wingspans as wide as 8 to 9 feet—that the aircraft companies wanted to verify. A team of GALCIT students suspended these test

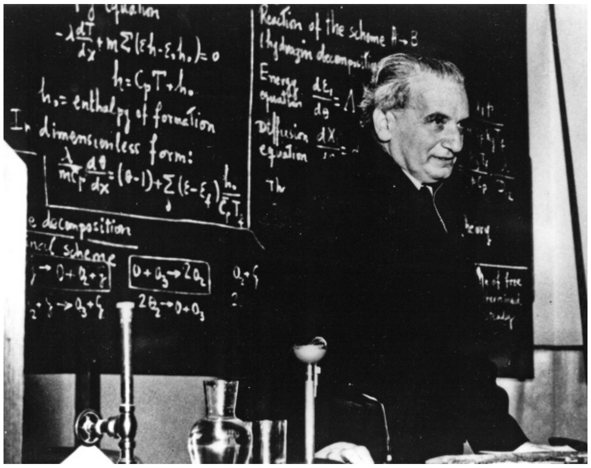


Figure 4. Kármán at the blackboard. He considered teaching his most important professional responsibility and, among his many attributes, his greatest single talent. [NASA/courtesy of nasaimages.org]

articles from seven balance points and subjected them to a range of airspeeds. Weight readings taken at the seven terminals, relayed to two operators in a control room, indicated the load carried by the models at various simulated altitudes and velocities. The staff then compared the data to aerodynamic calculations. At the end stage, Clark Millikan analyzed the findings and prepared reports for the firms, often recommending modifications of the designs.

Perhaps even more important than the actual engineering services Kármán rendered to Donald Douglas, Jack Northrop, and the rest, for the first time in American history a flight dynamics scholar of worldwide renown invited the leaders of the aircraft industry onto a university campus to derive practical benefits from the latest research. In addition, engineers employed by the regional aircraft companies and military officers involved in research and development soon appeared at the laboratory not just to observe the operation of the storied new wind tunnel, but also to take undergraduate and graduate courses, to earn degrees, and to attend Kármán's periodic special lectures *(Figures 4 and 5)*. Moreover, once they began to appreciate the increases in aircraft efficiency to be derived from GALCIT research, the bosses of Douglas, Northrop, Lockheed, and the others began to hire Kármán's graduates as full-time members of their design teams. As a consequence, in a relatively short period of time, a more science-based approach to materials, controls, and structures overcame the traditional reliance on empiricism in American aircraft design. Although other centers of university aeronautics existed in the United States prior to the creation of GALCIT, the Caltech model that integrated teaching,

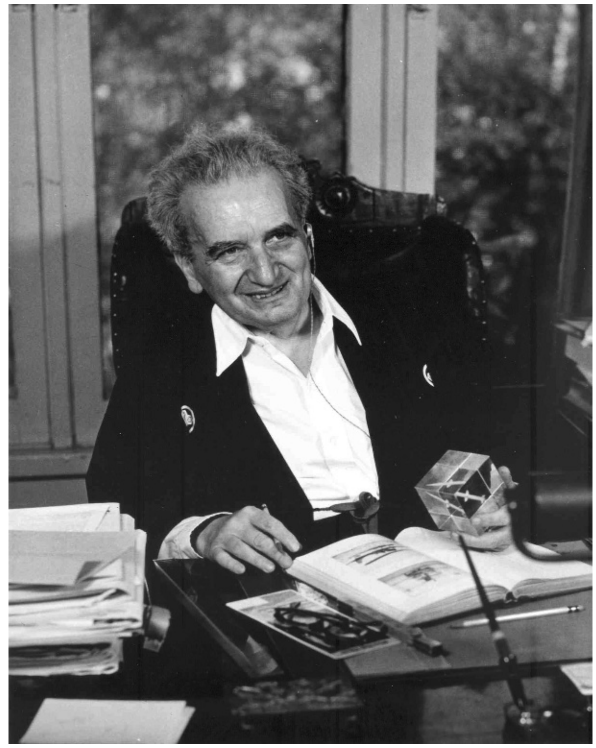


Figure 5. Kármán relaxing at his desk in his office at home, a place where many of his students found themselves late at night, after the parties on South Marengo Avenue had subsided, attempting to solve equations under his watchful eye. [Image courtesy of NASA]

basic research, and industrial consulting—a European model based upon Kármán's own experiences—gradually took root throughout the country *(Figure 6)*.

No project from this period did more to advance GALCIT research—or indeed, aeronautics as a whole—than the early series of Douglas Commercial (DC) passenger aircraft. The initial contact between the laboratory and the manufacturer occurred between Arthur Klein and Donald Douglas himself. Klein worked part time for the Santa Monica, California, airplane maker and persuaded him to refer some of his technical problems to the Pasadena campus. Although the Guggenheim family at first subsidized the work for Douglas, in time the design studies undertaken for him led to a booming consultancy for GALCIT and its wind tunnel.

Until February 1934, building airliners had not been a profitable business. But in that month, the balance shifted: the U.S. Army Air Corps abruptly canceled its airmail contracts with private carriers and began to fly the routes with military airplanes and crews. Deprived of lucrative orders for transport aircraft to haul the mail, the aircraft industry scrambled to adapt their designs for commercial use. For his part, Donald

Figure 6. Pausing for a cigar as he led a group of allied scientists into newly liberated parts of Europe at the close of World War II, Karman is flanked by Dr. Hugh L. Dryden (far left) of the National Bureau of Standards, and Drs. Ben Lockspeiser (second from left) and A. P. Rowe (far right), both English scientists. Kármán undertook the task of interviewing captured German scientists and seizing their papers and equipment at the behest of Commanding General of the Army Air Forces Henry H. ''Hap'' Arnold. [NASA/courtesy of nasaimages.org]

Douglas asked Arthur Raymond, one of his most able engineers, to help the company find a niche in the airline market. By now a part-time teacher on Kármán's faculty, Raymond brought the full resources of GALCIT to bear as his firm set its sights on designing the prototype DC-1 airliner, which went into production as the bigger, 14-seat DC-2, built expressly to carry human beings instead of sacks of letters and parcels.

During this time of design ferment, Douglas promoted Raymond to chief engineer and put him at the head of a group charged with fabricating an expanded DC-2. The resulting DC-3 airliner soon became the empress of American and world air travel and the first airplane to make money for the airlines on passenger seats, as opposed to airmail contracts. It attained its legendary reputation for speed, strength, durability, and comfort in large part due to the collaboration between the Guggenheim lab and the Douglas engineers. In its early flight tests, the DC-3 experienced severe wind buffeting, especially troubling because of its low wing mounting on the fuselage. Through calculations and experimentation, Kármán and his staff determined that a thin fairing, or fillet, fastened along the juncture of the top surface of the wing and the fuselage quieted the turbulence. Wind tunnel data demonstrated that this adaptation eliminated a dangerous leading edge, from which eddies of air (the notorious Kármán vortices) raced back, severely shaking the aircraft's rear structure. The discovery and remedy made the difference between a highly stable aircraft fit for the flying public versus one threatened by the dangerous turbulence that afflicted many other transports. In France, engineers soon referred to these little metal fairings as "Kármán fillets," a feature that found its way into many aeronautical designs.

One other Caltech research innovation greatly improved the seminal DC-3. Throughout the 1920s and into the early 1930s, airplane fabrication continued to rely for the most part on wood and fabric, as well as external supports. Manufacturers shied away from metal structures and skins, fearing that the weight and buckling qualities rendered them unsafe for the pressures associated with flight. But Kármán, working with Ernest Sechler, proved that by running stiffeners along its surfaces, sheet metal would flex, not deform. Constructed entirely out of a metal shell and strengthened by internal, rather than external, bracing, the DC-3 began service in 1935, and its immense and immediate popularity did much to broadcast its updated fabrication methods and materials to manufacturers everywhere.

The landmark improvements contributed by Theodore von Kármán[4] and his staff to the DC-1, the DC-2, and finally the DC-3—the most influential airliner in the history of aviation—cemented the reputation of GALCIT among the elite aeronautical institutions, as well as among the aircraft firms of the region *(Figure 7)*. In addition, by catapulting the laboratory's work to the far corners of the aviation world, these early projects represented at least a partial fulfillment of

Figure 7. Late in life, Kármán stands arm-in-arm with two old friends: his successor as director of the Jet Propulsion Laboratory, Dr. William H. Pickering (left), and Dr. Frank J. Malina, one of his first rocketry students. [NASA/courtesy of nasaimages.org]

Robert Millikan's[5] dream of fashioning Caltech into a competitor with the nation's great centers of scientific teaching and research. ❑

Michael H. Gorn, PhD, has worked as a public historian for more than 30 years, in the National Aeronautics and Space Administration, the Department of the Air Force, and the Environmental Protection Agency. He is the author of many books about the history of aeronautics and spaceflight, most recently NASA: The Complete Illustrated History *(Merrell, 2008) and* Superstructures in Space: From Satellites to Space Stations, a Guide to What's Out There *(Merrell, 2008). He is also the author of* Expanding the Envelope: Flight Research at NACA and NASA *(University Press of Kentucky, 2001), winner of the 2004 Gardner-Lasser Aerospace History Literature Award, presented by the American Institute of Aeronautics and Astronautics. E-mail: (care of) rjanssen@allenpress.com*

Endnotes

[1]This article is excerpted from Michael H. Gorn, *The Universal Man: Theodore von Kármán's Life in Aeronautics* (Washington and London: Smithsonian Institution Press, 1992), especially chapter 4, "A Magnet for Aeronautics." It also adapts material from chapters 1 to 3.

[2]Gorn, *Universal Man*, 45.

[3]Kármán, Klein, and Millikan did not have to start their deliberations from scratch. Caltech (then known as Throop College of Technology) started its aeronautics program as early as 1917, four years before Robert Millikan arrived on campus. Two faculty members—prominent mathematician Dr. Harry Bateman from Cambridge University and aviator Albert Merrill (who oversaw the construction of a 40 mile-per-hour open system wind tunnel at Throop)–started the program with limited success. Later, Clark Millikan, who took his doctorate from Bateman, and Arthur Klein joined Bateman and Merrill to form the backbone of the pre-GALCIT program at Caltech.

[4]Theodore von Kármán (1881–1963) pursued many important projects during the remainder of his long and active life. He continued to serve as GALCIT director until the early 1950s, during which time he sponsored the nation's first rocketry program on a college campus; continued to consult with many regional industry leaders (particularly Jack Northrop); founded and directed the Jet Propulsion Laboratory; advised Hap Arnold, commanding general of the Army Air Forces throughout World War II, eventually contributing the seminal postwar report, *Toward New Horizons*; and conceived of (and served as the first chairman of) the Air Force Scientific Advisory Board. During the early 1950s, he formed the NATO Advisory Group for Aeronautical (later Aerospace) Research and Development. In his last years he shuttled between Paris and Pasadena. Three months after being awarded the first National Medal of Science by President John F. Kennedy at the White House, Kármán died in May 1963, ironically in Aachen, Germany, where had gone to soak in the restorative hot springs. He is buried in Hollywood, California.

[5]Robert A. Millikan (1868–1953) served as chairman of the Executive Council of Caltech from 1921 to 1945, in addition to being the director of the campus's Norman Bridge Physics Laboratory, where he concentrated on cosmic ray research. Millikan received the Nobel Prize in physics in 1923 for his experimental work at the University of Chicago on the elementary charge of electricity and the photoelectric effect. In the end, he succeeded in his drive to bring distinction to the California Institute of Technology and is widely regarded as the father of the modern Caltech. See Judith Goodstein, *Millikan's School: A History of the California Institute of Technology* (Norton, 1991).


Will it work for the Global User?
Stiletto image courtesy of M Ship Co.
Shuttle image courtesy of NASA
Ask SAIC. We take it personally.
Testing in a global environment is a job our 44,000 employees take to heart everyday. For more information call.
SAIC T&E Support
Operational Test, Military Utility Assessment, Distributed Test, Joint Test & Evaluation Support, T&E Infrastructure
Contact: Mark E. Smith
mark.e.smith-2@saic.com
505-830-6757
SAIC®
From Science to Solutions

ITEA Journal 2009; 30: 469–472

Simulation and Analysis Facility (SIMAF)

Timothy Menke
Air Force Simulation and Analysis Facility, Wright-Patterson Air Force Base, Ohio
M. Walter March
Clear Creek Applied Technologies, Inc., Wright-Patterson Air Force Base, Ohio

The Simulation and Analysis Facility (SIMAF) is a state-of-the-art U.S. Air Force facility specializing in high-fidelity, virtual (manned), distributed simulation to support acquisition and test located at Wright-Patterson Air Force Base, Ohio. The SIMAF is charged with supporting capability planning, development, and integration in support of Air Force and Department of Defense acquisition program objectives. Key areas of emphasis include capability development and integration, network-centric system-of-system development, and electronic warfare.

Key words: Capability development; network-centric system of systems; communication; horizontal integration; reusability; realistic battlespace environment; distributed test.

Warfare is information-centric, and the future will continue to reward the combatants who are able to communicate across the force in real time, fuse those data, and swiftly apply cognitive reasoning to those data to yield an information advantage over the enemy. Data flows through electronic channels between our forces. The Simulation and Analysis Facility (SIMAF), located at Wright-Patterson Air Force Base (AFB), Ohio, recognizes the need to assess those electronic data between our platforms; command, control, communications, and computers (or C4); and weapons by bringing the best resources of the cross-domain community together through a distributed test bed *(Figure 1)*. The SIMAF was developed to support this objective.

The SIMAF has the ability to conduct large-scale live, virtual, and constructive (LVC) analytically based assessments within the facility or connect to other U.S. Air Force and Department of Defense (DoD) or coalition partners. From the concept of operations and tactics, techniques, and procedures development to the assessment of current or future systems, the SIMAF's expertise spans numerous domains *(Figure 2)*. Our analytical process is based upon system engineering principles with the objective of fostering software reuse and thereby increasing the return on investment for any dollar spent in simulation. Each assessment environment is uniquely constructed from the analytical objectives and implemented using an object-oriented real-time framework. Software and hardware integrated applications are built from that framework. One of our core capabilities includes the ability to build unique integrated hardware and software systems, conduct analyses of these systems in a real-time environment, and deliver usable data analysis products to the acquisition community and to do it quickly. One of our primary goals is to ensure these hardware and software systems are designed to allow for reusability: we enable massive reuse of software objects for each project. As our software object library grows and relationships with distributed partners' mature, project completion times are greatly decreased. Our ability to rapidly construct real-time virtual environments in record time is based on a system engineering process, a reusable object-oriented framework, and automated tools that allow the team to collaborate across the life cycle of an assessment. This complex mission is accomplished by our staff of 20 government employees and over 70 contractors.

The SIMAF was created by the Aeronautical Systems Center in 1995 to fill a gap in the application of manned simulation to support system requirements development and refinement. Since 1995, the SIMAF mission has matured from a single-system focus to a system-of-system focus. Over time, system interoperability matured into network-centric or network-enabled warfare concepts. Today, the SIMAF's mission centers upon decision support necessary to assess the horizontal integration of systems via tactical data links to generate desired battlespace effects. Our mission necessarily encompasses foundational capabilities vital to the successful execution of any mission such as tactical data links and electronic protection and

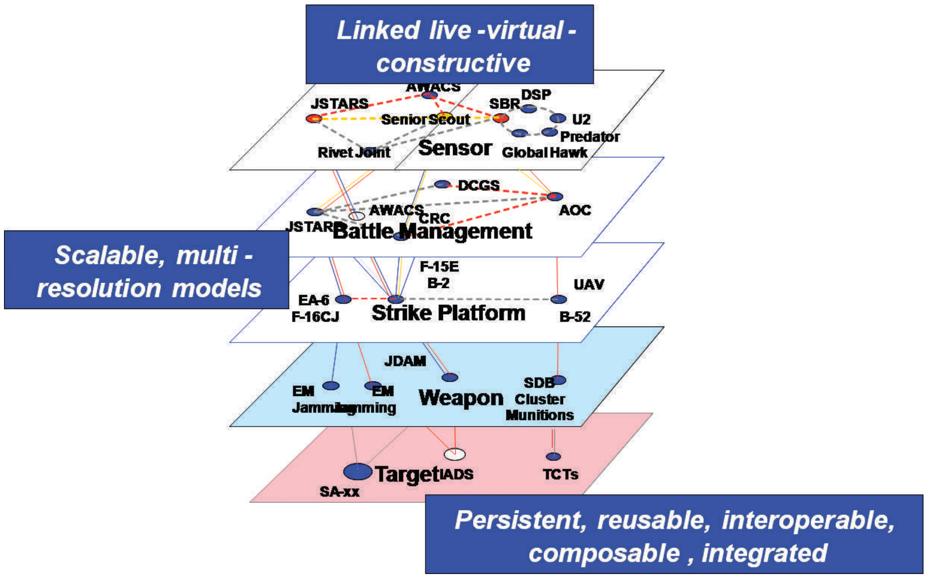


Figure 1. Representation of data flow in a virtual environment.

attack. Since 2000, the SIMAF has established key Air Force and Department of Defense partners via distributed simulation to bring the best of breed from across DoD in support of our customers.

In 2004, the Air Force–Integrated Collaborative Environment (AF-ICE) was created in response to an evaluation gap in the ability to access network-centric weapon systems and provide net compliance testing prior to fielding *(Figure 3)*. AF-ICE is the combination of existing Air Force live (test range assets), virtual (modeling and simulation facilities), and constructive (models) assets in the inventory to evaluate these

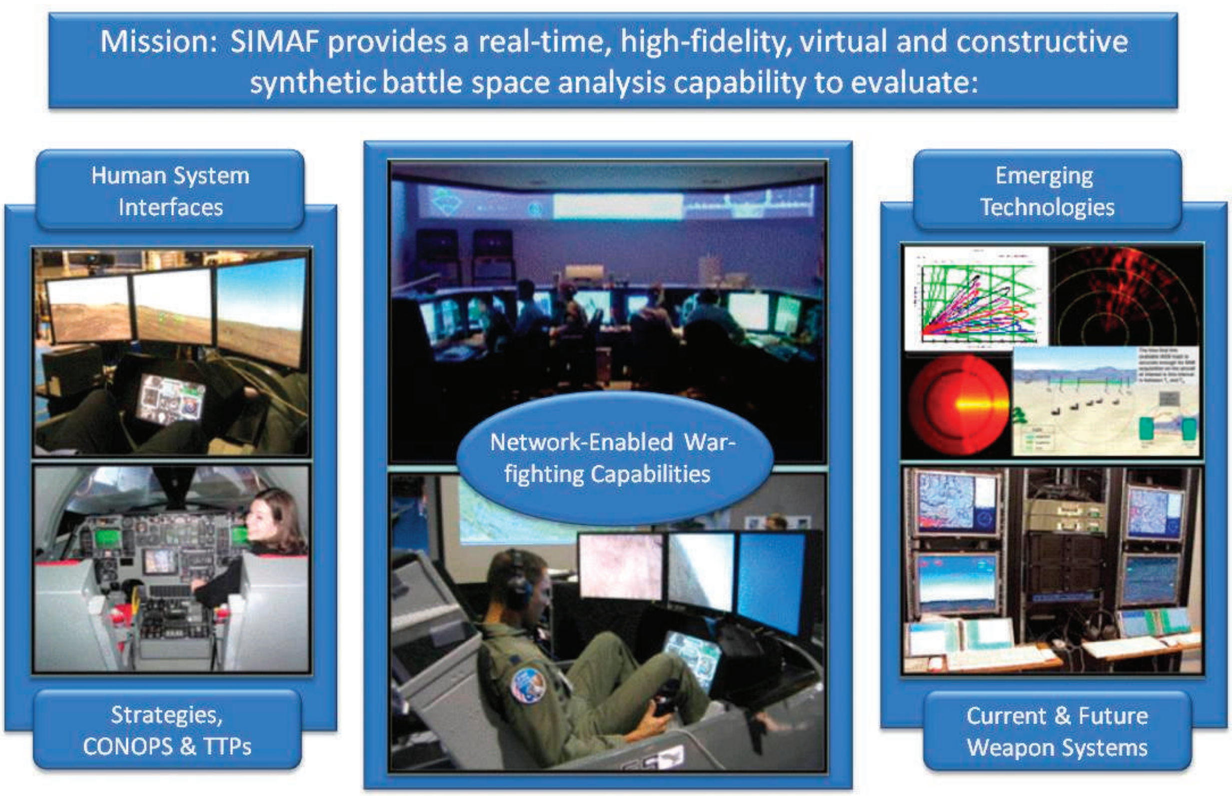


Figure 2. Integrated live, virtual, and constructive environment.

Figure 3. Air Force–Integrated Collaborative Environment.

complex systems. These facilities can be linked together via the Joint Mission Environment Test Capability (JMETC). AF-ICE has facilitated partnerships with other organizations within the LVC community. Examples of these partnerships include the Global Command and Control Innovation Center at Langley Air Force Base, Virginia; the 46th Test Squadron/Data Link Test Facility at Eglin Air Force Base, Florida; and other key service and industry partners such as the U.S. Navy at Patuxent River Naval Air Station, Maryland.

In addition to the relationships the SIMAF has with other Air Force and joint DoD test ranges, we have developed a strong relationship with the National Air and Space Intelligence Center and other members of the intelligence community. The combination of these relationships has allowed us to develop and maintain an ability to generate scalable, valid threat environments to support real-time assessments in a credible, integrated air defense system computer–based representation.

Current projects include Airborne Electronic Attack, Airborne Networking, and Unmanned Aerial System Sense and Avoid. The SIMAF continues to work with our partners to foster improved air and ground modeling and simulation infrastructure in support of joint network design, development, and assessment.

Given the history of building and prototyping capability and partnerships, the SIMAF also is among the thought leaders in facilitating, integrating, and executing events with the main focus to capture credible data useful for decision support. For example, the SIMAF is leading a new activity, Air Ground Integrated Layer Exploration (AGILE). The objective of AGILE is to provide a mechanism to address complex system-of-system assessment objectives by capturing the overarching customer objectives; rendering those objectives into requirements; and satisfying those requirements via analysis through the planning, technical integration, and execution in an LVC environment. The AGILE combined distributed test team will conduct a technical and an operational assessment of the information exchange requirements and broader system interoperability to support conventional mission threads, including Joint Close Air Support, Time Sensitive Target, and Joint Fires. The primary goal is to identify (interoperability) gaps, shortfalls, and overlaps with current systems and networks supporting Joint Fires within and between the Air Force, Navy, U.S. Marines, and U.S. Army air-to-ground communication layers. Additionally, AGILE will assist in defining requirements for the Command and Control Network Partnership and build upon the Joint Forces Command Joint Systems Integration and Interoperability Lab concept in a joint distributed LVC network to support developmental test and operational test agencies.

The LVC network for AGILE will be based on the existing JMETC infrastructure (a long-term SIMAF partner) tying various sites together for event execution. Potential AGILE Fire sites include the White Sands Missile Range, Central Technical Support Facility, SIMAF, Global Cyberspace Integration Center, Electronic Systems Center, Space and Naval Warfare Systems Command, Redstone Test Center, and 46th Test Squadron. The maturing JMETC environment is built on the foundation of an operationally representative construct that allows programs at all stages in the acquisition process the opportunity to quantitatively test/assess their performance.

Future of distributed test

With AGILE Fire (and other activities), the SIMAF is partnering with several organizations within the DoD to address critical factors in distributed testing. One of those factors is the nature of the historical acquisition process. Our acquisition and test culture is system-centric. Additionally, our acquisition and test processes are serial and independent processes. Finally, the service acquisition and test cultures do not readily lend themselves to the concurrent acquisition and testing of air and ground systems within a broader, more expansive concept of operations. The result is anything but interoperability. Perhaps the most difficult aspect is to change the culture and policy. System-of-system interoperability is reliant on distributed environments. The SIMAF has a focus on distributed testing, but overall success requires the community to increase the confidence level of program managers, primes, testers, and many other disciplines, to include the operational force and analysts, that the event data are credible. Furthermore, DOT&E needs to capture the continuum of activity leading up to a "system" operational test. An assessment environment that can provide early insight for program managers and primes through continual assessment opportunities, for all programs big and small, will increase the program managers' comfort level and will be seen as a benefit rather than a burden. This approach needs to map to a mission context and comprise combined service/joint assets so technical and cognitive effects can be understood. Moreover, regardless of where a system or technology is in the acquisition cycle—from milestone A through full operational capability—this continual assessment/insight–early influence business model provides opportunities to understand system attributes in a combined service environment without the gravity of a milestone decision. Using a flexible, disciplined approach within the acquisition community, from individual acquisition programs to emerging technologies that "touch" many platforms, will provide the means to effectively use distributed environments with confidence, reducing overall acquisition and test cost while providing increased confidence that systems will interoperate in the field.

Cleared for Public Release on September 22, 2009. Case number 88 ABW-2009-4125. ❑

TIMOTHY MENKE graduated from the University of Kansas with a bachelor of science degree in aeronautical engineering in 1985, attended Officer Training School, and was assigned to the Aeronautical System Centers at Wright-Patterson AFB. In 1990, he joined the federal civilian service working in the Air Force Research Laboratories and completed his master of science in aeronautical engineering at the Air Force Institute of Technology in 1992. Mr. Menke completed his juris doctorate from Capital Law School in 1998. His current assignment, with the Aeronautical System Center, is technical director in the Modeling and Simulation Division within the Capabilities Integration Directorate located at Wright-Patterson AFB. The Modeling and Simulation Division supports the development of a number of Air Force programs and capabilities utilizing high-fidelity, often emulative, virtual and constructive simulation. Linked with live test resources, the environment provides a mission effectiveness–based assessment capability to provide insight to acquisition and test decisions. E-mail: timothy.menke@wpafb.af.mil

M. WALTER MARCH's duties include providing input to the technical direction for the Aeronautical System Center Simulation and Analysis Facility in the design, development, and integration of multiservice live, virtual, and constructive (LVC) environments to provide decision support to the U.S. Air Force/Department of Defense acquisition and test and evaluation (T&E) community. He was the Air Force lead for the Office of the Secretary of Defense–sponsored Multi-Service Distributed Event; the overall integration lead for the Integral Fire 2007 LVC Event, which included 19 multiservice/Joint Forces Command sites from across the United States focused on Joint Close Air Support and Integrated Fires; and the integration lead for Persistent Fire 2009. He has over 22 years' experience in the T&E/Modeling Simulation & Analysis discipline. Mr. March has supported multiple T&E efforts to include the B-2 Bomber. As a support contractor, he was the Air Force Operational Test Center B-2 test manager; primary responsibilities during the Initial/Final Operational Test and Evaluation and Force Development Evaluation efforts included test plan development, pretest planning, test execution, data reduction, data analysis, and report writing. E-mail: martin.march@wpafb.af.mil

ITEA Journal 2009; 30: 473–480

A New Alpha-Omega Map for Acquisition Test and Evaluation

George Axiotis
Department of Defense, Washington, D.C.

Department of Defense (DoD) Acquisition Test and Evaluation (T&E) has been the gatekeeper to Major Defense Acquisition Program production since its formalization over 25 years ago. Yet, the landscape of the types, methods, and sources for war fighting systems has significantly evolved. The Department has studied and recommended action for Acquisition reform for decades, only tweaking in the margins for T&E. The time is right for DoD to consider a new approach to T&E, steering away from the "buy" decision to the more relevant "acceptance" and "operational" domains. This article outlines the issues and proposes a new "Alpha-Omega" map for T&E for the way we actually procure DoD systems.

Key words: Acceptance tests; acquisition reform; contractual necessity; field operations; low-rate initial production; near-peer threat; persistent engagement; users.

Department of Defense (DoD) leaders and numerous reform studies criticize the acquisition process for its inability to control spiraling costs and delays in getting systems to the user. In response, acquisition reform efforts to date focus on better requirements, efficient resources planning, as well as increasing feedback and accountability.[1] DoD Test and Evaluation (T&E), comprising the formal processes, policies, personnel, equipment, facilities, and consumables necessary to develop, certify, test and evaluate defense systems for production, has faced much of this criticism. In response, T&E reform focuses on process streamlining, reducing overhead, and further integration within the engineering process to better support the "buy" decision. The defense acquisition process, using T&E as a gatekeeper, is relatively unchanged since the Packard Commission recommendations.[2] Yet, the methods and players of DoD acquisition have fundamentally changed. Just as the Pentagon is embracing a new map for the application of military power based on an active strategy for the world, as it is today, DoD needs a similar active strategy for who is best served by T&E.[3]

This article examines the fundamental influences to Acquisition T&E and the results of major studies to date on reforming Acquisition T&E, and concludes that the current emphasis on "buying" as the "raison d'être" for T&E must be replaced with a new two-tiered framework and leadership that better support both contractual necessity and operations in the field. There is a new world of defense systems acquisition driving the need for a new map for T&E based on acceptance and operation, which I suggest are the *Alpha* and *Omega* of a new T&E order.

Good intentions

Acquisition T&E follows a linear engineering approach to reduce risk, build insight to meet contract delivery, and assess the delivered configuration in operational environments, verifying that the system works warranting further procurement.[4] Acquisition T&E is Service- and system-centric, managed through dedicated funding and contract vehicles with both developmental (DT) and operational testing (OT) supporting the "buy" decision. Public Law delays the decision until a "Beyond LRIP Report" is published following OT.[5]

Yet for all its formality, T&E plays a weak role in controlling what the Department actually buys. The current DoD Acquisition policy allows initiation of low-rate initial production (LRIP) just after Critical Design Review. What test results are available support this decision, but completion of testing or successful results are not formal prerequisites.

There is no DoD "*Approval for Service Use (ASU)*" decision based on successful test results. The only legal

The views expressed in this article are those of the author and do not reflect the official policy or position of the National Defense University, the Director, Operational Test and Evaluation, the Department of Defense, or the U.S. Government.

hurdle to proceeding beyond LRIP is a report by the Director of Operational Test and Evaluation (DOT&E) of an operational test where the test plan was approved in advance by DOT&E.[6] Again, as with LRIP, there is no requirement to pass the test, only to conduct it. Acquisition T&E today operates within a "buy" construct that neither demands minimal user-oriented testing before beginning the production process nor formalizes the full commitment to production through successful completion of testing.

Emphasis on speed

DoD Acquisition reform since the mid-1990s emphasizes faster cycle times through efficient management, capitalizing on emergent technologies, fielding of early capabilities, and continuous product evolutionary cycles. The objective is to maintain a competitive edge by getting to the "buy" decision faster. Yet, in World War II (WWII), compelling need put emphasis on production, with T&E supporting it. Production changes, additional requirements, and performance shortfalls based on experience in the field were the foundations for block upgrades.[7] Over 60 years later, the Mine Resistant Ambush Protected Vehicles procurement mimics this approach with an emphasis on T&E supporting production.[8]

DoD works to balance procurement for both a longer-term near-peer threat as well as near-continuous engagement against a less-defined extremist threat.[9] Senior leaders rely on the current acquisition construct in favor of one that responds better to ill-defined threats, requirements that grow and change rapidly, and technologies that evolve many times within the development cycle.[10] Today, as in WWII, the focus is getting the right capability to the field faster, but speed-to-user is not enough to drive significant change in T&E.

A new environment, really

There have been profound changes to DoD systems development and acquisition brought about by industry practice and by government policy that sets the new environment for T&E.[11] The following provides an illustrative snapshot:

- **Requirements process.** Requirements have steered away from the primacy of technocrats, planners, and buyers to the current end-user.[12] Component Commanders present unique challenges, as they focus on near-term needs and have different visions for how requirements are met and how advanced technology can be used. These users are less concerned with technology nuances, industrial influences, and specific capabilities. Yet, to manage successful acquisition, specificity is critical for configuration design and engineering.
- **Systems development.** Systems are more complex, and the ability to characterize fully end-state performance before fielding is a challenge. Lead systems integrators have increasingly less insight into the subsystems they are integrating and thus less confidence in understanding, and certifying to, actual systems-of-systems performance. The burden increasingly falls back to the Department to resolve, with risks not only from increasingly complex systems-of-systems integration with differing maturity but also from the globalization of defense industrial capability.
- **Industrial base.** The consolidation of the defense industrial base through the 1990s has left DoD with fewer options for competitive development of major capital systems. This drives systems to take on inherent design, engineering, production, and management practices with less government insight. Key components and materials will increasingly be available only from foreign sources with subsequently less control of the design and engineering.[13] The emphasis shifts from preproduction to as-delivered product adequacy.
- **Mismatched acquisition strategies.** The recently signed DoD 5000.02 Instruction relies on technologies being wrung out before initiating development, competitive prototyping used to find the "best of breed," and due diligence through T&E before production. Yet, technology evolves too quickly to tie acquisitions to fixed baselines with initial units differing in performance and utility from those later on in production. Lead systems integrators today deliver systems comprising subsystems in various levels of maturity and product life cycles.
- **Non–Service-unique systems.** In spite of the rhetoric, DoD does not buy capability; it buys "things" (systems) that are married to others, and users, to form war-fighting capabilities. While DoD is procuring more "joint" systems than ever before, the vast majority are still Service-centric asked to operate in ever increasingly joint environments.[14] It has become more difficult to characterize one system's adequacy for its own acquisition decision without interdependent systems that in and of themselves are of varying maturity levels.
- **Networked operations.** National defense strategy reflects ever-increasing multi-Service and Coalition operations.[15] These self-forming oper-

ations preclude fully understanding interfacing systems performance or concepts of operations to support an adequate operational test in advance of fielding decisions. While the Department's *Testing in a Joint Environment Roadmap* of 2004 set a vector to lash together the disparate testing capabilities within the Department and Industry, it can only go so far given ever-changing configuration baselines and unpredictable alterations of netted combat systems.[16] Testing and subsequent evaluations will focus more on in-theater assessments.

- **Expanded acquisition authority.** Once the domain of major Service Commands, acquisition authority has spread to user and mission-centric organizations such as Missile Defense Agency and Special Operations Command, each with their own processes. As such, systems developers will have less confidence in their system's performance as they have less insight into, or control of, interfacing systems. T&E will less likely depend on a priori knowledge of full system capability and default to rudimentary baseline assessments.
- **T&E beyond the Program Manager (PM).** System complexity and interconnectivity means that testers will find it difficult to build a test scenario that characterizes all desired performance points within shorter development time frames. Added is less confidence in the threat or expected concepts of operations as each user will likely tailor operations to their own needs. Acquisition will increasingly rely on tests outside the PM's control to build just enough insight for the decision needed. Capitalizing on other data from which to build consensus is key to Integrated T&E methodologies of the revised DoD 5000.02.[17]
- **Services are oversight.** Service T&E has undergone massive consolidation since the mid-1990s, which is not likely to be reversed anytime soon. Since 2000, the U.S. Army has consolidated much of its T&E organization and reduced its workforce between half and two-thirds. The U.S. Navy reduced personnel and substantially integrated its prime contractor/government testing. The U.S. Air Force further shifted DT control to prime contractors with commensurate reductions in its workforce.[18] The burden of conducting traditional Service DT has fallen more onto the contractors as part of the product acceptance process.

The above "snapshot" shows not only what ways have changed when the Department acquires its systems, but also that most aspects cannot be addressed without fundamental change in T&E. Much of what exists in today's "new" methodologies to help nudge the acquisition process along in this new environment ignores the Acquisition T&E world as it has become. The current processes quickly succumb under the weight of the endless reviews and forums.

No real change

DoD conducted three comprehensive studies on weapon system acquisition, with emphasis on T&E, to include the Defense Acquisition Performance Assessment (also known as [aka] DAPA Report), the Defense Science Board Task Force on Developmental Test and Evaluation (aka DSB Report on T&E), and the Joint Defense Capabilities Study (aka Aldridge Study). These authoritative studies produced a myriad of recommendations for T&E and Acquisition. While each report had its emphasis, their findings and recommendations for T&E were generally similar and grouped into four broad thrusts:

- gain organizational efficiencies by blurring the distinction between DT and OT,
- push discovery earlier in the process through more rigorous testing up front,
- increase transparency and streamline process overhead, and
- better utilize the planning and acquisition processes for joint war fighting needs.

All three studies hoped to efficiently push T&E to better support the "buy" decisions through process streamlining and combining DT and OT events where possible as part of an integrated T&E framework feeding a continuous thread of discovery. These reports form the basis of the DoD 5000 Integrated T&E strategy. Yet, "integration" is fundamentally an efficiency exercise of questionable purpose as early-on schedule and cost avoidance are lost through later rework and retest.

Unfortunately, neither report reassessed T&E's role or its customer. While testing is fundamental to systems engineering and contractual compliance, Acquisition T&E is seen as a "speed bump" to procurement. In trying to serve many masters, T&E became costly, less efficient, and its reports of questionable utility to both buyers and users.

A future of many masters

Future Acquisition T&E must support two acquisition extremes, the quick-reaction, less-defined threat and the long-term, near-peer threat.[19] It must also support near-term contractual necessities as well as longer-term product life-cycle processes. The emphasis is on early capability delivery for initial fielding. For many complex systems, the Department will only begin to understand what it has received once it is delivered and operating in the field. The Missile Defense Agency recognized this and developed T&E processes to support dedicated knowledge points that now form the basis for the revised DoD 5000.[20]

As in WWII, we find ourselves with users forward deployed and persistently engaged and needing 75 percent solutions in months. DoD must be more efficient and effective in getting information to the user and feedback from operations in the field.

A new model

Utility, not buying

Acquisition T&E must focus on its mission, not function, to support Acquisition and system complexity as they are today. T&E must uncover critical risks prior to initiating a program or, once begun, build knowledge to trade off risk. The focus today is to provide capability as soon as it is ready, with T&E the primary mechanism for fielding the right capability at the right time. This new model proposes it be separate from, but affiliated to, the buying decision.

While we acquire systems through the buying process, it is capabilities based on aggregates of constantly evolving systems that are delivered to the user. Authority to initiate development has become the initial production approval point reflecting the national commitment it is. The Acquisition process is no longer the tidy affair it once was. Yet, it is how DoD responds that is the basis for a new T&E model, which shifts emphasis from "buying" to the more relevant product acceptance and operational domains.

An "Alpha-Omega" model

The new model for T&E shifts the emphasis from buying to two basic, but not necessarily sequential, domains. The first includes activities to characterize sufficiently systems in support of contractual necessities, management, and initial fielding decisions—the world of acceptance tests or "Alpha Tests." The second includes the operational assessments made some time later to assess mission value added over the fielding life cycle or "Omega Tests."

The vision is a T&E process that accelerates the delivery of initial DoD capability by developers, while ensuring continuous evaluation of performance in the field for current operations and future capability development. This approach supports acquisition and life cycle activities such as the Department's Performance-Based Logistics and Training.

Alpha testing

Alpha Testing events are necessary to meet contractual requirements by capturing initial baseline capability for Service use. "Alphas" comprise all initial experiments, contractor development tests, quality tests, Service-unique interface and environmental compliance tests, security and accreditation tests. They are the necessary blending of contractor tests (CT) used to support delivery to the government with the traditional Service-oriented interface testing (DT) performed later on. Alpha Testing is a continuous aggregate of events, not necessarily fully completed events or *pass-fail* by their structure. Alphas are "owned" by a much broader community of stakeholders and not necessarily under any one single agent's control. Their results form the basis for decision gates ultimately for service use. Alpha Tests provide the basis for understanding delivered items at the time of delivery, not necessarily against a priori baseline parameters.

An Alpha Test construct capitalizes on all existing data sets, whether or not contractor derived, and is not throttled by concerns over the color of money, contracts, or ownership. It feeds on other Service efforts, direct and indirect PM efforts, training and fielding activities. Alphas fill the bin of system knowledge regardless of source. Alphas provide the PM, and those of affiliated efforts, the freedom to select the appropriate data from which to argue the case for delivery, up to and including Approval for Service Use (ASU). Where there is lack of data, the PM is obligated to fill the void or ensure that others do their share to help build the case for ASU.

An Alpha approach requires involvement by customers, users, test and oversight agents for insight and advice where practical or necessary given their control over ASU. Less oversight is required during Alpha Testing as the burden falls on the PM to build the case to deliver the incremental capability to the next user or integrator in the chain. This methodology is consistent with that used by sub-tier vendors delivering subsystems to the lead systems integrators and consistent with the Department's Systems Engineering Guide.[21]

Omega testing

Omega Tests are those scripted and unscripted, supervised and unsupervised demonstrations of systems operation in the field. Users, operational test agents, oversight, training, logistics, and doctrine agents focus on system utility and are less concerned with the buying decision. Omega Testing capitalizes on data and experience in the field, not as pass-fail since the Department has long since committed to the program, but to build on the baseline understanding of capabilities and limitations at ASU. Omega feedback also forms the basis for the next capability increment or decision to move on to new capabilities. Data and insight, through formal reports, assessments, or observations, are provided to the community at large to include operations research, requirements generators, product life-cycle managers, Program/Project Managers, oversight entities such as Service Chiefs and DOT&E, and training and doctrine agents.

A significant issue using today's OT&E construct is pegging deficiencies uncovered in complex systems-of-systems tests for a product-centered acquisition process. An Omega strategy broadens the responsibility as these events are funded through a myriad of single and combined sources to include Component Commanders, Training and Doctrine Commands, Research, Logistics and Engineering Activities, Intelligence Agencies, Programs, and other Service Acquisition agents.

This approach expands the community of Omega agents far beyond that limited by the Service Operational Test Activity (OTA) and removes the "black hat" image of today's operational testers. There would be less concern by Acquisition principals that OTA input blurs the role between system buying and fielding. Nevertheless, U.S.C. Title 10 must be revisited given the requirement to conduct an OT, and for DOT&E, an independent operational assessment must be made, prior to proceeding beyond LRIP. It is likely DoD will need consensus with Congress to either formalize a supervised period of Alpha testing on basic systems to support independent reporting or use the first Omega evaluation as the gatekeeper to further cross-Service capabilities. The latter would seem more appropriate as Congress and the DoD get a better picture of capabilities fielded and future needs with effectiveness judged through a broader evaluation lens.

Organizing to the Alpha and Omega

Service field activities would continue to function as life-cycle agents and as centrally or directly funded Alpha testers, supporting any Alpha event whether Service-specific or at Contractor sites. Alpha, being nonpartisan, can be managed either before formal program initiation, during program phases, or as part of postproduction life-cycle support. Much of this structure is already in place as test personnel at Department Major Range and Test Facility Base activities are direct customer-funded operations.

Service OTAs, freed from the grip of the acquisition process, support customers of all types. OTA and Omega would be funded through a much broader array of customers less tied to programs. The expeditionary OTA, or other agents tapped for such roles, deploy to theaters of operations or specific test sites to act as user test or evaluation agents. A much smaller senior cadre would be reserved for overseeing Alpha events supporting ASU decisions through working arrangements with program offices. Their portfolio of products and services would be greater than current program-centric assessments. The OTAs would be managed by the Services, overseen by DOT&E, and free to expand their operations worldwide to include foreign systems. This new and expanded role sets the OTA on a path to supporting future war fighting capability.

The emphasis is on empowering, with responsibility based on a closer working relationship between the developers and users. The Alpha-Omega strategy relies on three simple rules by which to frame progress and argue for ASU when appropriate.

1. What war fighting capability is provided (i.e., not the "thing" being procured)?
2. To what degree does it work, and how do you know (i.e., capabilities/limitations as delivered)?
3. What are the impacts to other systems (i.e., risk assessment across the Doctrine, Organization, Training, Materiel, Leadership and education, Personnel, and Facilities domains)?

Empowered testers

Testers and evaluators increase their impact on new program vectors. The emphasis is not on whether systems are good enough to buy (as they are already being bought) but rather what new vector must be set based on performance and deficiencies observed. OTAs plan, manage, and oversee Omega Tests as well as assess capability in the field, working with the users to vet future capabilities, upgrades, or changes to doctrine and concepts of operations. A new Joint Omega Executive provides both independent and collaborative insight of systems-of-systems operations in the field to support capability increments.

The right oversight

The Under Secretary of Defense for Acquisition, Technology and Logistics—through either the Director for Systems Engineering or the new Director for Developmental T&E[22]—oversees Alpha activities, ensuring that adequate insight and progress supports (along with user input) the decision when capabilities are fielded. This leader would also ensure sufficient capacity, training, and capability exists at T&E facilities. The DOT&E would oversee all Omega testing, advise on operational realism for Alpha events, and continue to report independently to Congress. Oversight agents would focus less on technical detail and more on validating that the achieved capability is usable and understood by the users.

A new Test and Evaluation Master Plan (TEMP)

A cottage industry has been built around TEMP production to document the PM and OTA commitment for T&E. Yet, in the pace of today's programmatic change, the TEMP becomes rapidly outdated. Under this process, the TEMP merely outlines the PM's top-level strategy for the next ASU decision focusing on system maturity, external resources, and likely data collection points for ASU.

Conclusions

The recently revised DoD instruction on acquisition strengthens the primacy of fielding, through acquisition, with T&E primarily supporting the latter. This article proposes a new map for Acquisition T&E supporting today's persistent engagement as well as the next near-peer threat. The DoD cannot wait for optimal solutions before fielding capabilities or rely solely on T&E as its gatekeeper. This new Alpha-Omega strategy, based on acceptance testing for delivery and operational use evaluations in the field, is on par with Acquisition as it is today, not on how we wish it to be. This strategy recognizes and accepts T&E's core role in engineering and contract compliance, as well as T&E's ultimate customer—the user.

This article examines how the acquisition environment has changed and how the process itself has evolved as it continues to adapt to this new reality. Nevertheless, recent authoritative studies on T&E have not recognized these fundamental changes in the landscape and have only recommended modest changes to T&E processes to speed it up a bit and make it cost a little less. T&E must emerge from its relegated shadows in acquisition to support a new customer set. The Alpha-Omega strategy hopes to change this by shifting the traditional OTA role out of the "buy" process into the more relevant fielding process as the agent of choice for a much wider set of customers to include not only Service acquisition and life-cycle agents, but also Component Commanders, Trainers and Doctrine agents, and Requirements developers.

The Alpha-Omega strategy for T&E supports bringing capability to the field faster, with better understanding of capabilities and limitations, across a broader set of systems of systems than current methodologies, streamlined or not, can ever do. The time is right for fundamental change. ❑

GEORGE AXIOTIS currently leads the Integrated Resources Analysis Team for the director, Operational Test and Evaluation (DOT&E) providing Congressionally mandated assessments of T&E resources used for Defense Department testing as well as oversight of Service test resources investments. Mr. Axiotis, a recent graduate of the Industrial College of the Armed Forces, has over 26 years in defense acquisition and T&E having been program, project, and test manager on four Defense Acquisition programs, a program sponsor as well as T&E advisor to two program executive officers. Before coming to the Pentagon, George was director of the Naval Sea Systems Command Test and Evaluation Office. E-mail: george.axiotis@osd.mil

Endnotes

[1]Department of Defense, "Joint Defense Capability Study of 2004". aka, "Aldridge Study: Joint Defense Capabilities Final report" (2004): 2–7. https://acc.dau.mil/CommunityBrowser.aspx?id=21805 (accessed April 17, 2009).

[2]Department of Defense, "President's Blue Ribbon Commission on Defense Management, National Security Planning and Budgeting" (aka The Packard Commission Report), 30 June 1986. President's Blue Ribbon Commission on Defense, Formula for Action: A Report to the President on Defense Acquisition. February 28, 1986 (1986): 11. www.ndu.edu/library/pbrc/pbrc.html (accessed April 3, 2009).

[3]Barnett, Thomas P. 2004. *The Pentagon's new map, war and peace in the twenty-first century*. New York: Berkley Books.

[4]Department of Defense, "Department of Defense Instruction, DoD 5000.02, Operation of the Defense Acquisition Process" 8 December 2008, (2008): Part 1: 12, also Enclosure 6: 51–53. http://www.dtic.mil/whs/directives/corres/pdf/500002p.pdf (accessed April 12, 2009).

[5]United States Code, Title 10, Subtitle A, Part IV, Chapter 141, Section 2399. "Operational Test and Evaluation of Defense Programs," Director, Operational Test and Evaluation (DOT&E), specifically: Section 2399: "The Secretary of Defense shall provide that a major defense acquisition program may not proceed beyond low-rate initial production until initial operational test and evaluation of the program is completed." www.law.cornell.edu/uscode/10/usc_sec_10_00002399-000-.html (accessed March 7, 2009).

[6]Ibid., DOT&E Responsibilities.

[7]Gropman, Alan. 1997. *The big 'L': American logistics in World War II*. Washington, D.C.: National Defense University Press.

[8]General Accountability Office. 2008. *Rapid acquisition of the Mine Resistant Ambush Protected Vehicles*. Washington, D.C.: GAO.

[9]Chao, Pierre. 2009. "Center for Strategic and International Studies (CSIS) Study on Defense Industrial Planning." Briefing to the Information Technology Association of America, 26 February 2009, Washington, D.C.

[10]James E. Cartwright, Vice-Chairman Joint Chiefs of Staff, General USMC. Briefing to the Senate Armed Services Committee on Systems Acquisition, 16 June 2009. http://www.armed-services.senate.gov (accessed July 2, 2009).

[11]Adolf, Pete, et al., "Developmental Test and Evaluation Defense Science Board Task Force Study." Department of Defense (2004): 6, 16. https://acc.dau.mil/GetAttachment.aspx?id=229302&pname=file&aid=37075&lang=en (accessed June 6, 2009).

[12]Department of Defense, Joint Defense Capability Study of 2004. aka, "Aldridge Study: Joint Defense Capabilities Final Report" (Joint Defense Capabilities Study January 2004): 2–7. https://acc.dau.mil/CommunityBrowser.aspx?id=21805 (accessed June 3, 2009).

[13]Guay, Terrence R. 2007. *Globalization and its implications for the defense industrial base*. Carlisle Barracks, PA: U.S. Army Strategic Studies Institute.

[14]Director of Operational Test and Evaluation. Deputy Secretary of Defense. 2004. "Department of Defense Testing in a Joint Environment Roadmap, Strategic Planning Guidance. Fiscal Years

2006–2011, Final Report, November 12, 2004". Washington, D.C.: Department of Defense (2004): 8, 12–14. http://www.dote.osd.mil/reports/TestinginaJointEnvironment-Public111204.pdf (accessed June 12, 2009).

[15]Department of Defense. "National Defense Strategy 2008": 17 www.defenselink.mil/news/2007 (accessed June 1, 2009).

[16]Ibid., 8, 18.

[17]Department of Defense. The three studies are: "Defense Acquisition Performance Assessment Summary (DAPA), A Report by the Assessment Panel of the Defense Acquisition Performance Assessment Project, For the Acting Deputy Secretary of Defense" (Defense Acquisition, Performance Assessment, Executive Summary, December, 2005): 5. http://www.defenselink.mil/pubs/pdfs/DAPA%2012-2%20WEB%20Exec%20Summary.pdf (accessed June 14, 2009). Includes "Defense Science Board Task Force Study: Developmental Test and Evaluation" (aka The Pete Adolph Study named after former Deputy Under Secretary of Defense, the Honorable Pete Adolph): and the Department of Defense, "Joint Defense Capability Study of 2004," aka "Aldridge Study: Joint Defense Capabilities Final Report", Joint Defense Capabilities Study January 2004.

[18]Adolf, Pete, et al. "Developmental Test and Evaluation Defense Science Board Task Force Study of 2004" (Defense Science Board 2004): 4–5. https://acc.dau.mil/GetAttachment.aspx?id=229302&pname=file&aid=37075&lang=en (accessed May 7, 2009).

[19]Department of Defense. "National Defense Strategy 2008": Washington, D.C.: DoD (2007): 8, 15–17. http://www.defenselink.mil/pubs/2008nationaldefensestrategy.pdf (accessed June 1, 2009).

[20]Honorable John J. Young, Jr. Under Secretary of Defense (Acquisition, Technology, and Logistics). Statement Before The House Armed Services Committee Subcommittee On Strategic Forces, April 17, 2008. www.dod.mil/dodgc/olc/docs/testyoung080417.pdf (accessed May 20, 2009).

[21]Department of Defense. Office of the Director, Systems and Software Engineering (SSE). "Systems Engineering Guide for Systems of Systems, Version 1.0." (AT&L/DSSE August 2008): 21, 24–25. http://www.acq.osd.mil/sse/docs/SE-Guide-for-SoS.pdf (accessed May 14, 2009).

[22]United States Senate, S-454, Senator Levin and Senator McCain. Senate Bill 454; "The Weapons Systems Acquisition Reform Act of 2009". Specifically, Sec 102, Director, Developmental Test and Evaluation. http://thomas.loc.gov/cgi-bin/query/z?c111:S.454 (accessed June 3, 2009).

Bibliography

Adolf, Pete, et al. *Developmental Test and Evaluation Defense Science Board Task Force Study of 2004*. https://acc.dau.mil/GetAttachment.aspx?id=229302&pname=file&aid=37075&lang=en (accessed April 3, 2009).

Barnett, Thomas P. 2004. *The Pentagon's new map, war and peace in the twenty-first century*. New York: Berkley Books.

Cartwright, James E. (Vice-Chairman Joint Chiefs of Staff). Briefing to the Senate Armed Service Committee on Systems Acquisition, 16 June, 2009. http://armed-services.senate.gov (accessed July 17, 2009).

Chao, Pierre. Center for Strategic and International Studies (CSIS) briefing to the Information Technology Association of America, 26 February 2009, Washington, D.C.

Department of Defense. Defense Acquisition Performance Assessment Summary (DAPA), A Report by the Assessment Panel of the Defense Acquisition Performance Assessment Project, for the Acting Deputy Secretary of Defense. Defense Acquisition, Performance Assessment, Executive Summary, 2005. http://www.defenselink.mil/pubs/pdfs/DAPA%2012-Department of Defense (accessed May 12, 2009).

Department of Defense Instruction, DoD 5000.02, Operation of the Defense Acquisition Process, 2008. http://www.dtic.mil/whs/directives/corres/pdf/500002p.pdf (accessed May 15, 2009).

Department of Defense. Joint Defense Capability Study of 2004. aka "Aldridge Study: Joint Defense Capabilities Final Report," Joint Defense Capabilities Study, 2004. https://acc.dau.mil/CommunityBrowser.aspx?id=21805 (accessed May 12, 2009).

Department of Defense. "National Defense Strategy 2008". Washington DC., 2007. http://www.defenselink.mil/pubs/2008nationaldefensestrategy.pdf (accessed April 6, 2009).

Department of Defense. Office of the Director, Systems and Software Engineering (SSE), Systems and Software Engineering. August 2008. Systems Engineering Guide for Systems of Systems, Version 1.0, 2008. http://www.acq.osd.mil/sse/docs/SE-Guide-for-SoS.pdf (accessed May 16, 2009).

Deputy Secretary of Defense. Department of Defense Testing in a Joint Environment Roadmap, Strategic Planning Guidance. Fiscal Years 2006–2011, Final Report, November 12, 2004. Director of Operational Test and Evaluation, Washington D.C., 2004. http://www.dote.osd.mil/reports/TestinginaJointEnvironment-Public111204.pdf (accessed June 1, 2009).

General Accountability Office. 2008. *Rapid acquisition of the mine resistant ambush protected vehicles*. Washington, D.C.: GAO.

Gropman, Alan 1997. *The big 'L': American logistics in World War II*. Washington, D.C.: National Defense University Press.

Guay, Terrence R. 2007. *Globalization and its implications for the defense industrial base*. Carlisle Barracks, PA: U.S. Army Strategic Studies Institute.

President's Blue Ribbon Commission on Defense Management, National Security Planning and Budgeting (aka The Packard Commission Report), 30 June 1986. President's Blue Ribbon Commission on Defense, Formula for Action: A Report to the President on Defense Acquisition, 1986. www.ndu.edu/library/pbrc/pbrc.html (accessed May 23, 2009).

United States Code, Title 10, Subtitle A, Part IV, Chapter 141, Section 2399. Operational Test and Evaluation of Defense Programs. www.law.cornell.edu/uscode/10/usc_sec_10_00002399-000-.html (accessed May 16, 2009).

United States Senate, S-454, Senator Levin and Senator McCain. Senate Bill 454; The Weapons Systems Acquisition Reform Act of 2009. Specifically, Sec 102, Director, Developmental Test and Evaluation. http://thomas.loc.gov/cgi-bin/query/z?c111:S.454 (accessed May 16, 2009).

Young, John J. Jr., Under Secretary of Defense (Acquisition, Technology, and Logistics). Statement before the House Armed Services Committee Subcommittee on Strategic Forces, April 17, 2008. www.dod.mil/dodgc/olc/docs/testyoung080417.pdf (accessed May 16, 2009).

ITEA Journal 2009; 30: 481–491

Climb Trajectory Prediction Software Validation for Decision Support Tools and Simulation Models

Jessica Romanelli, Confesor Santiago, Mike M. Paglione, and Albert Schwartz
Federal Aviation Administration, Atlantic City, New Jersey

The Next Generation Air Transportation System (NextGen) is the solution to capacity, safety, and efficiency problems that will result from an expected increase in traffic. Trajectory-based operations are identified in the NextGen Concept of Operations as a key capability required to ensure the success of NextGen; thus, it is essential that the accuracy of trajectory prediction software be tested and validated for all phases of flight. Trajectories are also modeled in fast-time simulation tools that are used to test future NextGen concepts and identify possible benefits or problems. The objective of this testing activity is to identify outliers during the climb phase of flight in the trajectory predictions of two decision support tools as well as in the trajectory modelers of two fast-time simulation models. The errors in trajectory prediction will also be examined by aircraft type in order to measure the accuracy of aircraft characteristics utilized in the tools.

Key words: Air traffic control; climb; cruise; descend; FAA; fast-time simulation; flight; Next Generation Air Transportation System (NextGen); prediction software; trajectory modeling.

Air traffic in the United States is predicted to increase threefold by the year 2025. The Next Generation Air Transportation System (NextGen) is the solution to capacity, safety, and efficiency problems that will result from this increase. The Federal Aviation Administration (FAA) is primarily responsible for the implementation of NextGen. The NextGen Concept of Operations identifies aircraft trajectory-based operations as a key capability required to ensure the success of NextGen. Four-dimensional (4-D) trajectory prediction algorithms predict an aircraft's horizontal and vertical position at some time in the future and are used for conflict detection, metering, and other applications in air traffic management decision support tools (DSTs). Fast-time simulation models also utilize 4-D trajectory modeling in research and development of new NextGen concepts. Therefore, it is essential that the accuracy of trajectory prediction software be tested and validated.

There are three phases of flight: climb, cruise, and descent. Recent National Aeronautics and Space Administration (NASA) analyses have shown that changes in an aircraft's phase of flight are associated with higher trajectory prediction errors as compared with cruising at a steady altitude. It has also been shown that errors in climb trajectory prediction differ among aircraft types (Gong and McNally 2004). The objective of this testing activity is to identify the trajectory accuracy outliers produced during the climb phase of flight by various aircraft types and other factors. Archived air traffic data from the Washington, D.C., Air Route Traffic Control Center (ARTCC) is utilized to compare the accuracy of trajectory predictors used in DSTs such as User Request Evaluation Tool (URET) and En Route Automation Modernization (ERAM), as well as those used in fast-time simulation models such as Airspace Concept Evaluation Simulation (ACES) and Reorganized Air Traffic Control Mathematical Simulator (RAMS).

DSTs aid air traffic controllers in making the safest and most efficient decisions in moving aircraft. URET was developed to help air traffic controllers safely handle a greater number of user-preferred flight profiles, increase flexibility, and increase system capacity (Mitre Corporation, 2008). ERAM combines the functionality of URET and the Host Computer System and provides the ARTCCs with surveillance and flight data processing, conflict probe functionality, and display support for the national airspace system (FAA, 2007). At the heart of these critical systems is the accuracy of trajectory predictions. Thus, analysis techniques to easily identify errors in the modeling of aircraft trajectories will help ensure these systems meet their goals of improved safety and efficiency.

Fast-time simulation is used in the validation of new concepts to obtain an understanding of potential benefits or problems. ACES is an agent-based fast-time simulator developed by NASA to "enable evaluation of the system-wide effects of proposed air transportation concepts intended to reduce delay, increase capacity, and accommodate the forecasted growth in air traffic" (NASA 2009). RAMS is developed and supported by ISA Software and features 4-D flight profile calculation, 4-D sectorization, and 4-D spatial conflict detection and resolution (ISA Software, 2008). Similar to the URET and ERAM operational systems, these simulation models also require trajectory predictions to be timely and accurate. The methods developed in this article will identify possible outliers in their trajectory modeling.

Overview of data and preparation

This activity focused on the accuracy of the trajectories created by DSTs and simulation models during the initial climb phase of flight. Thus, extensive data preparation was required to filter traffic data to only include flight tracks before their tops of climb were reached and to convert the recorded data to formats compatible with ACES and RAMS.

Base scenario

Recorded en-route air traffic from the ARTCC (referred to as ZDC) was utilized in this study. The recorded ZDC data were collected on March 17, 2005, and contained approximately 5 hours of flight data and approximately 2,200 flights. Using this recorded data, two scenarios were created to generate input files for URET and for ERAM test runs. For URET, this consisted of a single file containing Air Traffic Control (ATC) and track messages. This file was formatted into an ASCII pipe delimited version of the Host Computer System Common Message Set (HCS CMS) (FAA, 2004). This ASCII version of the CMS was developed during the URET Testing Program. For ERAM, the input scenarios consisted of two files, one containing ATC messages and the other containing the radar target message. A mode of the ATCoach simulator was invoked that reads the ATC clearances in one file and the radar data in a separate file, injecting them into ERAM and emulating the operational data flow (UFA, 2004). The two scenarios are slightly different due to the different methods of formatting the recorded data into appropriate URET and ERAM scenarios. These two test scenarios were originally created for ERAM's formal Run-For-Record (RFR) Flight Data Processing/Conflict Probe Tool Accuracy Test in August 2007 and were recycled for this experiment.

Since this focuses on studying trajectory accuracy during the climb phase of flight, the scenarios were truncated to only include aircraft that were departing in both scenarios. Of the original 2,200 flights, 627 departure flights were analyzed in this study. This filtering was performed after the scenarios were executed by their respective systems; only during the analysis of the results were flights removed.

Trajectory predictions of the decision support tools

Once the air traffic scenarios are injected into URET and ERAM, the predictions need to be captured and input into the various test tools for analysis. In order to compute the accuracy of any trajectory prediction, two data sets are needed: (a) the true (actual) flight paths, and (b) the trajectory predictions. In this study, the actual flight paths are derived from the recorded air traffic stored in CMS format. Both URET and ERAM have their own system analysis recording Synthetic Aperture Radar (SAR) capabilities. The SAR is where predicted flight paths are stored. The binary files that are produced are parsed with a set of scripts. This produces a trajectory file containing one or more 4-D trajectory predictions for every aircraft in the form of a sorted listing of the trajectory's predicted positions in time, stereographic x-y coordinates, altitude, and ground speed.

Preparation for simulation tools

RAMS and ACES are fast-time simulators that model aircraft flight paths. These tools simulate aircraft using a set of positional data, normally latitude and longitude coordinates. The tools generate a 4-D flight trajectory for every flight of a given air traffic scenario. The flight trajectories are created based on the input data and procedures specific to the individual tool; hence, they are often different between models. Since our study focuses on how close to actuality the models simulate the aircraft during the climb phase of the flight, the input flight paths for the simulation tools are created using the recorded data described in the Base Scenario section.

The scenarios needed to be prepared prior to injecting them into the models. The points of vertical transition in the scenario track data were calculated to define the time, speed, latitude and longitude coordinates, and altitude of each vertical event. A vertical event is a transition in the vertical profile from one vertical phase of flight to another. The three vertical phases of flight are level, ascending, and descending. *Figure 1a* illustrates a flight with four vertical events,

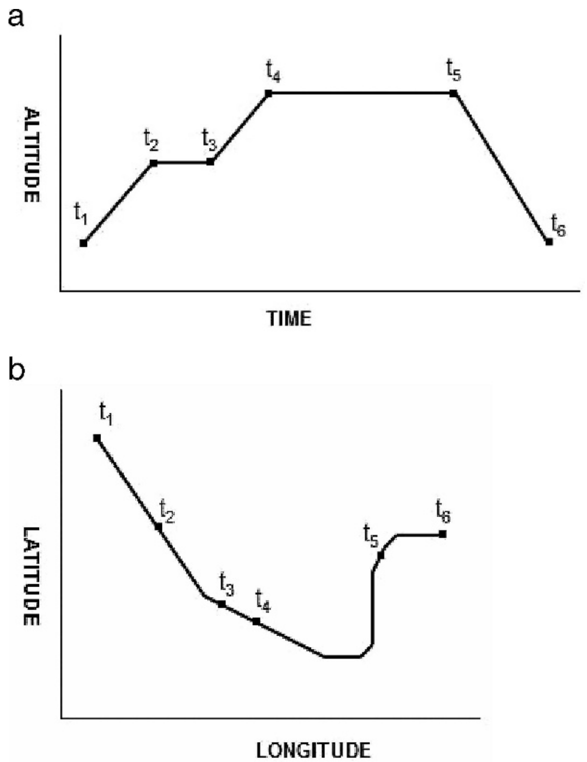


Figure 1. Vertical (a) and horizontal (b) views of vertical events.

where the x-axis is time, and the y-axis is the altitude of the aircraft. Furthermore, at time t_2 an ascent-to-level event occurs; at t_3, a level-to-ascent; at t_4, an ascent-to-level; and at t_5, a level-to-descent.

As previously mentioned, the simulations' tools use a set of horizontal coordinates to model the aircraft flight paths. In addition to the starting and ending coordinates denoted in *Figure 1a* by t_1 and t_6, respectively, the latitude and longitude coordinates at the time of each vertical event were used to produce the input air traffic. An example of the horizontal profile is presented in *Figure 1b*. The altitude and time at each vertical event is utilized in the model as well. The tools use this information to model the flights and generate their own vertical profile. This profile may deviate from the original inputted air traffic, which was based on actual operational data; thus this study focuses on measuring the amount of these differences.

Once the tools are executed, a 4-D trajectory is created for every flight. These trajectories are extracted into a trajectory file, which will be processed to measure the accuracy of the simulated trajectory when compared with the actual flight paths.

Trajectory prediction accuracy measurement

Trajectories are generated by the trajectory predictor (TP) that resides within a DST or by the trajectory modeler (TM) within a simulation tool. In DSTs, trajectories are used to alert air traffic controllers of potential conflicts in the future. Trajectories in simulation tools function as flight paths used to examine the effects of new airspace concepts. The accuracy of a TP or TM determines its overall performance. Measuring the accuracy of a TP or TM requires the actual flight paths as well as the flight paths' predicted trajectories or modeled trajectories. In order to measure the accuracy, the difference in altitudes of the actual and predicted or modeled paths is calculated. The details of these steps are described below.

Measuring vertical trajectory prediction accuracy

Trajectory prediction accuracy is measured by the difference between the trajectory predictor's path and the actual path flown by the aircraft. In order to measure this difference, the actual path of the aircraft needs to be obtained by examining radar surveillance reports and other ATC data such as flight plan amendments and altitude clearances. A set of data reduction and analysis tools is used to validate, synchronize, and store the data into database tables. Then another software tool is used to compare the inputted trajectories against the actual flight paths and calculate a set of metrics, quantifying the accuracy of the trajectory predictions, which is stored in another database table. The key metric in this study is the vertical error. Vertical error is the difference between the trajectory's predicted altitude and the actual track altitude (Paglione and Oaks 2007). *Figure 2* illustrates these two positions and the vertical error. The track altitude is labeled TK while the trajectory altitude is labeled TJ in the figure. A positive vertical error occurs when the trajectory's altitude is below the actual track altitude; hence, the error is negative when the trajectory's altitude is above the track altitude. The

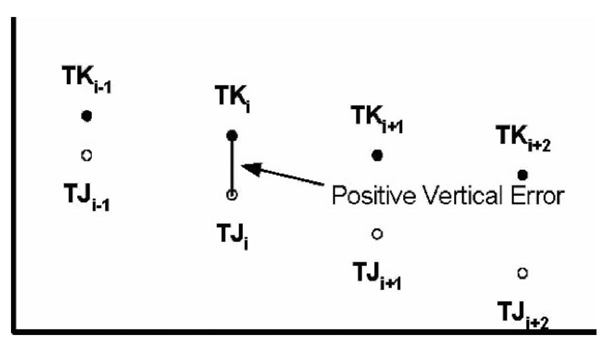


Figure 2. Vertical trajectory error. Adapted from "Implementation and Metrics for a Trajectory Prediction Validation Methodology" by Paglione and Oaks (Paglione and Oaks 2007).

errors are explored using descriptive and inferential statistics acquired by a statistical software package.[1]

Interval-based sampling

There are two parts in considering the accuracy performance trajectory predictions generated by DSTs. The first is the accuracy of a trajectory predicting the present position of an aircraft, and the second is the accuracy of a trajectory predicting the future position of an aircraft; both are extremely important in the conflict resolution process of an ATC.

Interval-Based Sampling Technique (IBST) is the trajectory accuracy sampling method developed by the Conflict Probe Assessment Team (CPAT) at the FAA William J. Hughes Technical Center and has been used in many FAA studies and test programs such as the Center Terminal Radar Approach Control (TRACON) Automation System and URET (Paglione & Oaks 2007). There are two main steps to the IBST. First, track points of an aircraft are sampled in succession a parameter number of seconds until the end of the track (see T_s in *Figure 3*). Then the trajectories are searched to find the most recent at a given sample time. Once the active trajectory is selected, the error in the trajectory is calculated iteratively for every look-ahead time value specified by the user. For the DSTs, a sample time of 120 seconds and look-ahead times of 60-second intervals from 0 to 900 seconds were used for this analysis. For the simulation portion of this study, a sample time of 10 seconds and no look-ahead times were used. Since there is only one trajectory for each aircraft modeled in the simulation tools, look-ahead time was not considered; however, sample time frequency was increased. With the combination of sample time (present time) and look-ahead time (future time) IBST creates data that can be evaluated to study the accuracy of trajectory predictions at current and future states.

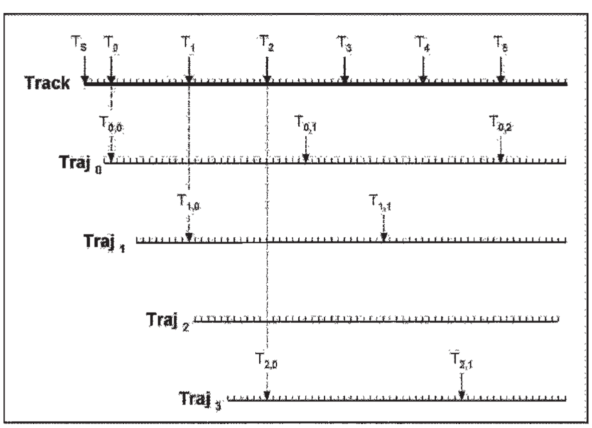


Figure 3. Time-line for the interval-based sampling technique. Adapted from "Implementation and Metrics for a Trajectory Prediction Validation Methodology" by Paglione and Oaks (Paglione and Oaks 2007).

Results

The results of this testing activity are measured by vertical error and by the absolute value of vertical error. Vertical error accounts for the direction and magnitude of the error, while its absolute value only provides the magnitude.

The following standard statistical measures will be referenced in presenting the results:

- the mean discussed is the arithmetic mean, which is the sum of the values divided by the number of values;
- the standard deviation is a measure of the variance of a set of data from its mean;
- the median is the middle of a distribution, such that half of the values in the data set are above the median and half are below; and
- N is the number of error measurements in the data set.

When the means of vertical error are discussed, it is the mean absolute value of vertical error that is being referenced since it more accurately illustrates the average magnitude of error.

Decision support tools vertical error

Initially, it was expected that the vertical errors of the DST ERAM would be less than those of URET. It is, in fact, a design requirement for ERAM to be at least as accurate as URET. Since several minor algorithmic enhancements were made in ERAM, most notably an improved radar tracking system, we intuitively expected the trajectories of ERAM to be more accurate than those of URET.

Results of the URET analysis indicated that it contained, on average, moderate vertical error that was somewhat balanced above and below the actual altitudes (see *Table 1*). The maximum vertical error for URET was 26,000 feet, and the mean vertical error for URET was 1,777.19 feet. A median error of −42 feet indicated slightly more error occurring above the true altitudes; in other words, there was marginally more error when the trajectory's altitude was above the actual altitude than there was when the trajectory was below the actual track.

The analysis of ERAM yielded similar results (see *Table 2*). The maximum vertical error was 25,941 feet, and the mean error for ERAM was 1,873.15 feet. The median vertical error was zero, which suggested the same amount of error existing above and below the actual track altitudes.

Table 1. Vertical error statistics for User Request Evaluation Tool.

Absolute vertical error	Value
Mean	1,777.19
Standard deviation	2,190.48
Minimum	0
Maximum	26,000
Median	1,000
N	30,233

Table 2. Vertical error statistics for En Route Automation Modernization.

Absolute vertical error	Value
Mean	1,873.15
Standard deviation	2,320.11
Minimum	0
Maximum	25,941
Median	1,017
N	26,494

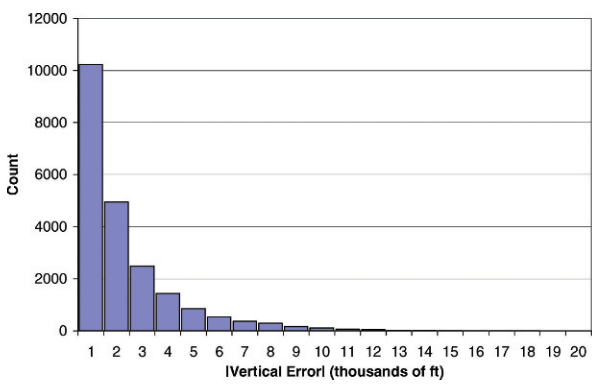


Figure 4. Histogram of absolute vertical error for User Request Evaluation Tool.

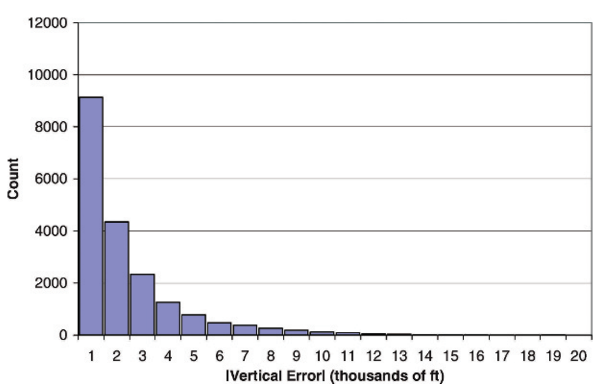


Figure 5. Histogram of absolute vertical error for En Route Automation Modernization.

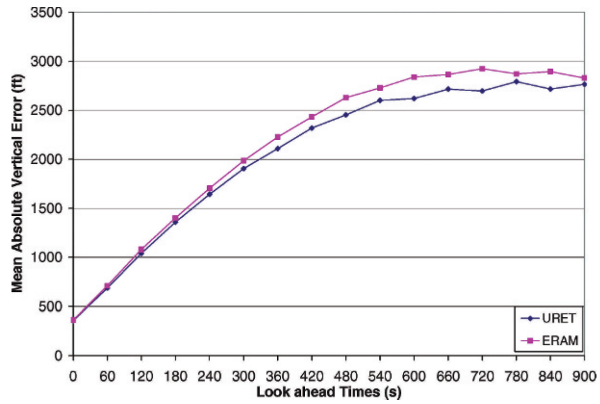


Figure 6. Plot of mean vertical error by look-ahead time.

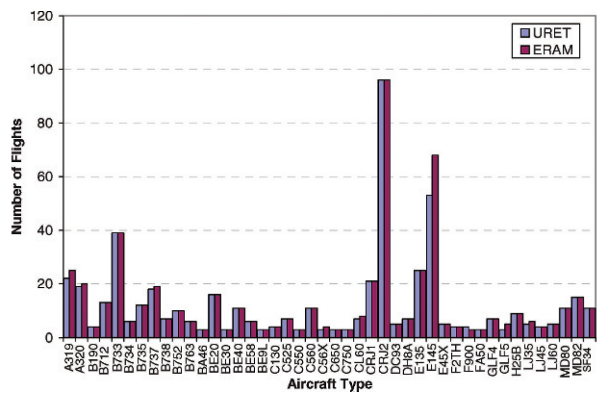


Figure 7. Histogram of frequency of flights by aircraft type in User Request Evaluation Tool and En Route Automation Modernization.

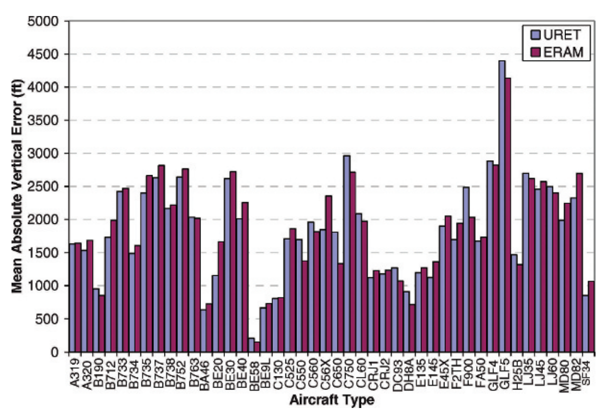


Figure 8. Histogram of absolute vertical error by aircraft types in User Request Evaluation Tool and En Route Automation Modernization.

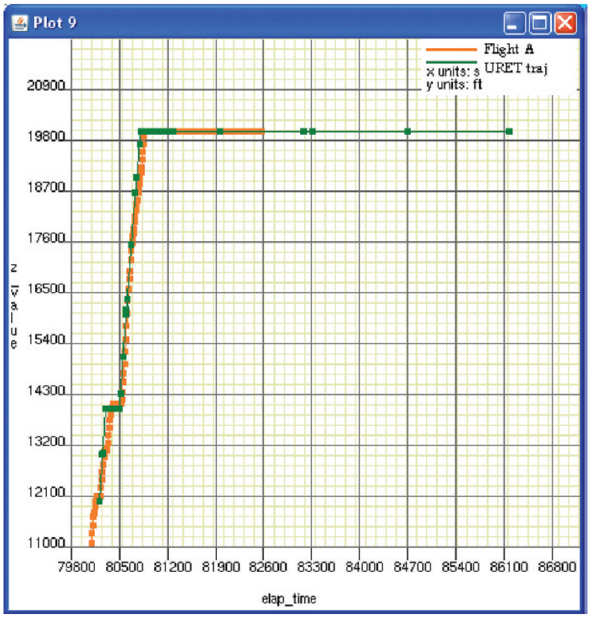


Figure 9. Time vs. altitude plot of flight A and its User Request Evaluation Tool trajectory.

As can be seen from the histograms in *Figures 4–5*, the distribution of vertical error in URET is similar to that of ERAM. Comparing the mean vertical errors of URET and ERAM, it is concluded that the trajectories of URET are slightly more accurate than those of ERAM during the initial ascent to the Top Of Climb (TOC). However, the median absolute vertical error of ERAM is very close to that of URET (within 20 feet), which indicates possible outliers in ERAM that caused the mean error to be high. Therefore, ERAM does not contain significantly more vertical error than URET.

Figure 6 shows the mean vertical error by look-ahead time for URET and ERAM. For each look-ahead time value, ERAM had a higher mean vertical error than URET. The graph confirms the statistics, which show that ERAM and URET had very similar errors, but ERAM has slightly more error than URET. It is interesting to note that for both ERAM and URET, mean vertical error and look-ahead time seem to have an exponential relationship based on the shape of the plot.

A comparison was also conducted on the errors of URET and ERAM based on aircraft type. This analysis was done on only those aircraft types that were found in output from both tools, thus the comparison was performed based on 91 different aircraft types. *Figure 7* shows the frequency of flights in URET and ERAM for each aircraft type, and *Figure 8* shows the mean absolute vertical error for each aircraft type in URET and ERAM. It is clear from the histograms that those aircraft types with high vertical errors did not occur frequently in the scenario, and those aircraft types that had a high frequency did not contain high vertical errors.

A simple difference measure was taken for the mean vertical errors of URET and ERAM for each aircraft type (URET error − ERAM error). It was determined that ERAM had a higher mean error for almost 59% of the aircraft types. Of these cases where ERAM had higher error, the average difference between URET

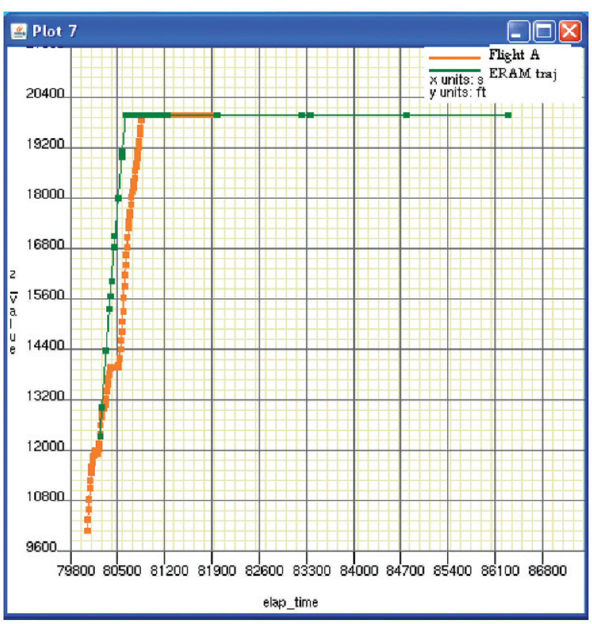


Figure 10. Time vs. altitude plot of flight A and its En Route Automation Modernization trajectory.

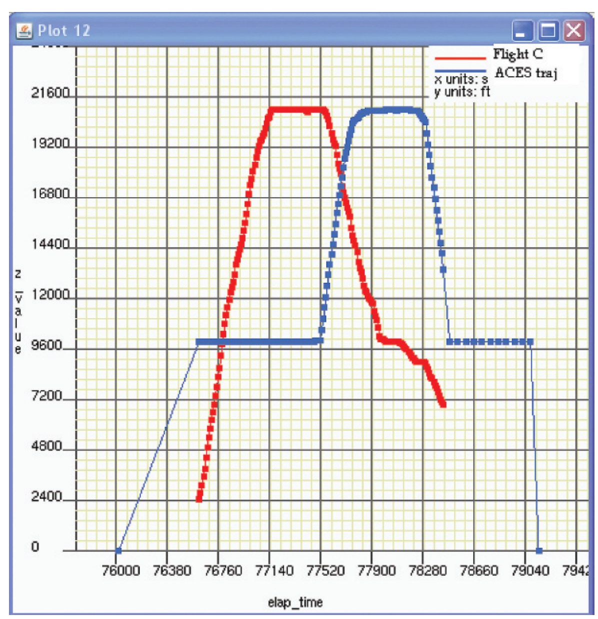


Figure 11. Time vs. altitude plot of flight C and its Airspace Concept Evaluation Simulation.

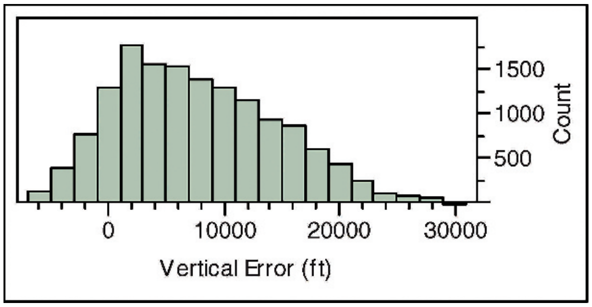


Figure 12. Histogram of vertical error for Airspace Concept Evaluation Simulation.

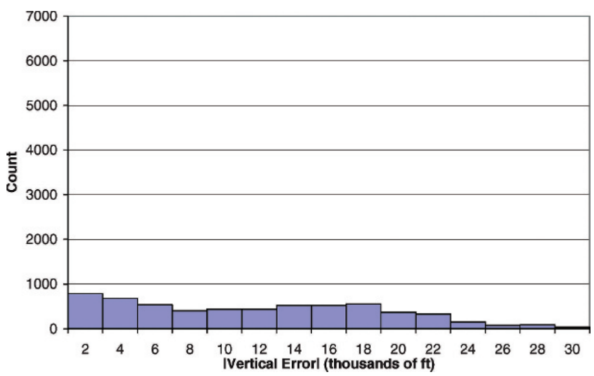


Figure 13. Histogram of vertical error for Airspace Concept Evaluation Simulation.

Table 3. Vertical error statistics for Airspace Concept Evaluation Simulation.

Absolute vertical error	Value
Mean	8,195.1
Standard deviation	6,277.09
Minimum	0
Maximum	30,000
Median	6,900
N	14,857

Table 4. Vertical error statistics for Reorganized Air Traffic Control Mathematical Simulator.

Absolute vertical error	Value
Mean	1,775.48
Standard deviation	1,705.45
Minimum	0
Maximum	10,991
Median	1,295
N	14,857

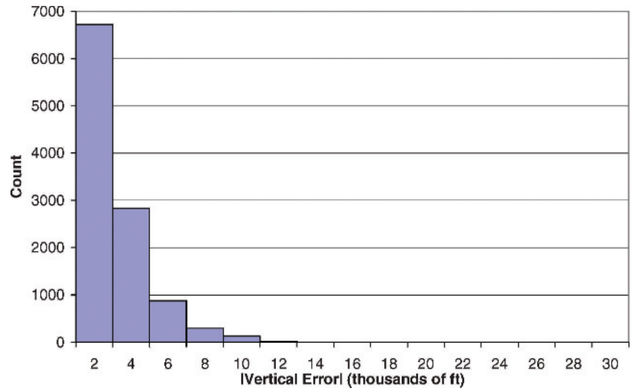


Figure 14. Histogram of vertical error for Reorganized Air Traffic Control Mathematical Simulator.

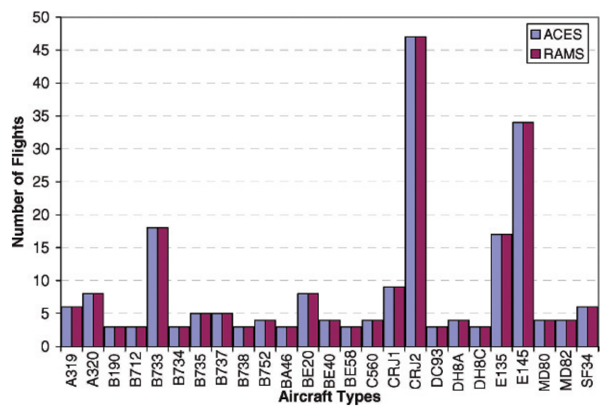


Figure 15. Histogram of frequency of flights by aircraft type for Airspace Concept Evaluation Simulation and Reorganized Air Traffic Control Mathematical Simulator.

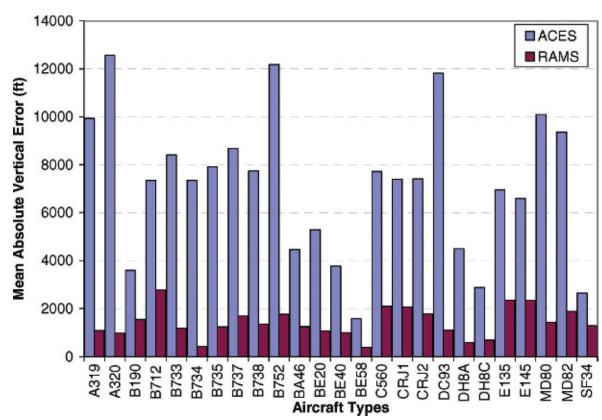


Figure 16. Histogram of mean absolute vertical error by aircraft type for Airspace Concept Evaluation Simulation and Reorganized Air Traffic Control Mathematical Simulator.

and ERAM errors was approximately 277 feet. When URET had higher error, the average difference was approximately 233 feet. This difference, although it may have been statistically significant, does not rise to the level of practical significance. For example, the aircraft's altitude data are only supplied to the nearest 100 feet. Thus, any difference of less than 100 feet is not significant; therefore, it cannot be concluded that URET and ERAM were systematically different.

Above is an example of a flight that was in both URET and ERAM. This flight contained one level period prior to reaching its TOC; however, URET captured this level period while ERAM did not. *Figure 9* shows the flight's altitude as a function of time in URET, and *Figure 10* shows the same information in ERAM. If ERAM had predicted the level phase of this flight, then URET and ERAM would have had minimal difference in vertical error for this flight.

Fast-time simulation tools vertical error[2]

Only flights with one continuous ascent to the TOC were modeled in ACES due to its inability to model interim altitudes. Thus, it was expected that the output of ACES would also include only continuous ascents. However, ACES forced all of the flights to have at least one level phase before reaching the TOC (see *Figure 11*). The level phase usually occurred at or near an altitude of 10,000 feet.

The additional level period in the ACES trajectories caused its errors to be very high. The maximum vertical error for ACES was 30,000 feet, while its mean error was 8,195.1 feet. In charts such as *Figure 11*, which plot the actual aircraft trajectory with the ACES trajectory, it appears that if the level period was not modeled in ACES, the errors would be much less. The median of the vertical error for ACES was 6,900 feet, which along with the histogram in *Figure 12*, indicates that the majority of the vertical errors occurred when the ACES trajectory was lower than the actual flight. This finding makes sense since the level period of ACES caused its trajectory to remain at a low altitude, while the actual flight continuously ascended to its TOC.

On the other hand, RAMS was able to model flights with continuous ascents as well as those which contained interim altitudes. The step climbs were modeled by creating navigational aids (NAVAIDS) at each location where a vertical event occurred (these are explained in the Preparation for simulation tools section) and instructing RAMS to reach each NAVAID at a required altitude.

Overall, results show that our expectations were met, and RAMS had minimal error. The maximum vertical error for RAMS was 13,038 feet, and the mean vertical error for RAMS was 1,807.6 feet. Its median vertical error was 0 feet, indicating an even distribution of error above and below the actual track.

In order to compare the two simulation models, the flights modeled by ACES needed to be extracted from RAMS since more flights were modeled in RAMS than in ACES. As stated above, the mean error for ACES was 8,195.1 feet. The mean error for RAMS flights that were also in ACES was 1,775.5 feet. *Figures 13–14* and *Tables 3–4* also show that the errors for RAMS were much smaller than the errors for ACES.

The output of RAMS and ACES had 66 aircraft types in common. *Figure 15* shows the number of flights for each aircraft type for ACES and RAMS. *Figure 16* shows the mean absolute vertical errors for each aircraft type in ACES and RAMS. These figures show that aircraft types that occurred frequently in the scenario had low vertical errors, and those with high vertical errors occurred very infrequently.

Results of the comparison of RAMS and ACES based on aircraft type showed that the mean absolute vertical error for ACES was higher than that of RAMS for 100% of the aircraft types. The average difference between RAMS and ACES errors was 6,109.4 feet. This is undoubtedly due to the error in ACES caused by the forced level period of each flight. It is thought that the difference between RAMS and ACES would be much smaller if ACES did not create the level periods.

Figures 17–18 show the ACES and RAMS trajectories for one flight. The flight's actual track had one continuous climb from about 7,000 feet to about 16,000 feet. *Figure 17* shows that the ACES trajectory began at ground level then climbed to about 10,000 feet where it leveled off. The ACES trajectory finally finished its climb to slightly below the flight's TOC altitude. As is evident when comparing the charts, RAMS was much more accurate than ACES in the vertical climb of this flight. The RAMS trajectory closely followed the actual track during the ascent to the TOC.

TOC time error

For each of the tools, the trajectories did not always reach the TOC altitude at the same time as the actual flight. *Figure 19* depicts one scenario where the trajectory reaches the TOC altitude before the actual flight and explains the calculation of the time error. The time error is the absolute value of the difference in the time between when the trajectory reached the TOC altitude and when the flight reached the TOC altitude.

Table 5 shows the mean, median, minimum, maximum, and standard deviation of the absolute value of time error. The means and medians of the time error followed an expected trend based on knowledge of the tools. ERAM and URET had similar low to moderate

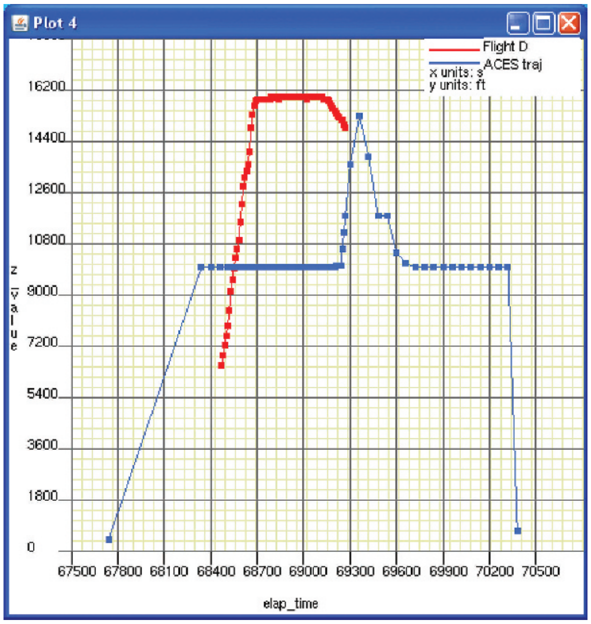


Figure 17. Time vs. altitude plot of flight D and its Airspace Concept Evaluation Simulation trajectory.

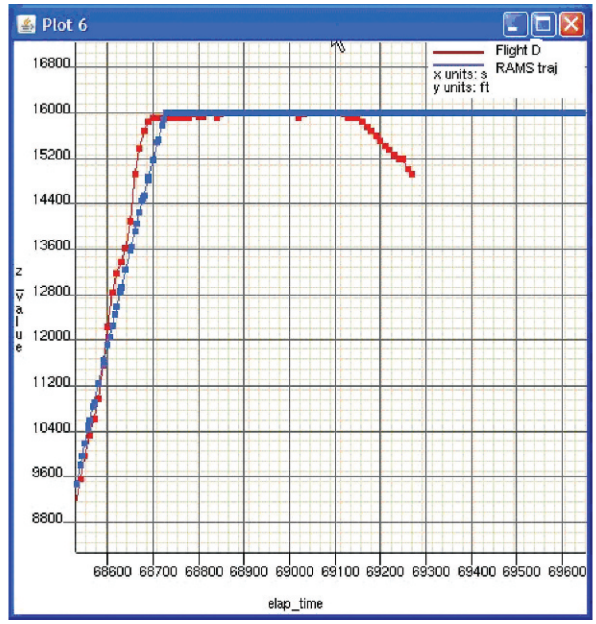


Figure 18. Time vs. altitude plot of flight D and its Reorganized Air Traffic Control Mathematical Simulator trajectory.

*Table 5. Time error statistics for All Four Tools.**

Absolute time error	URET	ERAM	ACES	RAMS
Mean	357.4333	234.7464	882.379	221.4544
Median	143	120	820	136
Minimum	0	0	270	0
Maximum	11,097	4,650	10,240	1,749
Standard deviation	639.9931	355.8754	666.8139	266.2661

*URET, User Request Evaluation Tool; ERAM, En Route Automation Modernization; ACES, Airspace Concept Evaluation Simulation; RAMS, Reorganized Air Traffic Control Mathematical Simulator.

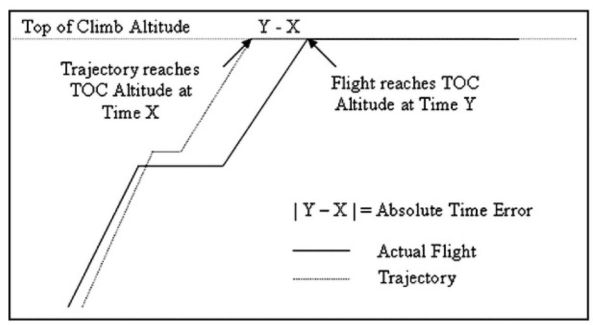


Figure 19. Top of climb time error diagram.

errors where ERAM was slightly more accurate than URET. The mean and median time errors of RAMS were also similar to those of the DSTs. Finally, as a direct result of the forced interim altitudes, ACES had very high error statistics.

Conclusion

The objective of this study was to test the accuracy of the trajectory predictor/modeler software during the ascent to TOC in DSTs URET and ERAM as well as in fast-time simulation models ACES and RAMS. Also, it is important to note that while the DSTs function as predictors of aircraft trajectories, the simulation tools do not make predictions; instead, they model the aircraft's flight path. After track data from ZDC were filtered, flagged for vertical events, and translated into ACES and RAMS format, a statistical software tool (JMP©)[6] was used to obtain the means and medians of the vertical error and absolute value of vertical error for each system.

It was found that URET and ERAM both contained a moderate amount of vertical error. The mean vertical error for URET was 1,777.19 feet, whereas ERAM had a slightly higher mean vertical error of 1,873.15 feet. The medians of absolute vertical error, though, indicate that ERAM is as accurate as URET. In a comparison by aircraft type, URET was more accurate than ERAM for over 50% of the aircraft types. This could be caused by possible outliers in the ERAM data since the data were based on its Run-For-Record results and contained some issues that caused outliers. The current version of ERAM may have resolved some of these issues.

An unexpected finding in this study was the limited capabilities of ACES in simulating actual vertical trajectories. There was no known feature in ACES that allowed the user to force a flight to level at given altitudes; as a result, only flights with a continuous ascent to their TOC were used in the ACES analysis. Also, ACES levels all flights at or near 10,000 feet possibly as a traffic management rule when aircraft are leaving the terminal area. This caused the vertical errors for ACES to be very high. However, RAMS was able to model all types of ascent profiles and, with a mean vertical error of 1,807.6 feet, proved to model trajectories that were almost as accurate as those predicted by the DST URET. When comparing the errors of ACES to those of RAMS, only flights that were in both ACES and RAMS output were considered and the results showed that the mean vertical error for ACES was 8,195.1 feet, while the mean vertical error for RAMS was 1,775.5 feet. This comparison of errors was broken down by aircraft type, and it was found that RAMS was more accurate for all of the aircraft types. However, it cannot be concluded that this was due to inaccurate aircraft characteristics since the vertical error for each aircraft in ACES was much higher than the errors in RAMS.

Overall, most of our expectations were met. ERAM was as accurate at predicting vertical trajectories as URET and, hence, fulfilled its requirements. RAMS was also proven to be roughly as accurate as URET. Finally, while ACES had high vertical errors, the cause of the errors has been determined, and it seems that if this problem had not occurred, ACES would have much smaller errors. ❑

JESSICA ROMANELLI started her career as a cooperative intern at the FAA William J. Hughes Technical Center (WJHTC) in 2006 and is currently a mathematician in the Simulation and Analysis Team at the FAA WJHTC in Atlantic City, New Jersey. Her current research interests include neural network modeling and air traffic trajectory simulation. Jessica earned her bachelor of science degree in mathematics at The Richard Stockton College of New Jersey in Pomona, New Jersey, in 2006 and completed her master of science degree in mathematical sciences at Rutgers University in Camden, New Jersey, in 2008.

Mrs. Romanelli is a member of the following professional societies: Institute of Electrical and Electronic Engineers (IEEE) and its affinity groups Women in Engineering (WIE) and Graduates of the Last Decade (GOLD), American Statistical Association (ASA), and South Jersey Human Factors and Ergonomics Society. E-mail: Jessica.Romanelli@faa.gov

CONFESOR SANTIAGO started his career as a cooperative intern for the Federal Aviation Administration and is currently a computer scientist in the Simulation and Analysis Team of the FAA located at the William J. Hughes Technical Center in Atlantic City, New Jersey. Mr. Santiago is the lead of the Rowan GUI Project, which is developmental work of visualization tools for the FAA. Furthermore, he is supporting a NextGen project investigating improvement to the separation management functions in the en route automation. He holds a bachelor of science degree in computer science and a master of science degree in electrical and computer engineering from Rowan University. Mr. Santiago has recently received the award for Best Paper in Conference at the 2008 International Testing and Evaluation Association (ITEA) Annual Symposium for the paper titled "Government and Academia Partnership to Test and Evaluate Air Traffic Control Decision Support Software." E-mail: Confesor.Santiago@faa.gov

MIKE PAGLIONE is the Conflict Probe Assessment Team lead in the FAA's Simulation and Analysis Group at the FAA William J. Hughes Technical Center, Atlantic City, New Jersey. He has extensive experience in air traffic control automation algorithms, simulation problems, analysis of decision support software, applied statistics, and general systems engineering. He is currently supporting the development, testing, and evaluation of FAA air traffic management software. Mike was FAA's Rutgers University Fellow from 1994–1996, accuracy test lead for the FAA's User Request Evaluation Tool, Program manager for the Joint University Program from 1999 to 2004, and is currently project lead on the Automation Metrics Test Working Group (a cross-organizational team developing and implementing metrics for the En Route Automation Modernization Program) and a local team lead supporting a NextGen project investigating improvement to the separation management functions in the en route automation. Mr. Paglione holds bachelor of science and master of science degrees in industrial and systems engineering from Rutgers University. E-mail: Mike.Paglione@faa.gov

ALBERT SCHWARTZ is an operations research analyst in the Simulation and Analysis Team at the FAA William J. Hughes Technical Center in Atlantic City, New Jersey. He has over 20 years of experience in modeling and simulation. Mr. Schwartz has a bachelor of science degree in information systems. E-mail: Albert.Schwartz@faa.gov

Endnotes

[1]JMP is developed by the SAS Institute and used here for all statistical calculations, see www.jmp.com for details.

[2]Testing was performed on RAMS version 5.29.06 and ACES version 510_v4.

References

Federal Aviation Administration (FAA) 2004. WJHTC/AOS-300, "ARTCC HOST Computer System Interface Requirements Document (IRD) for Air Traffic Applications," *NAS-IR-8217-0001*, Atlantic City: FAA, October 2004.

FAA, 2007. “En Route Automation Modernization (ERAM),” http://www.faa.gov/airports_airtraffic/technology/eram/ (accessed August 26, 2009).

Gong, C. and D. McNally. *A Methodology for Automated Trajectory Prediction Analysis*. American Institute of Aeronautics and Astronautics (AIAA) Guidance, Navigation and Control Conference and Exhibit, August 16–19, 2004, Providence, Rhode Island.

ISA Software, 2008. “RamsPlus.com,” http://www.ramsplus.com/Home (accessed August 26, 2009).

The MITRE Corporation 2008. “Mitre CAASD: Projects – User Request Evaluation Tool,” http://www.caasd.org/work/project_details.cfm?item_id=156 (accessed October 9, 2008).

National Aeronautics and Space Administration (NASA) 2009, NASA Ames Aviation Systems Division: ACES. http://www.aviationsystemsdivision.arc.nasa.gov/research/modeling/aces.html (accessed October 9, 2008).

Paglione, M. and R. D. Oaks 2007. *Implementation and Metrics for a Trajectory Prediction Validation Methodology*. American Institute of Aeronautics and Astronautics (AIAA) Guidance, Navigation and Control Conference and Exhibit, August 20–23, 2007, Hilton Head, South Carolina. AIAA.

UFA, 2004. “ATCoach® En Route Edition Interface Information LMTSS/UFA ERAM Interfaces,” *UFA-ICD-200401-001*, Version 2.7. Gaithersburg, MD: UFA, Inc. September 2004.

Acknowledgments

The authors would like to thank Mr. Christopher Hackett and Mr. Chu Yao for their help in the data collection and analyses for this article. We would also like to thank Mr. Robert Oaks for his skills in supporting the ACES simulations. Finally, we would like to thank Mr. Todd Lauderdale from NASA for addressing our questions about ACES, and Mr. Kenny Martin from ISA Software for his RAMS support.

ITEA Journal 2009; 30: 492–495

Operational System Testing of a Federal Aviation Administration Terminal Air Traffic Control System

Edward M. Gaguski
William J. Hughes Technical Center, Atlantic City, New Jersey

The Federal Aviation Administration (FAA) manages the largest live, 24-hour-a-day airspace system in the world using a Terminal Air Traffic Control System to accomplish the safe, orderly, and expeditious flow of air traffic at civilian and military airports throughout the United States and at military bases worldwide. The FAA executes operational testing of this complex system of systems by means of the Lifecycle Management process. This article describes the process the FAA uses to perform life cycle revisions of the live system of systems. During the in-service phase, the system experiences life cycle maintenance. The FAA assigns an Operational Test Team to perform the operational testing of life cycle revisions of the entire civilian and military air traffic control system. The Operational Test Team consists of test area teams that focus on each functional area of the system: air traffic control, technical operations, operational support, performance and capacity related testing, and configuration management. Different test environments to simulate the different environments at multiple operational sites are required. The FAA William J. Hughes Technical Center Laboratories provide the Operational Test Team with simulated environments that allow interface testing with National Airspace System (NAS) Automation Subsystems, surveillance sensors systems, and other external systems to ensure that the system can operate successfully in a system of systems. This approach promotes effective planning, reduces risks, decreases costs, and ensures that the right system is fielded.

Key words: Airspace, air traffic control, best practices, capacity testing, configuration management, FAA William J. Hughes Technical Center Laboratories, life cycle management, operational test team, performance testing, system of systems.

The Federal Aviation Administration (FAA) manages the largest and most complex airspace system in the world using the Terminal air traffic control system providing for safe and efficient movement of air carrier, civilian, and national defense aircraft at civilian and military airports throughout the United States and at military bases worldwide. The system enables air traffic controllers to visualize and identify the high volume of aircraft in their airspace and to provide safe, orderly, and expeditious separation of arriving, departing, and over-flying aircraft in high, medium, and low density airports. The software supporting this system is constantly evolving and improving by performing more tasks and reducing the amount of human verbal communication and note taking, thus allowing the air traffic controller to devote more time to his/her primary task of aircraft separation and less time toward support tasks such as communicating handoffs, and recording and communicating flight plan updates.

Terminal air traffic control system

Figure 1 depicts the Terminal systems in the National Airspace System (NAS). The Airport Tower has control over the aircraft from the time the aircraft leaves the gate until takeoff. Then the aircraft is transferred to the Terminal Radar Approach Control (TRACON) facility. The aircraft generally climbs to the ceiling of the TRACON within 20 nautical miles of the airport. Control of the aircraft is then transferred to the En Route/Oceanic Center. The cruise phase of the aircraft may be conducted with one or more En Route centers. These centers control the flight until it

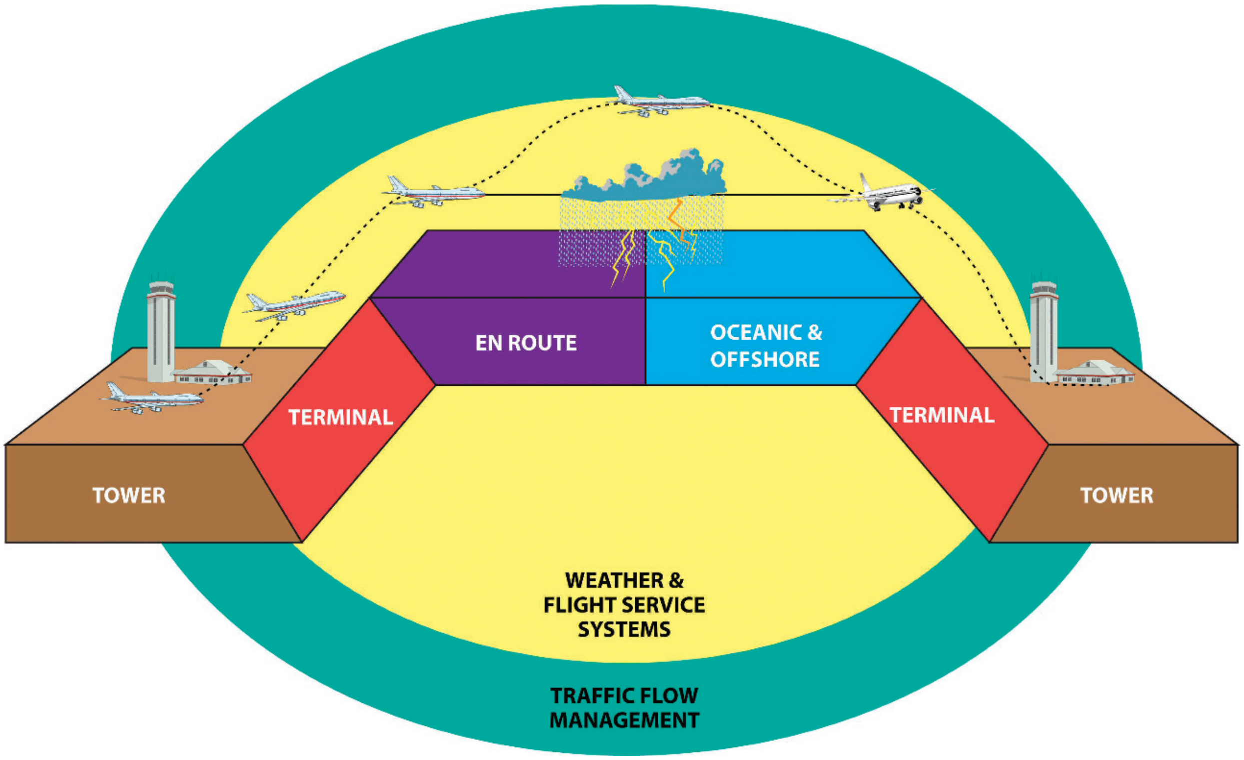


Figure 1. Terminal Systems in the National Airspace System.

reaches the arrival phase. During the arrival phase of the flight, control of the aircraft is transferred to the destination TRACON, generally at 40 nautical miles from the airport. The aircraft is finally transferred to the tower control for landing and taxi to the gate.

The Lifecycle Management process is organized into a series of phases and decision points including research and systems analysis, solution implementation, and in-service management. During the in-service phase, the system experiences life cycle maintenance. Each life cycle revision consists of a combination of engineering change proposals, program technical report solutions, and new functionality case files. Regression testing is also conducted for each life cycle revision to validate air traffic control operations and functionality, and to determine if the reliability, maintainability, availability, and supportability of the system have degraded.

Operational testing/ Lifecycle management

The FAA executes its operational testing under the Acquisition Management System policy by means of the Lifecycle Management process. Operational testing encompasses test and evaluation of a system's operational requirements. The primary objective of operational testing is to validate that a new or modified system or software build is operationally effective and suitable for use in the NAS. Operational testing consists of NAS operational integration testing, and Operational effectiveness and suitability testing. In addition, these tests verify external interface connectivity from the system under test to other FAA and/or vendor-provided systems.

The specific use of the Terminal air traffic control system in the NAS and the external interfaces to multiple systems require different test environments. These test environments are used to simulate an operational site to reach a suitability recommendation for each life cycle revision. The Terminal air traffic control system is required to receive, process, and transfer NAS information between TRACON facilities and Air Route Traffic Control Centers. The FAA William J. Hughes Technical Center Laboratories allows the team to conduct interface testing with NAS automation subsystems, surveillance sensors systems, and other external systems to ensure that the system can operate successfully in a system of systems occupying different environments. Flight data are received from and provided to external systems such as the Airport Movement Area Surveillance System, Final Monitor Aid, Precision Runway Monitor, Surface Movement Advisor, and for Noise Abatement.

The Operational Test Team

The Operational Test Team consists of test area teams that focus on each functional area of the system. These areas consist of air traffic control, technical operations, operational support, performance and capacity related testing, and configuration management. Each test area team includes a test lead and subject matter experts who contribute to creating scenarios to test the performance of the system and computer–human interface. The test area leads and subject matter experts are a mixture of FAA and contractor support personnel with experience as operational site personnel at FAA and Department of Defense (DoD) sites, engineers, computer scientists, and computer specialists. Air traffic, technical operations, and operational support personnel from FAA and DOD operational facilities are also invited to participate as test conductors and test observers.

The Operational Test Team develops a test plan and test procedures, conducts the testing, and reports test results. The operational test plan describes the planning and preparation activities required prior to the conduct of the test. It addresses the organizational roles and responsibilities, training, schedule, and staffing needs. The criteria for entry into and exit from testing and a description of the system under test as well as any test equipment or tools to be used are defined in the test plan. The requirements under test and the test procedures to be used including any risks or contingencies are also defined. The test report includes a summary of conclusions and a recommendation of the operational suitability of the life cycle revision. This recommendation is vital to the FAA decision to deploy or not to deploy the life cycle revision.

Air Traffic Control Team

The Air Traffic Control team tests essential FAA and DoD air traffic control services to ensure that the updated system performs to specifications and does not degrade current functionality. These complex functions include the conflict alert, minimum safe altitude warning, converging runway display aid, weather levels display, minimum separation indicator, and geographic plots/restriction areas. Testing is conducted through the use of live real-time surveillance sensors and other systems that make up the NAS infrastructure. Simulators are also used to allow testers and controllers to perform systems functionality testing without using live traffic and allow for repeatability.

Technical Operations Team

The Technical Operations Team utilizes essential system maintenance functions such as software and adaptation startup, system partitioning, certification, system administration, and monitor and control of the system resources to prepare the system for testing. Continuous data recording and data capture are functions that assist in isolating system failures both during testing and for field operations. The team ensures that the documentation inventory and hardware inventory are current and that procedures for removal and replacement of failed equipment are correctly documented. Interfaces for Surveillance sensors, EnRoute systems, and Terminal systems are verified, as are requirements for power shutdown and recovery.

Operational Support Team

The Operational Support Team tests essential operational support and data base management software functions for creating system builds and site adaptations. These functions include code version control and code verification for each build. The team performs the integration of a build containing software code along with site adaptation. The system adaptation and site databases must be generated and verified prior to testing.

Performance testing

Performance testing provides a measurement for comparing the results of one revision to another as a system is developed and improved. A means of assuring suitable system performance during deployment of a life cycle revision or hardware/technology refresh update is necessary to ensure that the performance of the system has not degraded. Simulators and scenarios are best utilized for such tests to acquire results that are repeatable and comparable between versions. Methods used to determine performance measurements include complex measurement equipment and/or video and playback techniques.

Capacity testing

Capacity testing assures that the system response to selected commands is processed in a suitable time to allow continued utilization by users. The system load is gradually increased until a workload level is achieved. A maintained workload level allows for collection of central processing unit and memory utilization data. As commands are executed, the system operational suitability, reliability, and availability are observed while the system load approaches capacity.

Configuration management

Configuration management is an integrated part of life cycle management applicable to hardware, software, firmware, and related technical data. Configurations must be validated by the test team to insure a stable, deliverable life cycle revision suitable for national release. The Configuration Management Team has the responsibility for managing, monitoring, and verifying changes to the baseline for the system.

Summary

The Operational Test Team continually strives for process improvement by applying effective verification and validation principles coupled with best practices to improve the quality of test and evaluation products and services. This approach promotes effective planning, reduces risks, decreases costs, and ensures that the right system is fielded. ❑

ED GAGUSKI is the operational test lead for the Standard Terminal Automation Replacement System (STARS) air traffic control system. He is assigned to the Terminal Automation Team within the Test and Evaluation Services Group at the Federal Aviation Administration's (FAA) William J. Hughes Technical Center. He is responsible for directing a team of FAA and contractor support personnel with experience as air traffic controllers, airway facility technicians, engineers, computer scientists, and computer specialists. The objective of the testing is to determine the operational suitability of the system, which is vital to the FAA decision to deploy or not to deploy the system to field sites. He has over 27 years of technical experience in software development and testing of Department of Defense (DoD) and FAA computer systems. He holds a master's degree in computer science from the University of Dayton, Dayton, Ohio, and a bachelor's degree in computer science from Jersey City State College, Jersey City, New Jersey.

Prior to being hired by the FAA as a computer scientist, Mr. Gaguski worked for the TRW Systems Integration Group as a senior systems engineer where he supported the FAA's Terminal Branch in the acquisition of terminal automation systems. Before his work with the FAA he supported various DoD programs as an employee for Logicon Incorporated, TRW Dayton Engineering Laboratory, and General Electric Aerospace. He also worked as a cooperative education intern with the U.S. Army Communications-Electronics Command, Software Support and Development Center at Fort Monmouth, New Jersey. His job titles include computer operator, research assistant, programmer, programmer/analyst, and software engineer. E-mail: Ed.Gaguski@FAA.gov

References

U.S. Federal Aviation Administration Acquisition Management System. 2006. *FAA acquisition system toolset, test and evaluation process guidelines*, Washington, D.C.: U.S. Federal Aviation Administration.

U.S. Federal Aviation Administration. 2004. *FAA test and evaluation gold standard and implementation guide,* Version 2.0, Washington, D.C.: U.S. Federal Aviation Administration.

U.S. Office of Management and Budget. 2007. *Acquisition management system: implementation strategy and planning*, Washington, D.C.: U.S. Office of Management and Budget (OMB).

Acknowledgments

The author would like to acknowledge the individual contributions of Fabian Avilla, David Cognata, John Fuller, Bill Griffith, Raymond Haines, Joan Hanson, Mike Headley, Grace Hwang, John Lawson, Kathy Moore, Alan Manalang, Dennis Sinko, Jackie Sirolli, Srivaths Sridharan, Joe Stasiowski, Madurai Vaidyanathan, and Lee Wong as members of the STARS Operational Test Team and for their dedication and professionalism.

2010 ITEA Journal *Themes*

The ITEA Publications Committee has established themes for the 2010 issues of *The ITEA Journal* and invites articles in the following areas:

The Role of T&E in Systems Engineering (March issue). Systems engineering is the engineering of complex systems and is intrinsically multi-disciplinary just as test and evaluation (T&E) are. Systems engineers, with their broad view of a program, are in a unique position to diagnose problems in the event of system failure. In T&E the operational requirements of a system must be decomposed to technical requirements and a strategy developed for measuring parameters that can lead back to assessment of mission performance. Systems engineering supplies the process and tools and along with integrated testing is the foundation for future T&E. Integrated testing is the collaborative planning and collaborative execution of test phases and events to provide shared data in support of independent analysis, evaluation, and reporting by all stakeholders. The purpose of integrated testing is to identify system deficiencies early, comply with accelerated schedules, and reduce cost. Design of experiments enables an efficient test design considering all key factors and conditions affecting performance. This issue examines all aspects of systems engineering as well as design of experiments; T&E workforce and training; verification, validation, and accreditation; standards, metrics, data, analysis, and more. ***(Manuscript deadline: December 1, 2009)***

User-Centric Systems (June issue). Systems of systems and network-centric systems are viewed as force multipliers deriving from the mutual connectedness of the elements and the perceived value of timely, critical information. Yet the right information provided to the right person at the right time does not guarantee success. The human is the key and in emerging complex systems of systems the human is more than an operator, and instead is part of the system, a node in the network. Testing a system includes objectively testing the user and requires characterizing human performance. Systems can become so complex that training the user to operate them is no longer possible; rather the systems must be designed with human limitations in mind. This issue examines cognitive performance and measures in addition to traditional form, fit, and function; instrumentation; personal protective equipment; human-machine integration; and situation awareness. ***(Manuscript deadline: March 1, 2010)***

Simulation – Where is T&E Today? (September issue). Simulation is not new and is known by terms such as modeling & simulation and live-virtual-constructive simulation. In one form or another it has been around T&E for more than 20 years and spawned simulation-based acquisition in the Department of Defense, simulation-based design in industry, and a host of other initiatives and hopes. Yet the predictions and expectations have not been realized and the capabilities are often oversold. Where is simulation today in the business of T&E? What has prevented realizing the full power of simulation, what needs to change, or have we arrived already? This issue looks at simulation past, present, and future in T&E and addresses technology, policy, history, success stories, and lessons learned; as well as simulation experience in operational testing, training, design, and other applications. ***(Manuscript deadline: June 1, 2010)***

Cyberspace Test and Evaluation (December issue). Cyberspace is the fifth combat domain, beyond air, land, sea, and space and is the realm of computers, networks, and software. The terrain of cyberspace is not physical but is virtual and ever in flux as network topology and system connectivity dynamically change. In the Department of Defense (DoD) the importance was recognized by creation of cyber commands. Beyond the DoD, cyberspace encompasses commercial networks, the communications industry, power distribution, commerce, transportation, and nearly everything that touches our lives and business today. Systems and networks are subject to continual attacks including spam, phishing, viruses, Trojan horses, worms, root kits, spyware and other malware, and distributed denial of service. Cyber-crime is multi-jurisdictional and spam is being replaced by scam. This issue looks at cyber-infrastructure, data-driven security, information assurance, information operations, electronic warfare, network electronic attack, and other cyber-threats and defenses. ***(Manuscript deadline: September 1, 2010)***

In addition: T&E articles of general interest to ITEA members and *ITEA Journal* readers are always welcome. Each Issue includes specialty features, each 2-3 pages long: **"Featured Capability"** describes unique, innovative capabilities and demonstrates how they support T&E; **"Historical Perspectives"** recall how T&E was performed in the past, or a significant test or achievement, often based on personal participation in the "old days" of T&E.; **"TechNotes"** discusses innovative technology that has potential payoff in T&E applications or could have an impact on how T&E is conducted in the future. **Interested authors**: should submit contributions to the **ITEA Publications Committee Chairman (itea@itea.org, attn.: Dr. J. Michael Barton)**. Detailed Manuscript Guidelines can be found at **www.itea.org** under the ITEA Publications tab.

ITEA Journal 2009; 30: 497–503

A Study on the Sound Source Positioning of an Initial Launching Missile in Noise Environment

Woong Park, Jae-Hyoun Ha, and Yong-Jae Lee, Ph.D.
Agency for Defense Development (ADD), Tae-an, Korea

This article proposes a basic conceptual design for the sound source positioning of an initial launching missile in a noise environment. Most of the sound positioning algorithms were studied from a stationary sound source in an ideal environment, such as an anechoic chamber, but the proposed algorithm is applied to a moving sound source in a noise environment using a Time Difference Of Arrival (TDOA) estimation. The TDOA estimation is based on a normalized cross-correlation algorithm in the time domain that doesn't need to be de-Dopplerized. The stationary sound source position is computed by TDOA and converted into a moving sound source position. Experiments have verified that the acquired sound source position is valid and can be applied to a missile flight test.

Key words: Missile flight test; sound source positioning; time difference of arrival.

In a missile flight test, the trajectory data, which are obtained through radar measurement, are critical for determining the success of a test, along with the telemetry data, which provide data onboard the missile. Trajectory data can be classified into three stages: initial, middle, and final trajectory data. However, it is reported that when a Line Of Sight (LOS) is not maintained or a missile approaches closely to the Earth's surface, the tracking performance of radar is remarkably reduced due to the scattering of radio waves. Radio waves can also be blocked by the nozzle plume from combustion products of metalized propellants. This phenomenon is prominent during the initial phase of a missile launch. During the launch, radar tracking is difficult and is currently supplemented with high-speed cameras to obtain trajectory data. However, high-speed image data cannot be displayed in real time, and quantitative trajectory interpretation from high-speed image data is difficult during missile flyout.

Accordingly, this study proposes an initial trajectory measurement method using sound source positioning to alleviate the problems of the existing initial trajectory measurement conducted by radars and high-speed cameras. In contrast with the current sound source positioning algorithm that tracks the position of a stationary sound source in an ideal environment, the proposed method tracks the position of a high-speed moving sound source from a missile in a noise environment, and is designed so that Doppler effect influence can be removed by mirroring the calculated value of the Time Difference Of Arrival (TDOA), which occurs in the process of sound arrival to the microphone using normalized cross-correlation in-time domain, and the diverging TDOA signal can be removed through a preprocessing filter. The proposed method is also designed so that the real position of a sound source from a missile can be tracked by calculating the position of the sound source using a TDOA signal and estimating the missile moving time. The method was verified by applying it to a flight test of a missile from which the initial trajectory data had been already secured through radar tracking, and comparing with the radar tracking trajectory data.

Method of sound source positioning

Research conducted thus far is mostly focused on stationary sound source positioning in an ideal environment such as an anechoic chamber. However, in order to track the sound source of a missile, which moves with high speed in the open air, the trajectory should be interpreted under the concept of a free field, and the frequency shift of the sound source according to the Doppler effect and the ambient noise, which occurs in a free field, should be considered.

Considering these problems, the sound source from a missile is assumed as a point source, and high-speed sound source positioning was conducted under the condition that the sound from a missile should be

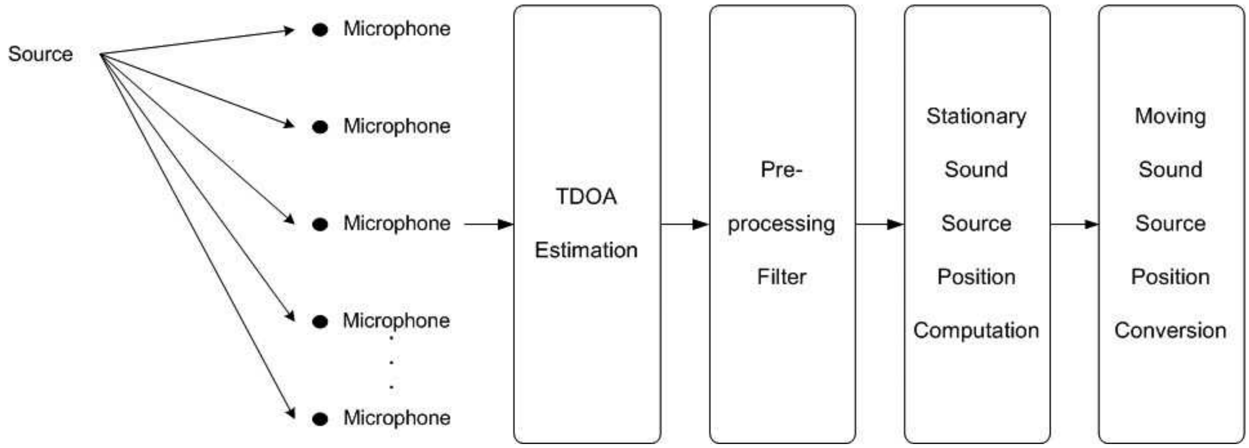


Figure 1. Sound source positioning process.

briefly perceived as a plane wave in the free field within 1 to 5 seconds after launching as prerequisite.

Figure 1 shows the process of sound source tracking as the TDOA is estimated from a sound wave reaching to an array of microphones. The noise is removed by the preprocessing filter, and the position of the sound source is calculated by geometric analysis based on plane wave theory subsequently. The calculated sound source is converted to a moving sound source position as the calculated sound source position is a stationary one.

Sound source positioning methods can be divided into two different categories of the method for finding a sound source at a far distance in a short time for a military objective, the acoustic holography method and acoustic intensity method for obtaining more detailed and specific information on the amplitude, the phase, and the propagation of a sound source rather than for tracking the moving position. In this work, the moving sound source position is the goal. Below the methods to estimate the time delay and incidence angle from a sound wave reaching the microphones are compared with each other.

Beam forming method

This is a method in which the incidence angle of a sound source is estimated by finding the largest sound power among the waves that reach to the multiple array microphones with different time delays by assigning weight values on the respective sound waves and adding the weight values to the power values of sound waves. However, this method can be problematic in that designing satisfactory beam patterns are difficult as the frequency range of the sound waves cover a wide band, and the Doppler effect according to the frequency shift of sound waves must also be considered.

Cross-power Spectrum Phase (CSP) method

In the case of the CSP method, the time delay is estimated through the correlation in the frequency domain as shown in Equation 1. Though this method is widely used in frequency analysis based on Fast Fourier Transform (FFT) as CSP can be expressed as a delta function, which has a time delay D as shown in Equation 2, and the calculation quantity for this method is relatively small, it is known that measuring time delay by this method in the free field presents problems when applied to the real signal with which noise is mixed, observing linear phase correlation between two signals is difficult. This method also has the problematic point that Doppler effect, which occurs in the frequency domain should be considered.

$$S_{ij}(f) = \int_{-\infty}^{\infty} R_{ij}(\tau) e^{-j2\pi f \tau} d\tau \tag{1}$$

$$R_{ij}^{(g)}(\tau) = \delta(\tau - D) \tag{2}$$

Normalized Cross-Correlation (NCC) method

The NCC method estimates TDOA in the time domain using the correlation between signals in which the time delay τ for measured signals can be figured out by Equation 3, and this delay time τ can be expressed as shown in Equation 4 by normalizing. As NCC is determined in the time domain, its value can be used independently from the Doppler effect; however, high-speed data sampling is necessary for calculating the exact TDOA.

$$R_{ij}(\tau) = \int_{-\infty}^{\infty} P_i(t) P_j(t-\tau) dt \tag{3}$$

$$R_{ij}^{(n)}(\tau) = \frac{R_{ij}(\tau)}{\sqrt{R_{ii}(\tau) R_{jj}(\tau)}} \tag{4}$$

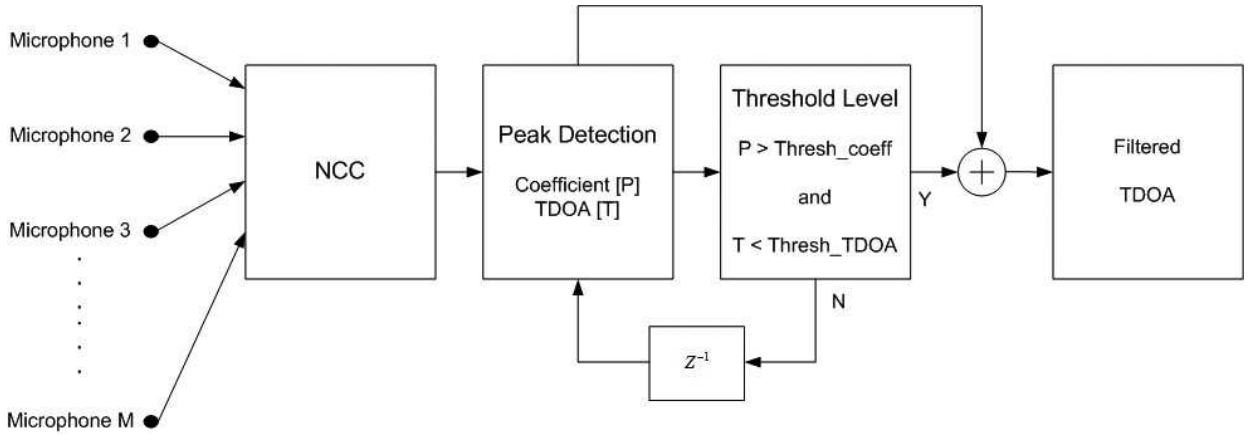


Figure 2. Preprocessing filter process.

Considering the advantage in this study, TDOA is estimated by the NCC method, which can be used independently of the Doppler effect and can be applied even in a noise environment.

Estimation of TDOA

In order to obtain the exact TDOA from the sound wave of a missile that is moving at a high rate of speed and covering a wide frequency band using the NCC method, high-speed data sampling is necessary. In this study, high-speed sampling was conducted at 1 MHz (16 bit for respective microphones), and TDOAs for individual blocks were calculated using a sliding window. The sliding window used was a 10 percent Tukey window with a size of 50-*ms*, and TDOAs were calculated moving at 10-*ms* intervals.

Preprocessing filter

The expected incidence angle range of a missile flight can be computed from the incidence angle range of sound waves that reach to the microphones, and the calculated result of this can be expressed as a TDOA range. By limiting this expected TDOA range and the correlation coefficient range, the needless noise of each window block can be removed. *Figure 2* shows the preprocessing filter, which is designed so that it can conduct NCC from sound waves arriving at the microphones, calculate the TDOA and correlation coefficient, compare these calculated values with preset range, and transfer the TDOA value only when the input data value is within the preset range to the next processing phase, otherwise refer to the previous value. The results before and after executing preprocessing filter are shown in *Figure 3*.

Computation of stationary sound source position

As seen in *Figure 4*, the position of a sound source can be expressed in polar coordinates as a combination where the azimuth angle between the reference

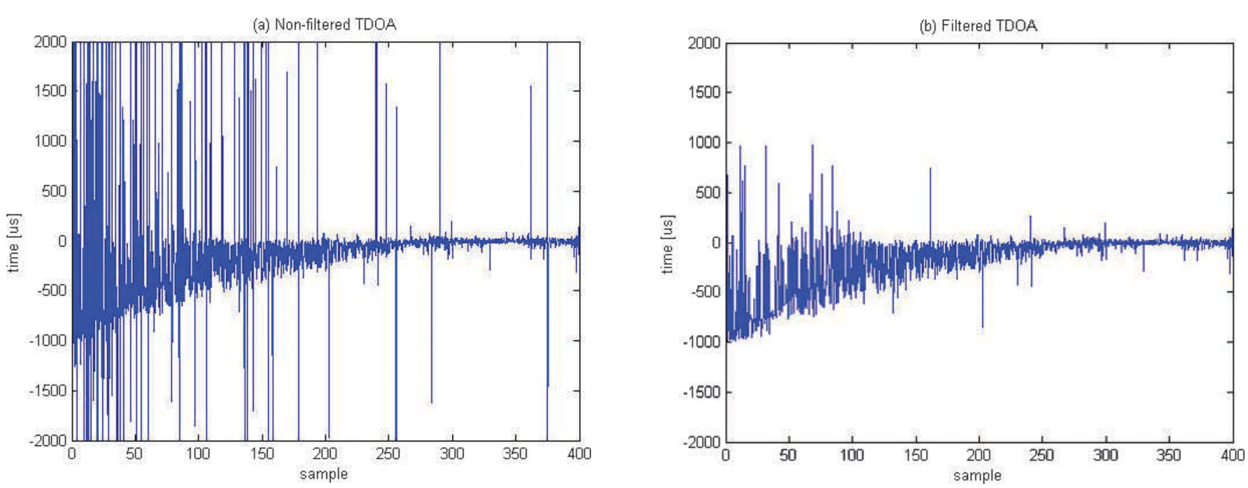


Figure 3. Nonfiltered Time Difference Of Arrival (TDOA) (left panel); filtered TDOA (right panel).

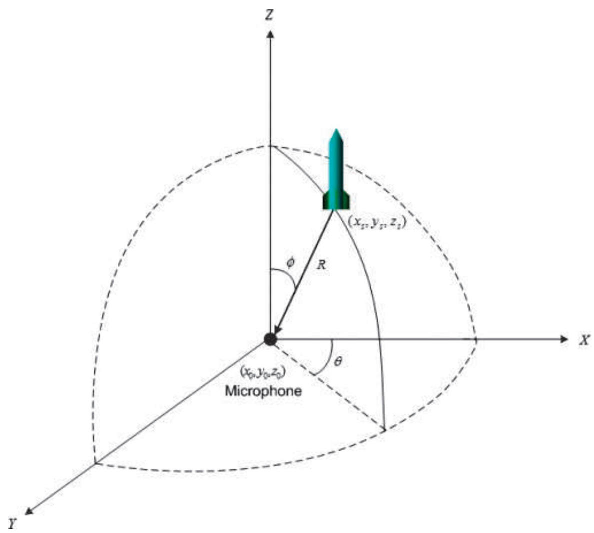


Figure 4. Stationary sound source position.

microphone and sound source is expressed as θ, and the elevation angle as φ, defining the position of a sound source, which is remote from the reference microphone position (x_O, y_O, z_O) with distance R as (x_s, y_s, z_s). Under the assumption that the sound wave generated from a missile is a plane wave, the stationary sound source can be tracked by geometric analysis as in *Figure 5*. In this study, the position in two divisions of the *XY* plane wave and the *YZ* plane wave of two-dimensional coordinates was calculated first, and next these two values were combined and converted into three-dimensional coordinates to make the calculation brief.

Figure 5 shows the microphones, which are remote from the reference microphone with distance d_i and d_j each. Here, R_i, R, and R_j, respectively, indicate the direct paths through which the sound waves reached the microphones from the sound source. The TDOAs can be calculated as below using the path difference and sound speed value, c.

$$\tau_i = \frac{R_i - R}{c} \tag{5}$$

$$\tau_j = \frac{R - R_j}{c} \tag{6}$$

When complete measured values of TDOAs (τ_i, τ_j) can be obtained, the relation between the values τ_i, τ_j, which are figured out by cosine formula, the distance to the sound source R and the azimuth angle θ can be proposed as below.

$$\tau_i = \frac{-R + \sqrt{(R^2 + d_i^2) - 2Rd_i \sin\theta}}{c} \tag{7}$$

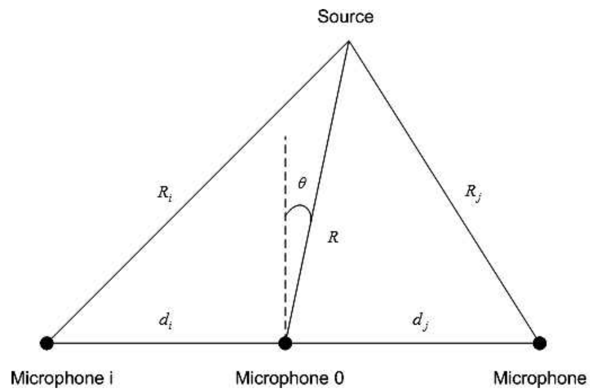


Figure 5. Geometrical structure between microphones and sound source.

$$\tau_j = \frac{R - \sqrt{(R^2 + d_j^2) - 2Rd_j \sin\theta}}{c} \tag{8}$$

Equations 7 and 8 are nonlinear function s of R and θ. These formulas can be substituted into Equations 9 and 10 by approximating using Taylor series.

$$R = \frac{d_i d_j (d_i + d_j)\cos^2\theta}{2c(d_j \tau_i - d_i \tau_j)} \tag{9}$$

$$\theta = \sin^{-1}\left[\frac{c(d_j^2 \tau_i + d_i^2 \tau_j)}{-d_i d_j (d_i + d_j)}\right] \tag{10}$$

In the same way, by estimating and combining the values of R, that is, the distance in the *YZ* plane and the elevation angle φ, these values can be converted into three-dimensional coordinates. Here, considering the value of R, the error in τ_i and τ_j in the denominator occurs more sensitively, compared with the case of incidence angle. In this study, only the incidence angle value is directly calculated and opted, and the distance data were opted from Doppler radar data.

Conversion of moving sound source position

The sound source generated from a missile is not stationary in reality but moves at high speed. Though the speed of a missile just after launch is subsonic, it exceeds the sound speed only 2 to 3 seconds afterward. Accordingly, the measured time by sound source position after missile launch should be converted into the real time when the sound source appeared. Figure 6 shows the transfer process of a sound wave that reaches the microphones after launch. T, the time that a sound wave reaches the microphones after the movement of a missile, can be expressed as the sum of t, the time that the missile moves over distance R, and t, the time that a sound wave reaches to the microphones.

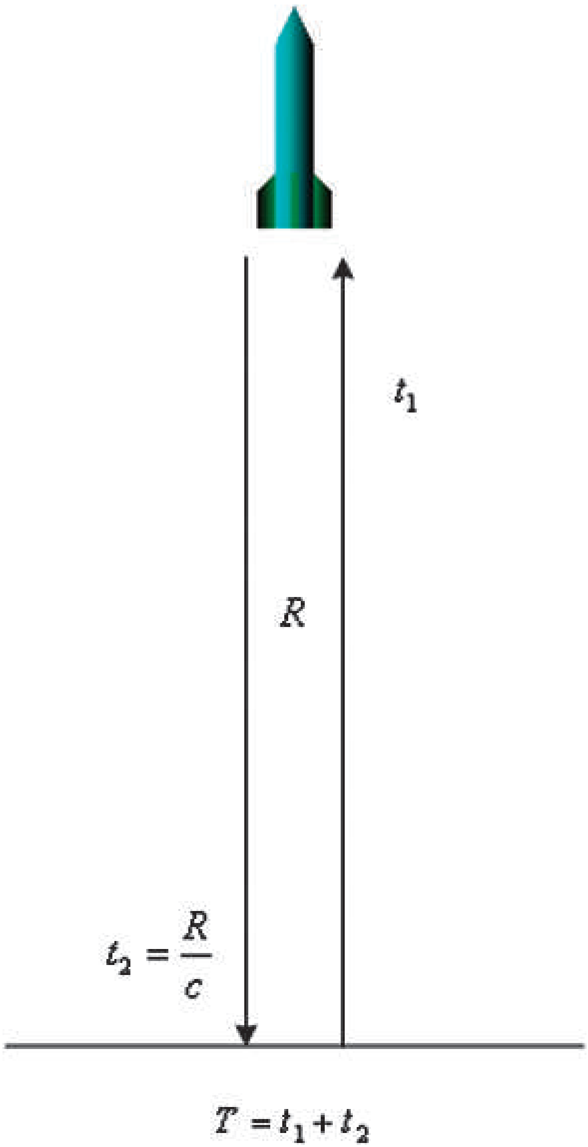


Figure 6. Propagation process of sound wave to a microphone.

Here, just as the total spent time T can be measured from the time recorded by being synchronized with missile launching, and t_2, the time that a sound wave reached to the microphones, can be grasped from the relation between missile movement distance R and sound speed c, the missile movement time t_1 can also be calculated. In this case, it is assumed that the sound speed is fixed. Accordingly, the sound source position of the movement of a missile in reality can be expressed by both the missile movement time t_1 and the corresponding sound source position.

$$T = t_1 + t_2 \tag{11}$$

$$t_2 = \frac{R}{c} \tag{12}$$

$$t_1 = T - \frac{R}{c} \tag{13}$$

Experiment

It remains to apply the proposed method to a missile flight test for which the initial trajectory value is not known, to examine the validity of the method in principle. However, to verify the result, the method was applied to an existing test in which an initial trajectory value had already been determined by radar, and the two result values were compared with each other. The experimental configuration is shown in *Figure 7*. The B&K, Ltd., 4958 microphone array was arranged in nine spots in center, azimuth, and elevation directions by 0.5-m intervals, and the sound source was measured at 16-bit resolution and 1-MHz sampling rate by respective channels through anti-aliasing filters.

The total measurement time was 5 seconds after the missile launch, and measurement starting time was based on the moment that an infrared (IR) detector sensed the flame of the missile launch for synchroni-

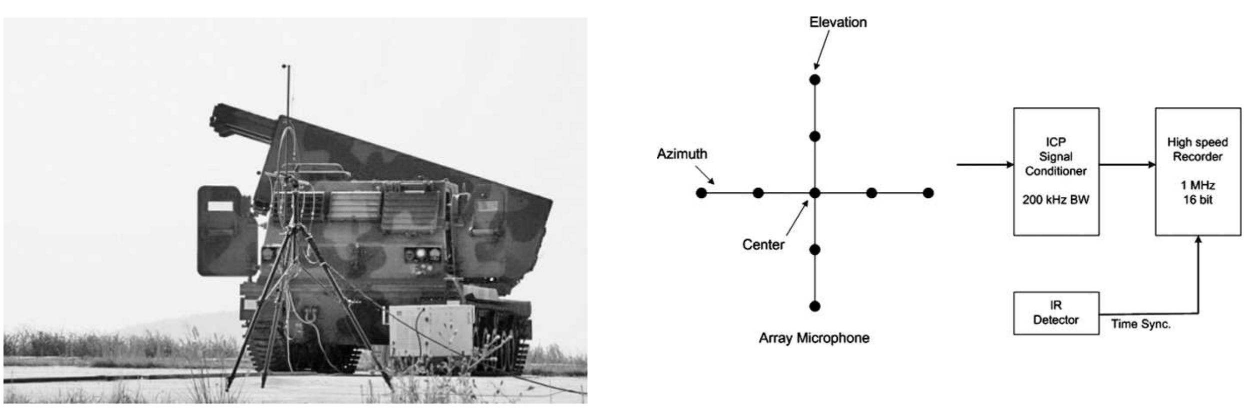


Figure 7. Experiment configuration.

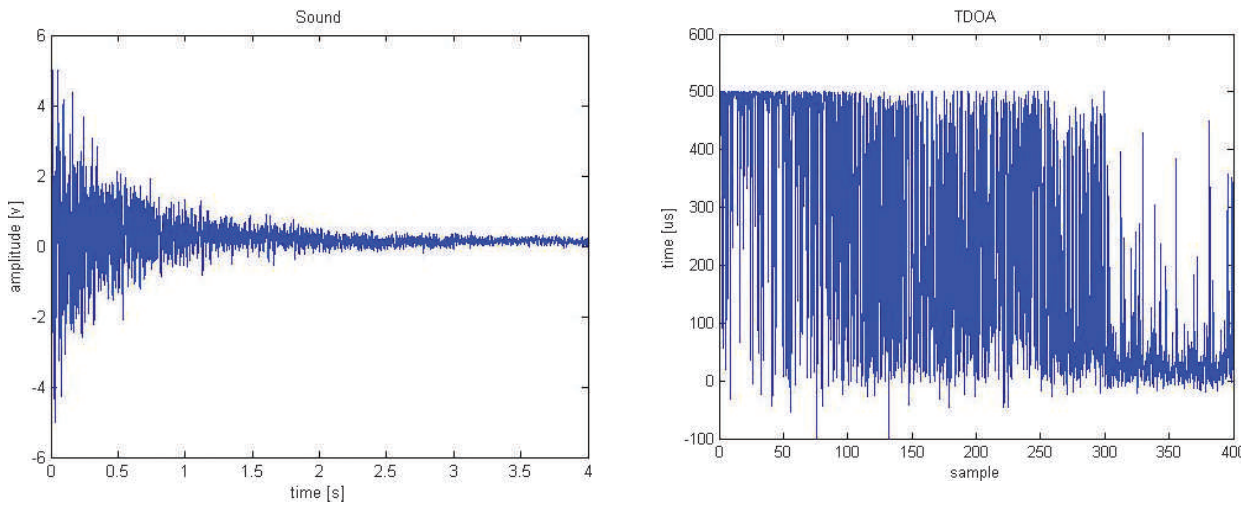


Figure 8. Sound source level (left panel); Time Difference Of Arrival simulation (right panel).

zation. The sound source emerged in a very large scale at the time of launch; however, the scale reduced drastically afterwards. For this phenomenon, if the input range of the data acquisition system is set to fit the size of the sound source at launch time, the resolution becomes smaller afterward; hence, fine resolution data are unavailable. Accordingly, in this experiment, the input range of the data acquisition system was set and sized for the sound source at 1 second after launch; the sound source position tracking was also conducted in the block of 1 to 5 seconds after launch in the actual data processing

TDOA simulation

In order to verify the TDOA estimation algorithm prior to its use, checking NCC between the sliding windows of the originally received sound source and the one made by 500-µs delay from this original sound source was conducted. The result is illustrated in *Figure 8*. Here, TDOA appeared appropriately to be the peak value of NCC values in the experiment. Afterward, it was discovered that the correlation values fell drastically, and the TDOA appeared to be inexact as the sound wave was processed longer. The reason, it is presumed, is that the received sound source scale fell drastically according to time passage; hence, the signal-to-noise ratio (SNR) also decreased. In this experiment, the TDOA was estimated by peak detection after the calculation of NCC from the signals received at the microphones of heading center microphone and azimuth, elevation directed microphones, and subsequent adjustment by the preprocessing filter.

The actual sound speed was applied as c = 339.78 *m/s*, which was determined by Equation 14 at the experiment temperature $t = 13.8^\circ$.

$$c = 331.5 + 0.6t \tag{14}$$

Since the tracked sound source position should be converted to the coordinates of the sound source position corresponding to missile movement, the moving sound source position coordinates were figured out using actual missile movement time. *Figure 9* illustrates the converted sound source coordinates compared with the radar coordinates, in which the

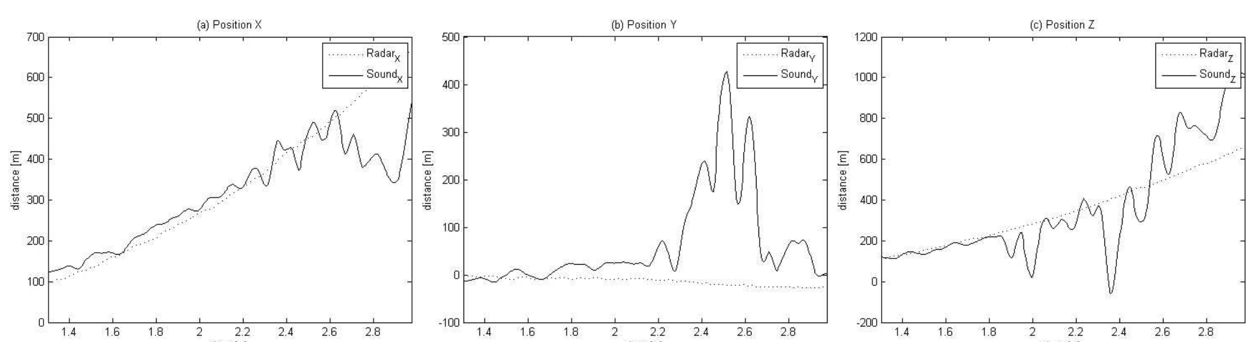


Figure 9. Result of moving sound source positioning: X axis (left panel), Y axis (center panel), Z axis (right panel).

resultant sound source position appears to be similar to the coordinates measured by the radar for 1.3 to 2.3 seconds after missile launch, which is the time block before the sound wave correlation decreases. Specifically, it is found that the tracking result for the *X* direction is better than for the *Y*, *Z* axis, which is presumed to be due to the phenomenon that reflection effect is less manifested on the *X* axis (the azimuth angle) compared with the *Y* axis (the elevation angle).

Conclusion

This study is a basic conceptual design for sound source positioning of a high-speed moving missile using microphones in a noise environment, to supplement the measurement provided by radars and high-speed camera images in a flight test of a missile and to verify the validity of the technology application through a simple model. We plan to develop more credible technology models, to investigate effective noise removal algorithms and algorithms for data fusion of sound source position of the sounds which are measured from multiple places in order to enhance the exactness of measurement in the future. ❑

WOONG PARK is senior researcher at the Agency for Defense Development in Korea. He earned a master of science degree in electronic engineering from Soongsil University and has been working for test and evaluation of guided missile systems at Defense Systems Test Center (DSTC) since 2001. His research is focused on the technique of moving sound source positioning. Recent work includes the sound visualization of moving sound source. E-mail: parkwoong@add.re.kr

JAE-HYUN HA is a researcher at the Agency for Defense Development in Korea. He earned a bachelor of science degree in electronic engineering from Kumoh National Institute of Technology. He was employed as an engineer at Samsung-Thales from 2002 to 2004. He has been working for test and evaluation of guided missile systems at Defense Systems Test Center (DSTC) since 2005. His research interests include real-time data acquisition and visualization. E-mail: hahyoun@add.re.kr

YONG-JAE LEE is a principal researcher and works at the Agency for Defense Development in Korea as a project team leader. His research interests include the data fusion algorithms of target tracking and modeling and simulation for flight test control. He earned a master of science degree and a doctor of philosophy degree in electronic engineering, all from Ajou University. E-mail: yjlee@add.re.kr

References

Bau-Yen Tsui, James. 2004. *Fundamental of global positioning system receivers: A software approach*. Hoboken, NJ: John Wiley & Sons.

Bendat, Julius S., and Allan G. Pierson. 2000. *Random data analysis and measurement procedure*. Hoboken, NJ: John Wiley & Sons.

Benesty, Jacob. Time-delay estimation via linear interpolation and cross correlation. 2004. *IEEE Transactions on Speech and Audio Processing* 12 (5): 509–519.

Carter, G. C., A. H. Nuttal, and P. G. Cable. 1973. The smoothed coherence transform. *IEEE* 6: 1497–1498.

Foy, Wade H. 1976. Position-location solutions by Taylor-series estimation. *IEEE Transactions on Aerospace and Electronic Systems* AES-12 (2): 187–193.

Friendlander, Benjamin. 1987. A passive localization algorithm and its accuracy analysis. *IEEE Journal of Ocean Engineering* OE-12 (1): 234–245.

Knapp, C. H., and G. C. Carter. 1976. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing* ASSP-24: 320–327.

Omologo, Maurizio, and Piergiorgio Svaizer. 1997. Use of the crosspower-spectrum phase in acoustic event location. *IEEE Transactions on Speech Audio Processing* 5 (3): 288–292.

ITEA Journal 2009; 30: 504–512

Methods for Peer Behavior Monitoring Among Unmanned Autonomous Systems

Rick Dove

Kennen Technologies, Questa, New Mexico

As multi-agent weapon systems manifest more autonomous operation and work in teams, unpredictable emergent behaviors will occur in both individual agents and in teams. These group behaviors will add new dimensions to testing, posing potentially explosive centralized monitoring and evaluation tasks with large groups. The impossibility of specifying and testing all potential situational conditions is recognized, and the safety of the testing environment itself is of concern. A prior companion article explored a socially attentive foundation for peer behavior monitoring among Unmanned Autonomous Systems (UAS) under both test and mission conditions. This article explores certain complexity issues of aberrant behavior detection in UAS, draws parallels with human cognition capabilities, and provides a technology foundation for massively parallel behavior-pattern detection.

Key words: Aberrant behavior; control penetration; expertise; parallel recognition capabilities; pattern recognition; peer-peer socially attentive monitoring; unmanned autonomous systems; very-large-scale integrated (VLSI) circuit.

The trend toward increased autonomy in unmanned weapon systems has raised concerns about methods for testing these devices both individually and in tactical group maneuvers (DoD 2008). Increased autonomy is generally enabled and permitted by increased intelligence of the artificial kind in Unmanned Autonomous Systems (UAS). Intelligent systems, be they human or artificial, exhibit behaviors in response to situational conditions. Situational conditions are unpredictable and infinite in potential variety, leading to emergent behaviors at both the individual and group level. For UAS in warfighting, emergent behavior is necessary and desirable when it is appropriate and useful, and potentially a major problem when inappropriate.

Range testing can never duplicate the situational variety that will arise in warfighting, any more than prequalifying the capabilities and performance of Michael Vick as a sports team player was able to avoid later behaviors that reflected poorly on all players by association (subsequent rehabilitation not withstanding). UAS that run amok in any way will reflect poorly on all UAS—eroding necessary public trust.

In a prior companion article (Dove 2009b), Arkin (2007) and Moshkina and Arkin (2007) were cited for identifying the important need of UAS conformance to rules of ethics, rules of war, and related high-level behaviors expected by the public of weapons toting UAS. Infractions can be devastating to continued public acceptance as well as to life and property. Range testing alone cannot assure safety under warfighting conditions. It is suggested that testing for appropriate behavior become a continuous process throughout the operational life of UAS, carried out by peer-peer monitoring.

Peer behavior monitoring occurs naturally and constantly in social animals. Each member of the group evaluates the others for adherence to social norms and threats to social coherence and security. Rogue elephants, for instance, are the result of banishment for unacceptable behavior. Social insects are known to restrain and even kill members of the group that overstep certain social bounds (Monnin et al. 2002; Heinze 2003; Flack et al. 2006; Ratnieks, Foster, and Wenseleers 2006).

Humans monitor the behavior of others in ways more sophisticated and more complex than animals of lesser cognitive capability. The process is often carried out formally as a test for granting new candidates membership in a group. Initial tests are typically for similar values, compatible behaviors, acceptable capabilities, and even for synergy in mission-based groups such as sports teams or Special Forces.

A revealing example of human peer-behavior monitoring and punishment was recently published in Myers (2008). Professor Myers studied social reactions in on line game play. He played by the rules of the game but not by the cultural rules of the dominant player group. The degree of escalating retaliation as the group turned against him is an interesting study in human peer-behavior policing.

Unlike traditional approaches at sophisticated behavior detection and classification through reasoning, this article suggests an approach inspired by human expertise studies, where it appears that a conclusion is driven by a vast quantity of simultaneously accessible "experience" patterns rather than a compromising sequential search or reasoning process. This approach is both suggested and conceptually possible with new processor architectures offering massively parallel pattern-recognition capabilities. Many of these architectures are inspired by human cortical learning and classification models but may not offer ready post-learning algorithm cloning nor behavior-capability transparency. A single-processor architecture without integrated learning that features massive parallel classification capability for explicit patterns avoids these potential limitations, and will be used to establish a conceptual foundation for peer-peer socially attentive monitoring.

Detection-complexity leverage

Progress in pattern recognition has come in the form of trying harder with more elaborate recognition algorithms, pattern-tuned special-purpose processors, multi-core processors and clustered servers, multiple graphic processors, and massively parallel supercomputers. All of these approaches continue to make tradeoffs among the same forces in tension: accuracy, time, and cost.

Biological capability is the benchmark for pattern recognition. Machines, like people, cannot recognize situations of which they have no prior knowledge. A healthy person over a lifetime builds up a wealth of experience patterns, stored in memory, adding details and variations as repeated exposure reveals new levels of nuance. How biological entities achieve this remains as conjecture, but it is clear that patterns are developed, retained, and applied in the necessary and constant sense-making of everyday life.

Klein (1998) suggests his Recognition Primed Decision model to explain how humans make decisions without apparent deliberation or reasoning. Well known for his studies of professional firefighters making appropriate choices for situation response almost immediately, he describes the Recognition Primed Decision model as one that uses intuition (pattern recognition) to qualify the first viable action, without conscious weighing and decision making.

Research indicates that human expertise (extreme domain-specific sense-making) is strongly related to meaningful pattern quantity. According to an interview with Nobel Prize winner Herb Simon (Ross 1998: 98–104), people considered truly expert in a domain (e.g., chess masters, medical diagnosticians) are thought unable to achieve that level until they've accumulated some 200,000 to a million meaningful patterns, requiring some 20,000 hours of purposeful focused pattern development. The accuracy of their sense-making is a function of the breadth and depth of their pattern catalog. Of interest, in biological entities, the accumulation of large expert-level pattern quantities does not manifest as slower recognition time. All patterns seem to be considered simultaneously for decisive action. There is no search and evaluation activity evident.

On the contrary, automated systems, regardless of how they obtain and represent learned reference patterns, execute time-consuming sequential steps to sort through pattern libraries and perform statistical feature mathematics. This is the nature of the computing mechanisms and recognition algorithms generally employed in this service.

Ross (2006) talks about the expert mind, and Herb Simon presents a "chunking" explanation for how chess masters can manage and manipulate a vast storehouse of patterns. Ross ties this chunking discussion into the common understanding that the human mind seems limited by seven plus-or-minus two elements in working memory: "By packing hierarchies of information into chunks," Simon argued, "chess masters could get around this limitation, because by using this method, they could access five to nine chunks rather than the same number of smaller details."

Psychologist George Miller (1956) wrote "The magical number seven plus or minus two" that provided the underpinning for Simon's suggestion. Miller's article is a great and rare reading pleasure as well as a rich storehouse of information, far beyond the simple seven-digit limitation to which common reference has reduced it. Of importance, Miller's concept of chunking into hierarchical levels of patterns-of-patterns appears highly relevant in attempting to build pattern-recognition algorithms that exhibit capabilities seen in humans. Subsequent research (Cowan 2001) carries this study of chunks and limits further and makes a case for the number four plus-or-minus one as a more likely limit.

Similar chunked-hierarchy architecture is reported by researchers at MIT (Serre, Oliva, and Poggio 2007). Serre's doctoral dissertation (Serre 2006) describes

> *"a quantitative model that accounts for the circuits and computations of the feedforward path of the ventral stream of visual cortex," and claims "that this may be the first time that a*

neurobiological model faithful to the physiology and the anatomy of visual cortex . . . achieves performance close to that of humans in a categorization task involving complex natural images." (Serre 2006)

Though Serre's work is focused on image recognition, it is inspirational in its fit with the platform developed in this article and will surely guide subsequent steps in this investigation.

This section will close by noting a tie to the discussion in Dove (2009b) of multi-agent trajectory behavior recognition. Hockey legend Wayne Gretsky is renowned for his field sense (Kahn 2007)—knowing where his teammates are without looking and knowing where the puck will be next. Though what sensory mechanisms are involved may be illusive for now, this is expert pattern recognition involving the trajectories of bodies and objects in motion, rather than static chessboard configurations or medical diagnostic symptoms. Intille's American football-play identification from visual image pattern recognition did not have the vast quantity of patterns associated with expertise (Intille 1999, 2001), nor did it have to respond in real time; but Intille's work can offer initial guidance on how an artificial mechanism might represent tactical choreography patterns far in excess of a football playbook.

Technology leverage

Brain circuitry understanding and models of parallel pattern-recognition algorithms with brain-like results at MIT (Riesenhuber and Poggio 1999), at the San Diego Neuroscience Institute (McKinstry, Edelman, and Krichmar 2006), and at Numenta (George 2008) are already being fabricated as experimental Very-Large-Scale Integrated (VLSI) circuits at Stanford (Merolla and Boahen 2006) and at the Ecole Polytechnique in Lausanne (Schemmel et al. 2006). These VLSI chips combine analog and digital circuitry to emulate simple models of neuron/synapse circuitry, and integrate pattern learning with pattern detection.

The integrated nature of learning before recognition may make rapid cloning of the information in these chips difficult, and the nature of the learned patterns may be difficult to verify for behavior boundaries. For example, Jeff Krichmar, a senior fellow of the San Diego Neuroscience Institute, said in a recent interview:

"Put a couple of my robots inside a maze, let them run it a few times, and what each of those robots learns will be different. Those differences are magnified into behavior pretty quickly." (Kotler 2009)

These chips promise remarkable capabilities, but they may also raise some problems for weapons toting UAS test and verification.

On the other hand, conventional stored-program sequential-instruction processors pressed into massive pattern recognition service are severely constrained by trade-offs among speed, cost, and accuracy.

A new VLSI pattern-detection processor architecture, shown partially in *Figure 1*, does not contain integrated learning, can be unambiguously loaded with detection patterns, decouples the speed/accuracy trade-off, and renders the cost/accuracy trade-off negligible (Dove 2009a). The architecture features massively parallel, dynamically configurable Feature Cell Machines (FCMs), which simultaneously process the same data stream. Low-cost VLSI fabrication, unbounded scalability, and high-speed constant-rate throughput independent of pattern number and complexity break current trade space constraints.

This decoupling of the speed/accuracy trade-off constraints enables new possibilities for employing pattern recognition. In particular, the massive quantity of simple patterns associated with expert performance can be investigated as an alternative to time-consuming accuracy-compromising computational heuristics. In one sense it sounds like a brute force approach: enumerating all possible patterns of interest, rather than developing an elegant heuristic. On the other hand, the biological benchmark appears to use this massive-pattern-quantity approach; while "elegant" approaches are made necessary by the nature of the computational mechanisms employed—not the problem in need of a solution—and they extract a cost in both accuracy and time that can be avoided.

The processor architecture and how it eliminates these trade-offs is explained in Dove 2009a. Currently an emulator is used for investigating parallel algorithm development, with field programmable gate array (FPGA) processor prototypes employed for large data streams while VLSI design is in process.

Some understanding of the processor architecture is necessary. Referring to *Figure 1*, a partial view of the processor's architectural concept shows massively replicated detection cells. A quarter to half million such cells on a single VLSI chip appears possible for early generation silicon. These cells are independent units, with four dynamically configurable elements to consider:

1. an activation status,
2. a 256-element feature-vector designating all byte values of interest,
3. a set of pointers to other cells that will be activated if this cell is "satisfied," and

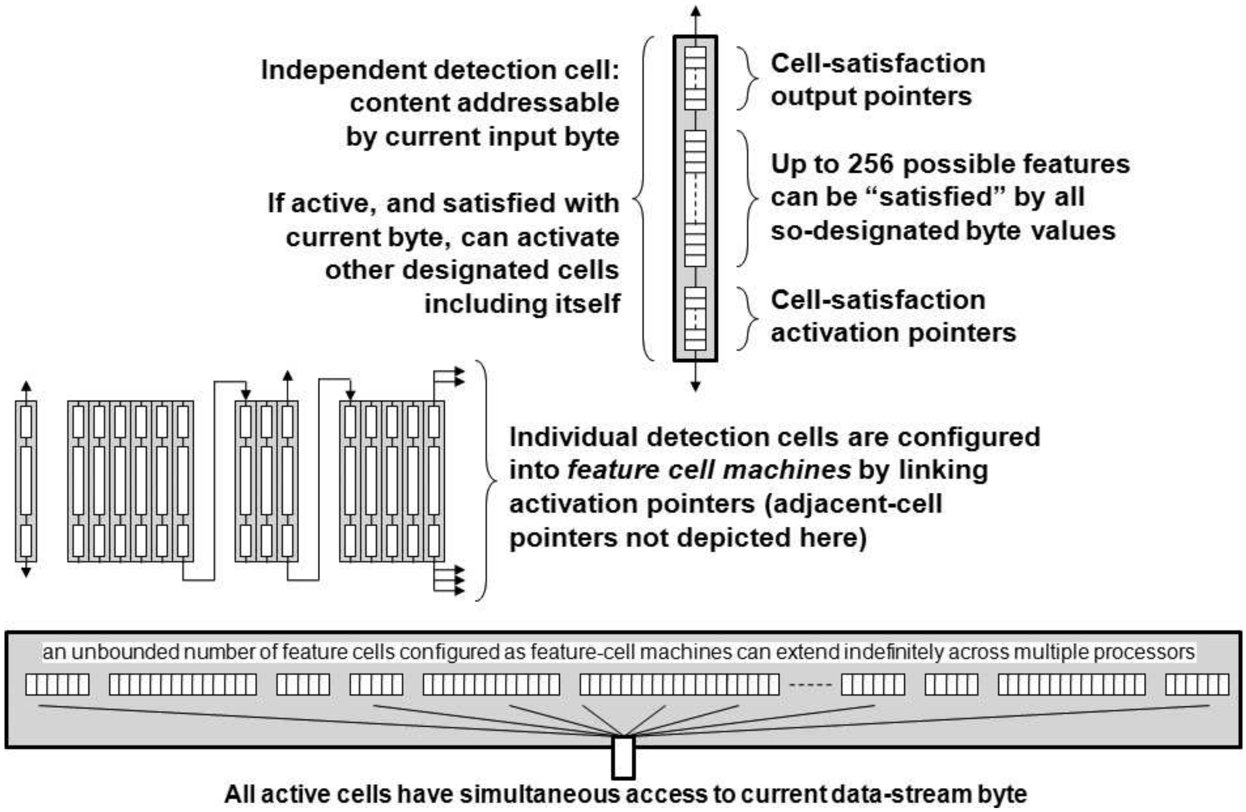


Figure 1. Configurable feature cells and feature cell machines.

4. a set of pointers to output transforms that can logically combine designated cell-satisfactions into buffered output and reset designated processor status.

In operation, an external controller feeds data stream bytes in sequence to a current-byte register in the processor. The presentation of each new byte triggers the beginning of a detection cycle. The current byte acts as an index into the feature vector for all active cells simultaneously. If a cell is active and the current byte value is one of interest, as designated in the feature vector, it is said to be "satisfied." That satisfaction will activate (for the next cycle) all other cells according to the satisfied cell's activation pointers and will cause designated output transformations to occur according to this cell's output pointers. A cell's activation pointers may include one that reactivates itself, as cell activation is effective for a single cycle only. Note that a cell can respond to any number of data-stream byte values, which enables value-based as well as syntactic feature-based classification.

Multiple processors can be employed in parallel and serial arrangements to increase throughput speed and/or reference-pattern capacity. Interleaving packet-based data streams, for instance, across multiple processors can be used to increase throughput speed. Presenting the data stream "current" character to multiple processors simultaneously can be used for unbounded reference-pattern scalability.

With this architecture, groups of detection cells can be configured into FCMs, similar to finite state machines, by setting activation pointers to pass activation successively through a group of successively "satisfied" cells. One cell may activate many other cells, so that multiple pattern branches may become simultaneously active. Any number of such FCMs may be configured within the total cells available within a processor. Typically such FCMs are created to detect (classify) specific patterns or subpatterns of interest.

The next section will provide some simple examples that could be useful in aberrant behavior detection methods for the concepts presented earlier.

Classification techniques

Pattern recognition has two distinct approaches, and a third that blends the two. For a full treatment see Jain, Duin, and Mao (2000).

- Statistical Approach—In this approach an unknown entity or situation is characterized by a set number of features and measured values for each

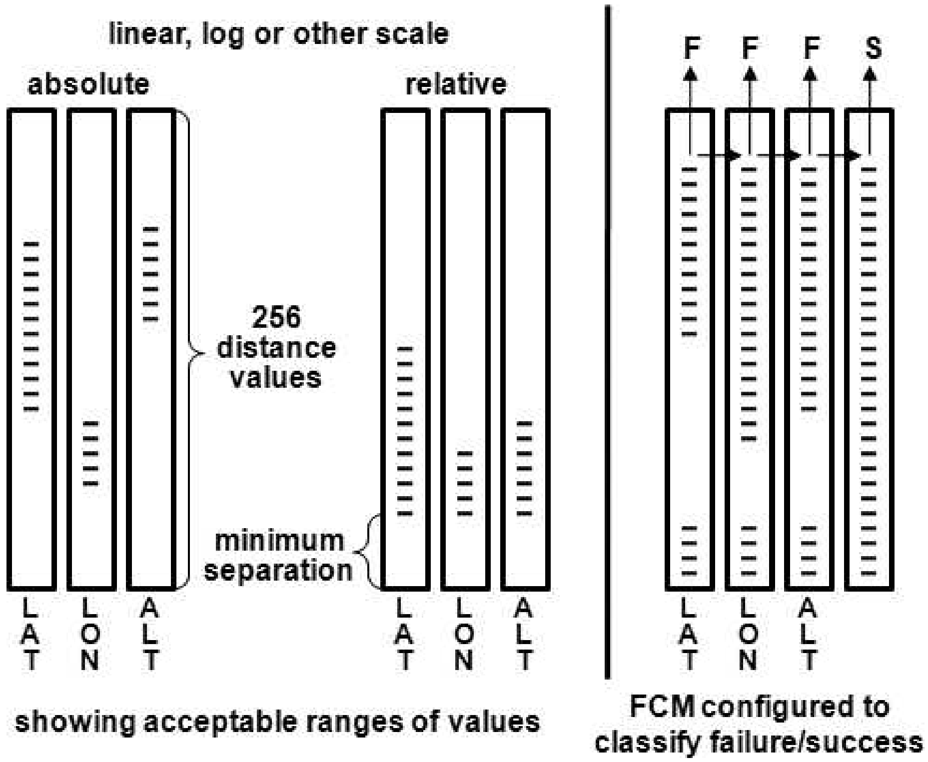


Figure 2. Some possible ways of encoding an envelope of acceptable values for latitude, longitude, and altitude. Multiple detection cells and data-stream bytes could be used for each. Minimum separation ensures two Unmanned Autonomous Systems (UAS) do not get dangerously close. A detection cell only has one satisfaction line, so the Feature Cell Machine (FCM) on the right must satisfy on failure with the complement of the permissible range. The final cell matches on anything and signals success.

feature (a person characterized by height and weight; a danger characterized by velocity, distance, and heading). Mathematically the features become dimensions in a multidimensional space, and the values for each of those features then place an unknown entity or situation at a point in that multidimensional space. Regions of the space are associated probabilistically with pattern classification (man or woman; dangerous or suspicious or benign).

- Syntactic Approach—This approach is structural in nature and generally hierarchical, where patterns are composed of subpatterns, which are in turn composed of subpatterns, with the lowest level subpatterns being simple recognizable primitives. In syntactic pattern recognition, a formal analogy can be drawn between the structure of patterns and the syntax of a language. Language parsing is common usage for this approach, but other patterns such as waveforms and multi-agent trajectory paths, which can be constructed from primitive structural components, lend themselves to syntactic recognition. Syntactic patterns are composed of primitives that follow rules about how they may be combined in relation to each other. Using the common linguistic metaphor, these rules form a grammar of allowable pattern structures.
- Augmented Grammars—This approach combines the two above, which may be done in a variety of ways to suit the raw sensor data, the computational resources being employed, the difficulty of feature extraction, and the speed vs. accuracy trade-offs dictated by the application. In a general sense, augmented grammars have syntactic elements and semantic elements, mixing structural relationships and feature values.

The processor described here is well suited to the syntactic approach, appears highly promising for augmented grammar approaches, and has utility for some statistical approaches.

A few general basic techniques will be shown to give some idea of how detection cells can be organized as FCMs, and how such FCMs can be organized to discriminate syntactic structure, feature values, and pattern groupings.

Feature value discrimination

One likely classification of undesirable behavior might be a UAS that is not where it is expected to be. Perhaps it

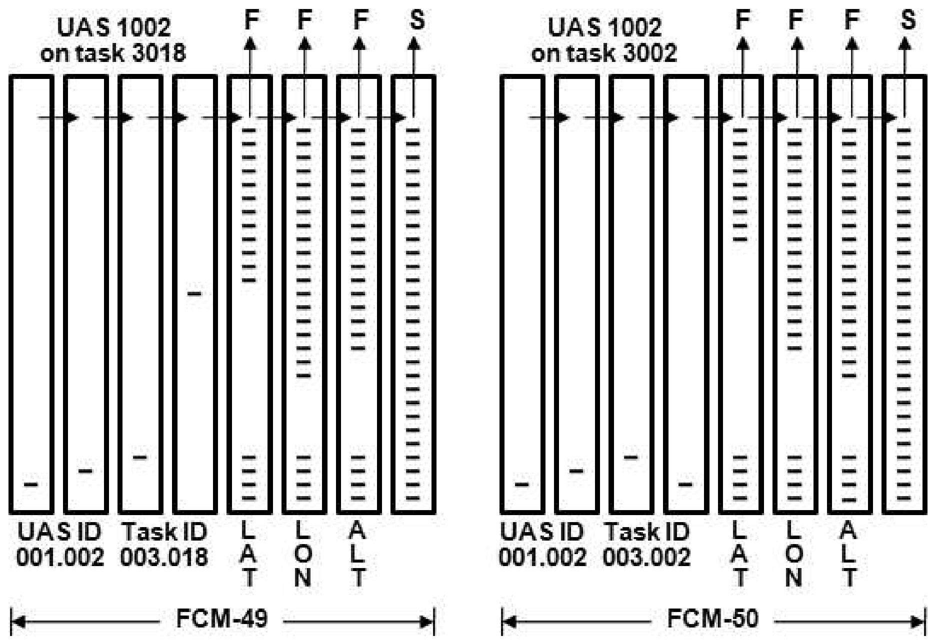


Figure 3. This example employs a packet approach to data-stream packaging. Packets here have a two-part header: the first two cells/bytes signify the Unmanned Autonomous Systems (UAS) associated with the data, and the second two signify the task that this UAS is currently supposed to be executing. Detection of the header activates the correct positioning envelop for that UAS on that task. Two Feature Cell Machines (FCMs) are shown for the same UAS on two different tasks 49 and 50.

has developed a mobility malfunction, missed a cue signaling a new task, been incapacitated by the enemy, or redirected by an unauthorized command. If a team of UAS is coordinated in accordance with a specific plan, each agent is expected to maneuver within some absolute or relative location envelope. This envelope may be narrow during travel to a target area, larger during engagement, and different among some members of the team when sub-groups are deployed on separate tasks. The example in *Figure 2* shows two ways to encode the location of a UAS by three global positioning system coordinates: latitude, longitude, and altitude. The absolute case can detect a UAS that wanders outside of its expected travel envelope for a given leg of a journey. The relative scale might be in relationship to a monitoring team-member during some phase or task of planned teamwork.

Figure 3 shows how the location envelope for two different tasks assigned to the same UAS at different times during a mission might be configured. Imagine thousands or tens of thousands of such reference patterns all prepared to classify incoming data as acceptable or not, all localized to multiple specific UAS. The example FCM could be expanded to include additional behavior data for a specific UAS in a specific task, or separate FCMs could handle separate behaviors and be associated appropriately with data-stream packet headers.

A final technique example is offered to indicate one of the ways this processor could be used to weight different features or subpatterns within a total pattern. *Figure 4* indicates some of the higher level aggregation and output capabilities of the processor conceptually (Dove 2009a). In this example, the down counters are employed to give different weights to different features of a pattern. A down counter can be initialized to some value when a configuration load is sent to the processor. Output pointers associate specific down counters with specific FCM satisfaction lines. *Figure 4* depicts four possible classifications for a large number of features, where one of those features carries a weight of 2, another carries a weight of 3, and the rest each carry a weight of 1.

The examples shown are all simplistic and not indicative of the range of possibilities. They were chosen to show some specific techniques that broaden the purely syntactic applications readily associated with this re-configurable replicated detection cell architecture.

Concluding remarks

The leverage discussions in this and the prior article (Dove 2009b) are complementary, each having a potential role in a total solution platform. The starting point in the prior article was social attentiveness in certain biological systems that monitor and enforce peer behavior. The end point was a conceptual example of technology that can approach the pattern capacities and speeds of biological systems in bounded sub-domains of behavior interest.

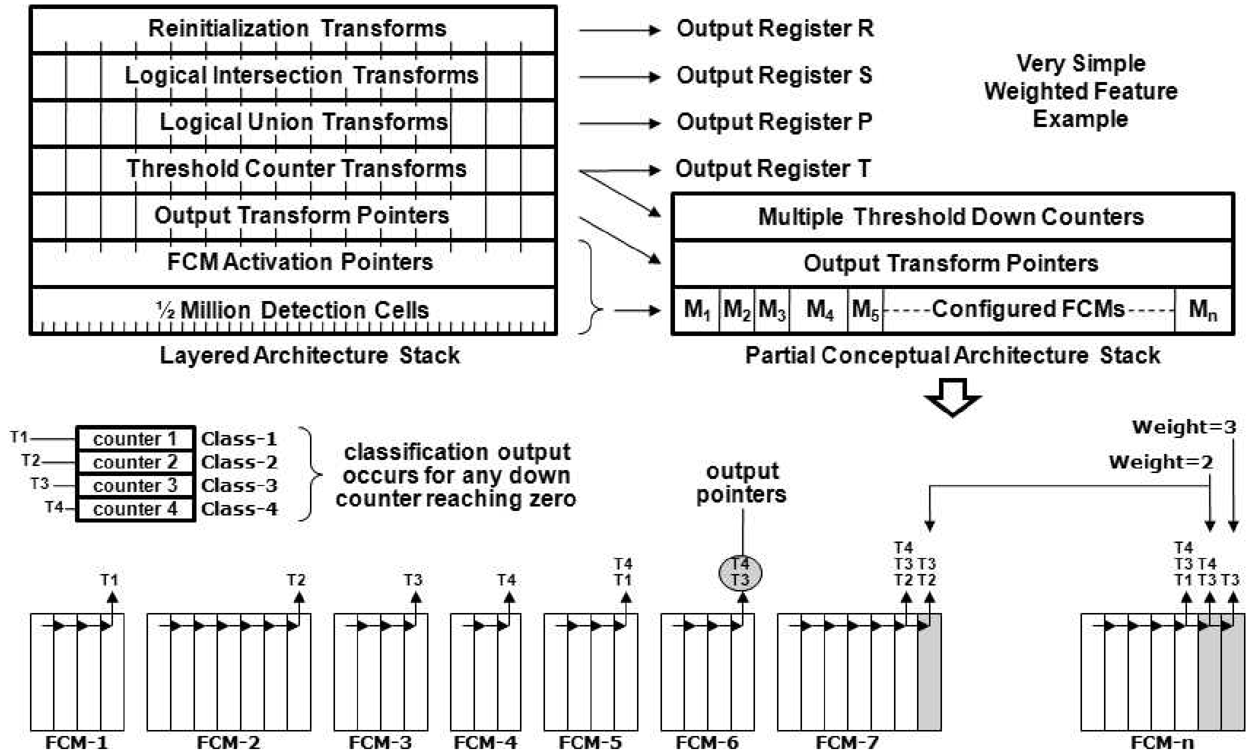


Figure 4. This conceptual depiction of additional features of the processor's architecture, at the top, places the Feature Cell Machines (FCMs) and their association with down counters in context. The bottom depiction shows a large number of multi-element FCMs, which are used to determine a classification. Classification-1, for instance, may have its down counter initialized to 3, while Classification-4 may have its down counter initialized to 4. Thus, Classification 4 can occur in multiple ways, whereas Classification-1 must have one each of the three specifically designated features.

The work reported here attempted to find sufficient connective concepts between the two end points that would warrant a next stage investigation. It is not suggested that the connecting leverage points discussed here and in the prior article are the only way to approach the problem effectively, nor that they are completely sufficient, but rather that they appear promising as a foundation for a solution path worth exploring.

This is preliminary work that sought and constructed a basis from which to investigate algorithms that can detect safety- and security-threatening behavior among UAS working in teams; where speed, accuracy, and breadth of comprehension are key performance factors. The work suggests that a promising basis exists in combining recognition of social behavior and trajectory behavior with a technology that can manage a vast quantity of stored reference patterns structured and accessed in a feedforward chunked hierarchy. Of importance, that technology must employ parallel recognition capabilities that eliminate any need for time-consuming search or sequential algorithms.

Social comparison theory guides us to a comparison of an agent's behavior pattern against behaviors of others on the team, against mission plans, against defined patterns of normal behavior, and against defined patterns of aberrant behavior. Trajectory behavior classification could be considered a special subset of social behavior detection for loosely choreographed teamwork.

Expertise theory, if it can be called that, guides us to a need for an extremely large number of reference patterns that can be simultaneously compared relative to a dynamic situation, eliminating time for sequential evaluation and reasoning steps, and eliminating much of the otherwise selective monitoring and pattern simplification that increases uncertainty.

This work was prompted by the growing concern for testing methods that can keep pace with the growing intelligence of UAS. It is suggested that a very different approach is required, one that never stops, one that carries the test environment into the operational environment, on board every UAS eventually.

There is precedence for and experience with this approach: the training and vetting of Special Forces operatives. One point to note is that we appear comfortable moving operatives into field status after some "testing" period, even though we know they will face situations that have not been tested. Another point to note is that these operatives on mission are always being evaluated by their peers, who rely on the integrity of each and every member of a team.

UAS will operate outside of ready observation and are subject to attrition by enemy destruction. In such live cases, aberrant behavior must be detected and evaluated to sense control penetration by the enemy, as well as malfunction, that threatens the mission or might provide a disabled UAS to the enemy for recovery post-mission.

Continued study will investigate appropriate classes of behavior for monitoring, the nature of expert-level detection capability, and suitable pattern representations for whatever technology is employed. ❑

Rick Dove is a founding partner of Kennen Technologies, an adjunct professor in the School of Systems and Enterprises at Stevens Institute of Technology, and chairman of Paradigm Shift International. He is author of Response Ability—The Language, Structure, and Culture of the Agile Enterprise; *and* Value Propositioning—Perception and Misperception in Decision Making. *He is an active researcher and consultant in agile and self-organizing systems. He has 40 years' experience at start-up, turn-around, and interim executive management. He holds a bachelor of science degree in electrical engineering from Carnegie Mellon University and has done graduate work in computer science at U.C. Berkeley. E-mail: rick.dove@kennentech.com*

References

Arkin, R. C. 2007. *Governing lethal behavior: embedding ethics in a hybrid deliberative/reactive robot architecture*. Technical Report GIT-GVU-07-11. Atlanta, Georgia: Mobile Robot Laboratory, College of Computing, Georgia Institute of Technology, http://www.cc.gatech.edu/ai/robot-lab/online-publications/formalizationv35.pdf (accessed May 15, 2009).

Cowan, N. 2001. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24 (1): 87–185. http://www.bbsonline.org/documents/a/00/00/04/46/bbs00000446-00/bbs.cowan.html (accessed May 15, 2009).

DoD (Department of Defense) Office of the Secretary of Defense. 2008. *Unmanned and Autonomous System Testing (UAST) Broad Agency Announcement <BAAUAST0002>*. April 4, 2008. Washington, D.C.: DoD.

Dove, R. 2009a. Pattern recognition without tradeoffs: Low-cost scalable accuracy independent of speed. In *Proceedings Cybersecurity Applications and Technologies Conference for Homeland Security (CATCH)*, March 3–4, Washington, D.C., 255–260. Piscataway, NJ: IEEE. http://www.kennentech.com/Pubs/2009-PatternRecognitionWithoutTradeoffs.pdf (accessed May, 15, 2009).

Dove, R. 2009b. Paths for peer behavior monitoring among unmanned autonomous systems. *The ITEA Journal* 30 (3): 401–408.

Flack, J. C., M. Girvan, F. B. M. de Waal, and D. C. Krakauer. 2006. Policing stabilizes construction of social niches in primates. *Nature* 439 (7075): 426–429.

George, D. 2008. *How the Brain Might Work: A Hierarchical and Temporal Model for Learning and Recognition*, Ph.D. dissertation, Department of Electrical Engineering, Stanford University, June. http://www.numenta.com/for-developers/education/DileepThesis.pdf (accessed May 15, 2009).

Heinze, J. 2003. Reproductive conflict in insect societies. In *Advances in the Study of Behavior*, ed. P. Slater and J. Rosenblatt, 1–57. New York: Academic Press.

Intille, S. S. Month 1999. Visual recognition of multi-agent action. Ph.D. dissertation, MIT. http://web.media.mit.edu/~intille/papers-files/thesis.pdf (accessed May 15, 2009).

Intille, S. S. 2001. Recognizing planned, multiperson action. *Computer Vision and Image Understanding* 81 (3): 414–445. http://web.media.mit.edu/~intille/papers-files/cviu01.pdf (accessed May 15, 2009).

Jain, A. K., R. P. W. Duin, and J. Mao. 2000. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (1): 4–37.

Kahn, J. 2007. Wayne Gretzky-style 'field sense' may be teachable. *Wired Magazine*, May 22. http://www.wired.com/science/discoveries/magazine/15-06/ff_mindgames# (accessed May 15, 2009).

Klein, G. 1998. *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Kotler, S. 2009. Here Come the Neurobots. *h+ Magazine*, June 5. http://www.hplusmagazine.com/articles/ai/here-come-neurobots (accessed July 15, 2009).

McKinstry, J. L., G. M. Edelman, and J. L. Krichmar. 2006. A cerebellar model for predictive motor control tested in a brain-based device. *PNAS* 103 (9): 3387–3392. http://www.pnas.org/content/103/9/3387.full.pdf+html (accessed May 15, 2009).

Merolla, P. A., and K. Boahen. 2006. Dynamic computation in a recurrent network of heterogeneous silicon neurons. In *Proceedings IEEE International Symposium on Circuits and Systems*, May 21–24, Kos, Greece, 4539–4542. http://www.stanford.edu/group/brainsinsilicon/pdf/ISCAS2006_merolla.pdf (accessed May 15, 2009).

Miller, G. A. 1956. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review* 63 (2): 81–97. www.musanim.com/miller1956/ (accessed May 15, 2009).

Monnin, T., F. L. W. Ratnieks, G. R. Jones, and R. Beard. 2002. Pretender punishment induced by chemical signaling in a queenless ant. *Nature* 419 (6902): 61–65.

Moshkina, L., and R. C. Arkin. 2007. *Lethality and autonomous systems: Survey design and results*. Technical Report GIT-GVU-07-16. Atlanta, Georgia: Mobile Robot Laboratory, College of Computing, Georgia Institute of Technology. http://www.cc.gatech.edu/ai/robot-lab/online-publications/MoshkinaArkinTechReport2008.pdf (accessed May 15, 2009).

Myers, D. 2008. Play and punishment: The sad and curious case of Twixt. In *Proceedings of The [Player] Conference*, August 26–29, Copenhagen, Denmark.

Ratnieks, F. L., K. R. Foster, and T. Wenseleers. 2006. Conflict resolution in insect societies. *Annual Review of Entomology* 51: 581–608. http://www.people.fas.harvard.edu/~kfoster/RatnieksetalAnnreventomol2006.pdf (accessed May 15, 2009).

Riesenhuber, M., and T. Poggio. 1999. Hierarchical models of object recognition in cortex. *Nature Neuroscience* 2 (11): 1019–1025. http://cbcl.mit.edu/projects/cbcl/publications/ps/nn99.pdf (accessed May 15, 2009).

Ross, P. 1998. As I see it: Flash of genius. *Forbes*. November 16. http://www.forbes.com//forbes/1998/1116/6211098a.html (accessed May 15, 2009).

Ross, P. 2006. The expert mind. *Scientific American* 295 (2): 64–71.

Schemmel, J., A. Grubl, K. Meier, and E. Mueller. 2006. Implementing synaptic plasticity in a VLSI spiking neural network model. In *Proceedings International Joint Conference on Neural Networks*, July 16–21, Vancouver, BC, Canada, 1–6. http://www.kip.uni-heidelberg.de/Veroeffentlichungen/download.cgi/4620/ps/1774.pdf (accessed May 15, 2009).

Serre, T. June 2006. Learning a dictionary of shape-components in visual cortex: Comparison with neurons, humans and machines. Ph.D. dissertation, Massachusetts Institute of Technology. http://cbcl.mit.edu/publications/ps/MIT-CSAIL-TR-2006-028.pdf (accessed May 15, 2009).

Serre, T., A. Oliva, and T. Poggio. 2007. A feedforward architecture accounts for rapid categorization. *Proceedings National Academy of Sciences* 104 (15): 6424–6429. http://cvcl.mit.edu/Papers/SerreOlivaPoggioPNAS07.pdf (accessed May 15, 2009).

Acknowledgments

This work was supported in part by the U.S. Department of Homeland Security award NBCHC070016. Curt Harris invented the processor architecture.

ITEA Journal 2009; 30: 513–522

Testing Unmanned Autonomous System Communications in a Live/Virtual/Constructive Environment

Eric Paul Parker, Ph.D., Nadine Elizabeth Miner, Ph.D., Brian Peter Van Leeuwen, and James Brian Rigdon

Sandia National Laboratories, Albuquerque, New Mexico

Test and evaluation (T&E) of unmanned autonomous systems (UAS) is a challenge for decision makers because of their complexity, network communication dependencies, and requirements for rapid deployment. Sandia National Laboratories has developed a live/virtual/constructive (LVC) T&E framework used to emulate live UAS while incorporating network communication effects between platforms within mission-specific virtual/constructive simulations. This framework possesses sufficient fidelity to support UAS decision making and measure performance based on system-of-system (SoS) contributions. In this framework constructive assets can affect and/or be affected by live and virtual assets via high-fidelity communications models and system-in-the-loop technologies. Using system-in-the-loop enables assessment of distributed communications that rely on wireless networks without deploying real wireless networks. Models of wireless assets and supporting platforms are included along with real applications. Experimental results include simulated networks of commercial and military radios and demonstrate the importance of incorporating wireless communication network effects when evaluating distributed SoS operation.

Key words: Simulated wireless communication; simulation environments; shared wireless spectrum; system in the loop; communication routing protocols; cyber disruptive techniques; modeling; military radio; communication capacity; interference; robotic vehicles.

Given the increasing pace of development for new technologies, the Test and Evaluation (T&E) community is finding it necessary to adapt its approach to evaluating the effectiveness of new and deployed technologies. This is especially true with respect to Unmanned Autonomous Systems (UAS). These systems have shown tremendous benefit to the warfighter in current operations and are being deployed on an accelerated schedule that leaves little time for traditional T&E. When dealing with UAS, there is the additional fact that they are complex, network centric systems, and testing cannot be limited to physical aspects of individual systems alone (Macias 2008):

> *"In the future that we are facing today, UASs will be deployed on a timeline of months instead of years. Systems that are being developed in industry and academia have utility today to a warfighter who is facing enormous challenges. The question of how to start testing these systems in parallel with development may require us to move beyond the traditional test focus and towards a test strategy that covers the entire acquisition cycle from cradle to grave. The challenges of testing UASs are moving from simple system test toward the world of complex systems engineering."*

Testing a complex system requires that all of the system's components be included. This is especially true of technologies that must function in a network centric environment. In addition, to meet the increasing demands of UAS T&E, the Department of Defense needs a seamless, robust, and rapidly deploy-

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

able Live/Virtual/Constructive (LVC) T&E framework. To meet these requirements, Sandia National Laboratories has extended an LVC framework developed using internal research and development funds to support UAS T&E.

Using a LVC framework for UAS T&E provides a number of benefits. The LVC T&E framework provides the operational context needed to evaluate performance of new technologies that must function with existing systems and of mature technologies being applied to new problems with respect to mission objectives, execution, environment, and personnel. It also provides the ability to 1) identify requirements for future T&E infrastructure (with respect to instrumentation, data capture, and processing), 2) to vary the scale of problems addressed, 3) to incorporate the complexity (bandwidth, latency, fidelity) required when engineering advanced technologies, 4) to preplan range operations, and 5) to shorten exercises by using fewer live assets. An LVC T&E framework allows fewer live test fires, requires less instrumentation and telemetry for each asset, and reduces the time needed to put test instrumentation on a UAS. Finally, another major benefit of using the LVC test-bed approach has to do with obtaining radio frequency spectrum permission. Wireless communication spectrum is controlled and licensed for specific use. Permissions must be obtained if performing live radio frequency radio experiments in controlled frequency bands. In the LVC test bed, only the wireless control link providing connectivity between the live robot and the simulator is a live radio link. Because this is not the wireless communication link under study, it can communicate over the license-free industrial, scientific, and medical (ISM) band.

Previous approaches used for SoS T&E have involved performing experiments where the entire experiment is simulated or experiments where the entire experiment is live. Several efforts have described T&E frameworks that combine simulation, emulation, and real devices. Research being done at UCLA (Zhou, Ji, and Bagrodia 2006) attempts to combine real and simulated wireless systems but segments the wireless physical layers into isolated real and simulated segments with no interaction between the two segments. The approach using isolated segments of the wireless physical layer is limiting, and the capability to merge the real wireless physical layer with the simulated wireless physical layer is necessary to obtain overall system realism in experiments. Other wireless networked system evaluation test beds deploy a large number of real wireless radios that can be configured for specific experiments (Hadjichristofi et al. 2007). These test beds provide realism when examining systems that can be emulated with the test bed radios and topology; however, this approach is limiting in its ability to examine advanced radio concepts that have not yet been developed in hardware. These examples demonstrate that while there is significant emphasis on simulation-based live training, there is no indication that there is a systematic effort to determine the best approach to link live and constructive communications. The research described in this article addresses testing LVC assets and their supporting network communications in a systematic fashion with a variety of communication equipment represented with simulated models, emulated or real devices, and the ability to employ the methods simultaneously. Using this approach, simulated network communications can provide data transport between virtual, constructive, and live elements.

The LVC test bed

Sandia National Laboratories has developed an LVC Integrated Control Framework (LVC ICF) and an associated test bed for testing and evaluating SoS effectiveness for achieving mission objectives. The modeling and simulation (M&S) capabilities of the LVC ICF, and test bed include physics-based technology models, high-fidelity terrains, manned and unmanned vehicles (ground and air), both wired and wireless communications, several types of unattended ground sensors (UGS), and cognitive models of personnel.

SoS T&E requires that the communication network be represented as realistically as possible. Creating real networks with real hardware is challenging or, in some cases, impossible. For example, design and incorporation of new military radio technologies often requires SoS T&E prior to the real hardware being available. The LVC test bed brings together simulated network communications with real applications running on live platforms. The test bed enables rapid and cost-effective SoS operation analysis including the impact on communications due to network limitations and environmental effects (Parker et al. 2009; Rigdon, Parker, and Miner 2008). Through the LVC test bed, evaluation of real system applications that depend on wireless communications is possible without deploying real wireless networks. For example, the impacts of impaired wireless communication performance on overall system operation can be assessed. Issues such as end-to-end data packet latency or effects of dropped data packet rates on application performance can also be assessed). In addition, communication network technologies such as Quality of Service (QoS) or multicast protocols can be assessed. Increasing scale with respect to the number of platforms and the

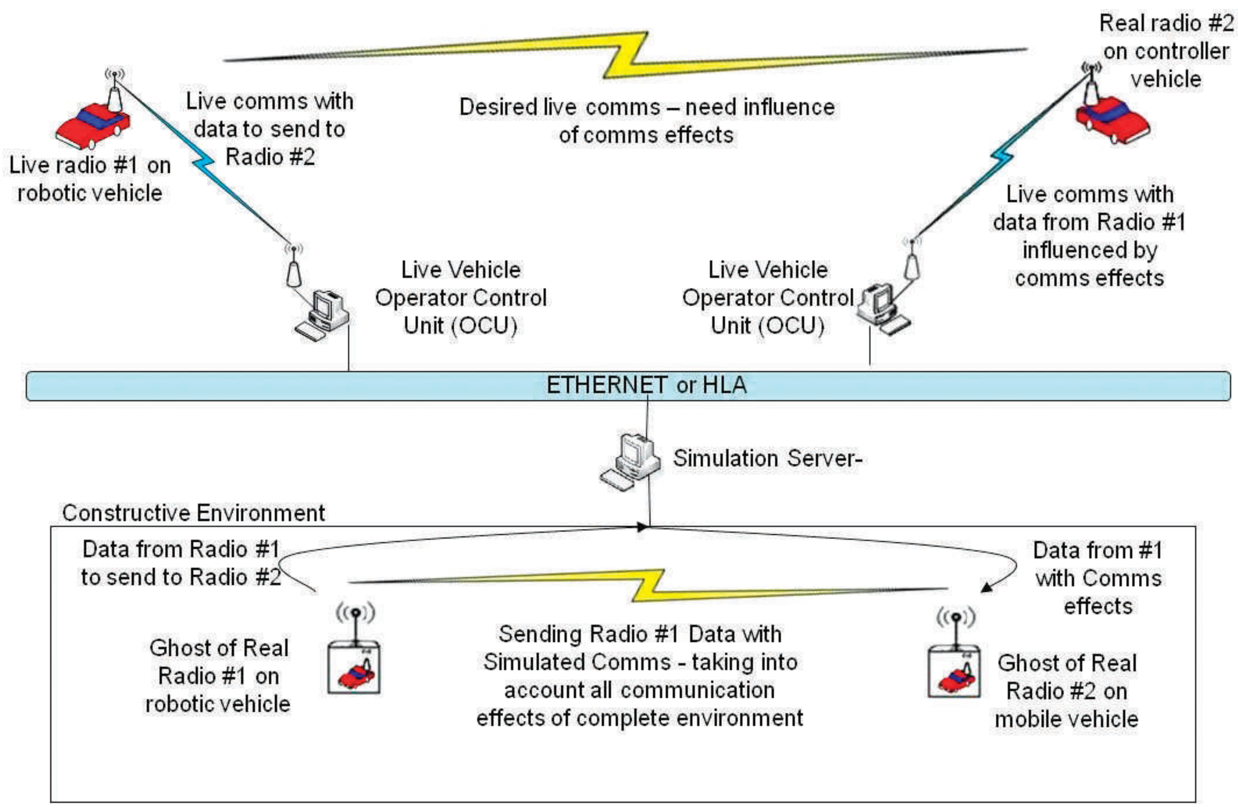


Figure 1. LVC experiment—two live vehicles using a simulated communication network.

resulting impact on the wireless network can also be assessed.

In this article, the LVC test bed is described and an example scenario is presented with results showing the value of utilizing a framework such as the LVC test bed. The test bed represents network communication by leveraging network communication simulation environments such as OPNET Modeler (OPNET Technologies, V. 15.0, 2009), Joint Communications Simulation System (JCSS V. 9.0, 2009), and QualNet (Scalable Network Technologies, V. 4.5, 2009). The shared wireless spectrum between live and constructive elements is realistically represented in the test bed by employing JCSS.

The LVC test bed and the integration of simulated network communications are described in the context of an experiment involving an unmanned ground vehicle (UGV). The experiment includes multiple UGVs, the software control system, the operator control unit (OCU), and a shared wireless communication network as shown in *Figure 1*. The system design depends on continuous, reliable flow of data from each autonomous robot to the OCU. Similarly, the UGV depends on continuous updates of differential GPS data, system state data, and streaming video and audio data transmitted via the supporting wireless communication network (especially when executing behaviors requiring a higher level of autonomy). For this experiment to produce useful data, LVC elements interact as if they were nodes in a single integrated system. Constructive elements are modeled and simulated, and interact with the live and virtual elements.

The experiment is complicated because LVC UASs communicate over a shared wireless communication network that must account for multiple nodes, real or constructive, accessing the network. Thus, real and constructive UASs transmitting at the same time cause radio channel interference if they are in close proximity; however, if they are separated by an adequate distance, they can share the same radio frequency spectrum. Because this shared wireless spectrum is an important aspect of system operation, it has to be realistically represented in the test bed.

Additionally, because a live UAS is included in the experiment and no physical cabling is used to provide a communication link to the live UAS, a wireless communication link is provided. The live communication link is a live radio system that is part of the UAS under test and is a readily available wireless communication link using IEEE 802.11 LAN/MAN Wireless LANS (IEEE 802.11g, 2003). The live wireless communication is considered an artifact of the experiment in that this link is not intended to suffer from shared wireless spectrum interference or other wireless channel degradations. The live wireless link has data throughput greater than the communication link that is being assessed and is represented in the simulation. In addition, the live wireless communication link is used for immediate emergency stop of the UAS (i.e., e-stop). This safety stop is necessary as a precaution to prevent personnel harm or equipment damage.

Given that the LVC test bed combines live platforms with constructive platforms that potentially share the same physical area, a means to provide for the live wireless communications to affect the simulated wireless communications and for the simulated wireless communications to affect the live wireless communications is necessary. In the LVC test bed, the shared wireless communication channel spectrum impacts and other channel degradation are accounted for in the simulation environment. All platforms participating in the experiment have a representation in the simulation. Live platforms that are moving through live terrain are represented in the simulation as "ghost images" of the live platform. Because all platforms are represented in the simulation, the simulation is where all wireless communication impacts are performed.

Modeling networked communications

This project's focus is to provide a platform for developing and exploring methods for integrating, testing, and evaluating communications between distributed LVC entities. The key to this effort lies in performing high-fidelity simulations of live, virtual, and constructive asset communications. This is accomplished by importing live asset attributes such as radio characteristics and geographic position into a model that represents the live asset in the communication simulator. The communication simulator also includes detailed models of the scenario terrain and models of the wireless communication channels for the scenario. This enables live asset communications to be affected by virtual and constructive assets because live communications are passed through the simulated environment. With this approach, for example, the communication two real assets can be jammed by a constructive adversary. Key aspects to implementing the network communications in LVC experiments are

- Live platform communications must be affected by virtual and constructive platform communications and simulated spectrum impairments.
- Live mobile platforms will have experiment control communications that are not affected by simulated channel impairments. As long as a high quality wireless link exists between the controller and mobile platform, critical experiment control can be communicated to real platforms, such as emergency stop (i.e., e-stop).
- Intended live mobile platform communication equipment can be modeled with simulated radio systems.

Our approach to creating LVC experiments requires that communications between live platforms (with the exception of e-stop) are imported into the simulation, affected by virtual, constructive, and other live communications traffic, and then exported to the live transceiver. Thus, live wireless platforms have a real wireless link and a simulated wireless link. To obtain high-fidelity models of the various internet protocol (IP) communication network protocols and devices, we had the test bed leverage network communication simulation environments. For the experiment discussed in this article, the OPNET Modeler and the JCSS are used to simulate the communication network in real time. To enable the import and export of application data and network control data, we used the OPNET system-in-the-loop (SITL) module in the Modeler and extended it for use in the JCSS. Real hardware to be tested is connected via SITL to the simulated radio models. Additionally, the LVC test bed is used to assess military equipment using military radio systems, making the availability of military radio models critical to this research. Thus the integration of JCSS, which has models of military radios, such as Single Channel Ground-to-Air Radio System (SINCGARS), the Enhanced Position Location Reporting System (EPLRS), and various satellite-based systems, plays a key role in the LVC test bed.

Our experimental approach also has the following benefits:

- It enables performing experiments that require a greater number of platforms with radios than are available as real hardware.
- In addition to platforms such as ground-based robotic vehicles that are represented as live and/or constructive, an LVC experiment may also include aerial or space-based platforms that carry communication payloads used for relays or data servers. Aerial or spaced-based platforms are difficult or impossible to obtain for an experiment, yet their influence on an experimental outcome can be substantial.
- It provides realistic representation of platform network architectures that may include routers, firewalls, data encryptors such as a high assurance internet protocol encryptor, or other network devices.

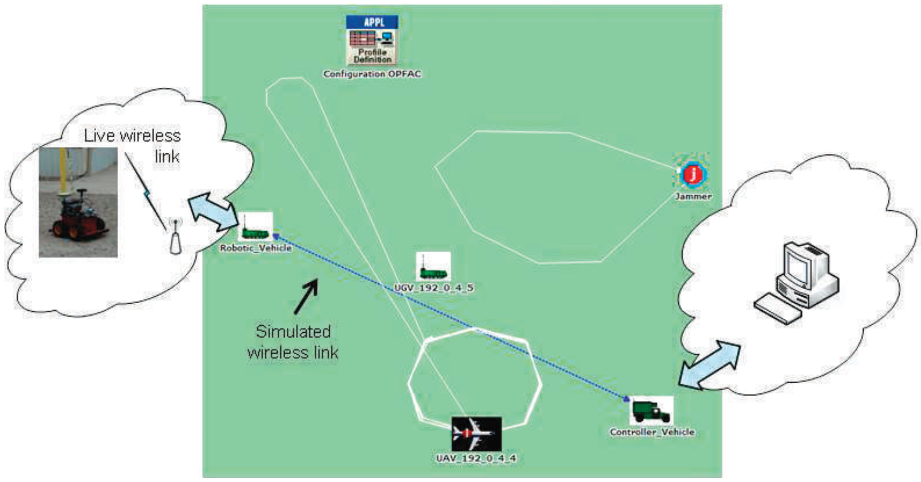


Figure 2. Live platforms communicating via an OPNET/JCSS simulated network.

- It enables incorporation of simulated electronic warfare capabilities such as jamming or other cyber disruption techniques.

LVC experiments

Two experiments were conducted to measure the performance of different communication architectures and configurations for a conceptual UAS that involves distributed sensors over a multithousand square meter area. The sensors have low profile antennas and require an unmanned aerial vehicle (UAV) to support relay of their communication links. The experiments include a controller manned by a live human operator, a single tele-operated UAS, and a constructive UAV with a communication payload that provides for the effects of message relay between the controller and UAS. In addition, we also include a constructive jammer disruption to the communication link.

Experiment 1: Wireless Network with Relay Platforms

The first experiment, shown in *Figure 2*, includes a simulated and a live wireless communication link. A single simulated wireless link provides communications between two wired-network local area networks (LANs), representing the robotic controller and UAS. Each LAN, shown in *Figure 3*, includes real and simulated workstations and servers along with other network devices. The live wireless link is an IEEE 802.11 based wireless link providing communication between the live platforms, a Pioneer UGV, and the live UGV controller that interfaces to the simulation. Constructive platforms include the UAV platform circling overhead and the supporting wireless IP network. The control messaging and the sensor data are transported via the simulated wireless network. In the scenario, no direct communication exists between the constructive control vehicle and the constructive UAS because they are equipped with low-gain antennas that are positioned close to the ground. All communication is relayed through a constructive UAV equipped with a communication payload. The UAV has multiple missions and during the experiment, moves away from its relay position for a brief period, resulting in a degradation of the communication link. In this experiment, simulated communication effects are obtained using OPNET Modeler together with JCSS and the SITL module.

The analysis objective of the first experiment is to assess UAS operation with a live operator controlling a live UAS. The simulated wireless network is instrumented to collect data throughput, dropped packets, and end-to-end latency. Several experiments are performed using different radios and protocols. More specifically, a mobile ad hoc network routing protocol operating with a high-data-throughput radio is assessed first with the Ad hoc On-demand Distance Vector (AODV) protocol and second the Optimized Link State Routing (OLSR) protocol. The primary distinguishing factor between these two ad hoc routing protocols is that AODV is a reactive (i.e., on demand) routing protocol that performs route discovery only when needed, and OLSR is a proactive routing protocol that periodically performs route discovery. A proactive routing protocol typically has much shorter route discovery and end-to-end latency because routes are current in the routing tables. However, a proactive routing protocol has a greater amount of overhead control traffic to maintain the routing tables. In contrast, AODV performs a route discovery only when a route is required to the destination, resulting in a greater end-to-end latency if routes are changed but a lower amount of overhead control traffic. Overall, the experiment enables the T&E of various ad hoc routing

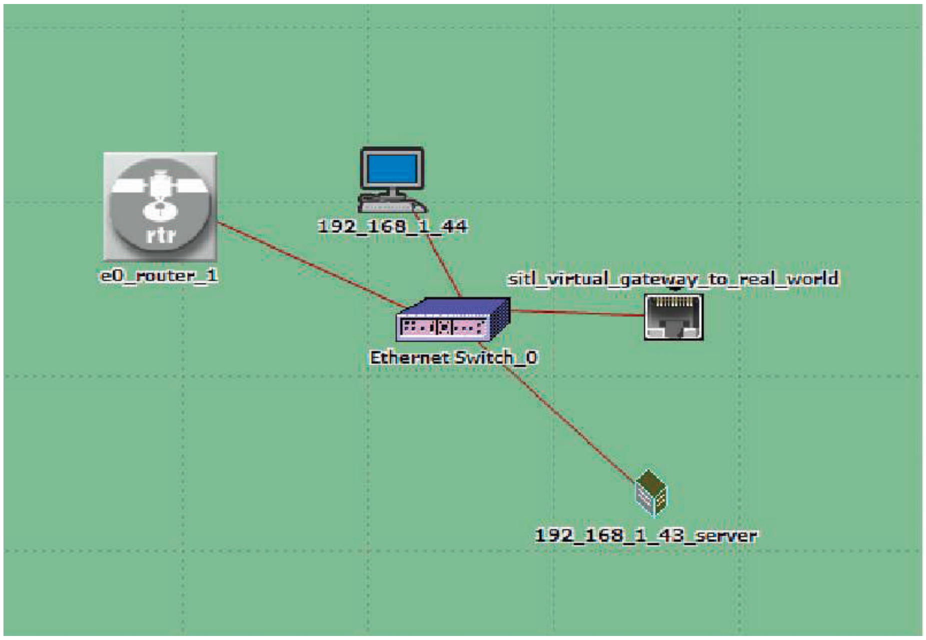


Figure 3. Communication devices, including SITL interface, comprising robotic vehicle.

protocols along with various radio physical-layer characteristics such as data rates and power levels.

The experiments are used to evaluate the responsiveness of the system under various network architectures and configurations. The instrumentation of the simulated communication network provides for network performance data collection that can provide quantitative performance measurements. In addition, the experiments assess the data throughput of the supporting relay platforms as the scenario jammer moves through its trajectory.

The experiment begins with selecting the UGV as its relay platform because it is in closer proximity to the source and destination platform. As the experiment progresses, the jammer moves closer to the UGV and disrupts the connectivity with the source and destination platforms. Because the constructive platforms at both the robotic vehicle and controller are configured with mobile ad hoc routing protocols, the platforms reroute through the UAV. The UAV provides a data forwarding capability until approximately 10 minutes into the experiment. At approximately 10 minutes, the jammer has moved into a position where it is causing interference to both the UGV and UAV relay platforms. At this time there is no connectivity between UAS and the controller. At approximately 15 minutes the jammer has moved away and the UAV reestablishes connectivity.

Experiment 1: Results

Figures 4 and 5 illustrate the data throughput of the UGV and UAV relay platforms for a system configured with AODV routing and OLSR routing. The results illustrate the data throughput as the system transitions from the UGV relay to the UAV platform. In this experiment, the constructive part of the experiment affected the live communication between the live platforms as if a live UGV or UAV relayed the messages. The live messages were also affected by routing protocol transitions and periods of no communications.

In this experiment the latency of networks configured with AODV and OLSR were also examined. The network employing AODV, which is a reactive protocol, required approximately 0.18 seconds to discover a new route when the jammer moved to a position where the original route could not be used. In contrast the OLSR configured network was able to switch routes in much less time because it had proactively discovered routes before they were needed. *Figures 6 and 7* illustrate the end-to-end latency for application-layer packets transported by the simulated network. The results indicate that OLSR has a lower end-to-end latency compared with AODV.

Experiment 2: Wireless network using EPLRS military radio

In a second experiment we use the LVC test bed to assess the tele-operated UAS using the EPLRS military radio. EPLRS uses virtual circuits, called needlines, to set up communications between EPLRS radio sets. A "needline" is a common set of time and frequency resources shared among two or more EPLRS radio sets to exchange data. The needlines

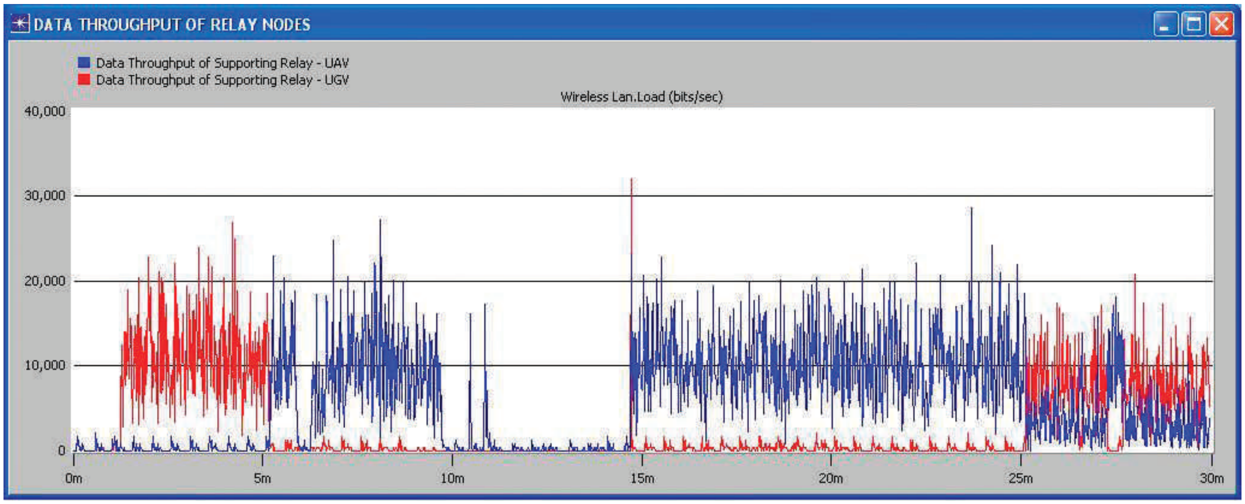


Figure 4. Data throughput of relay platform—AODV routing.

can be either many-to-many or one- to-one, and every EPLRS radio set can support many needlines at the same time, thus enabling radio sets simultaneously to send and receive different information from different radio sets. The configuration of time and frequency resources has a direct impact on the communication capacity of a needline while the choice of waveforms provides a tradeoff between data rate, transmission range, and error resiliency of the needline. The LVC test bed provides a means to evaluate system operation of various EPLRS networks.

The experiment configuration using EPLRS radios to support robotic system control is notional; however, the main point of the experiment is to illustrate the use of the LVC test bed.

Experiment 2: Results

Figures 8 and 9 illustrate experiment results of data throughput and end-to-end latency of an EPLRS configured with two Carrier Sense Multiple Access (CSMA) needlines. The first CSMA needline must always exist as a Private Virtual Circuit (PVC) needline and is used by all EPLRS platforms to exchange control messages with each other. The second CSMA needline is used to communicate vehicle control messages and return sensor data. The CSMA needlines

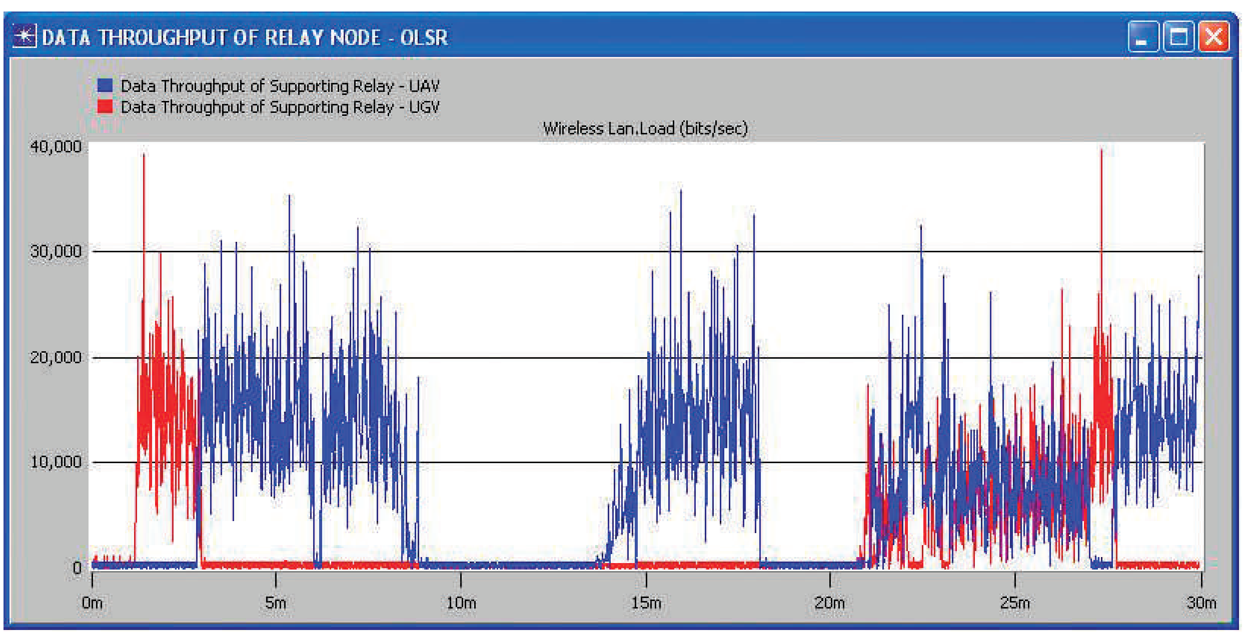


Figure 5. Data throughput of relay platform—OLSR routing.

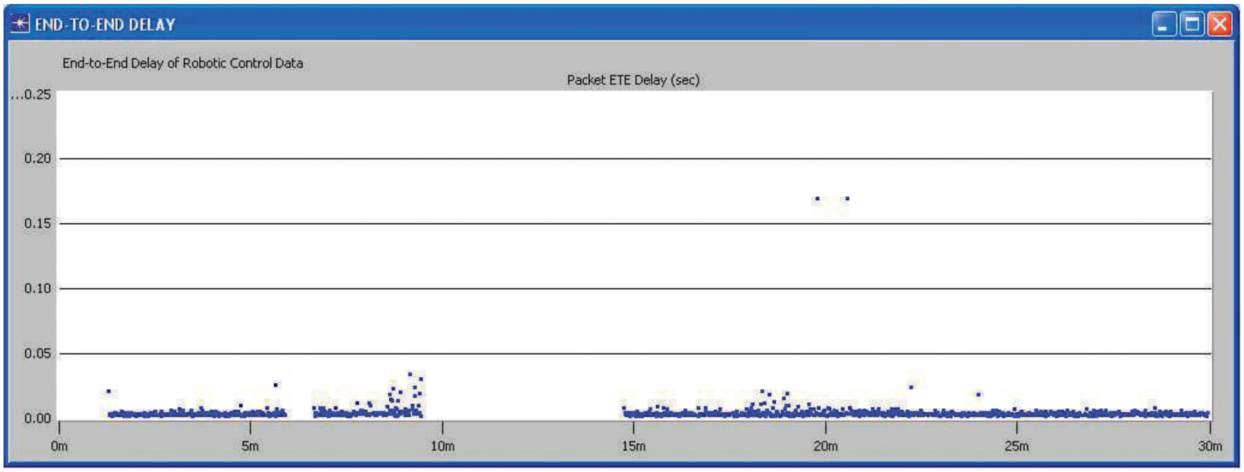


Figure 6. End-to-end latency when using AODV routing.

are configured for two-hop relay coverage; thus, each transmission is silent for two consecutive timeslots after each transmission. The period of silence is necessary for the relay or forwarding of messages by relaying platforms. In EPLRS, this is not a dynamic configuration and the silent period occurs whether the message is being relayed or not. *Figure 8* illustrates the data received while the jammer is moving along its trajectory. The EPLRS communication is more resilient than the radios used in Experiment 1 with the jammer interference because it has better signal-to-noise at the receiver. *Figure 9* illustrates the end-to-end delay associated with the EPLRS. The end-to-end delay is larger than the radios used in Experiment 1. The increased end-to-end delay is larger for EPLRS because of the use of timeslots and the wait period for each transmission for the relay requirement of the supporting needline.

Conclusions and future work

To incorporate communication effects of the overall system, i.e., LVC, in an LVC experiment, all system communication must exist in the constructive domain. To accomplish this, all live communications are brought into the constructive domain where the effects of virtual and constructive communications can influence the live communications and live communications can affect constructive communications. The LVC test bed provides an experimental environment for assessing systems of systems that are dependent on wireless communication networks and various mobile platforms such as the UAS.

The LVC test bed

- provides a means to incorporate military radio technology in LVC experiments;

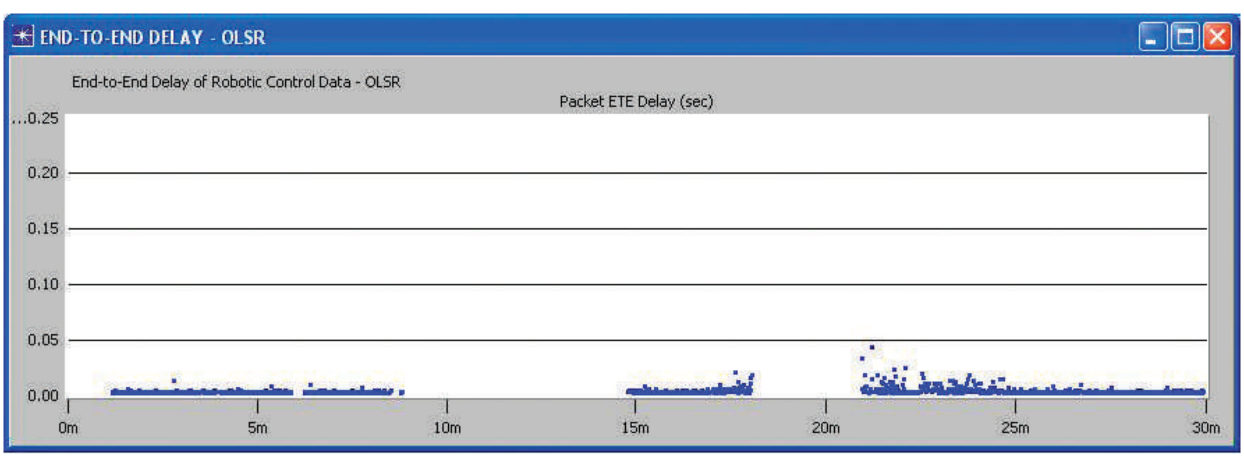


Figure 7. End-to-end latency when using OLSR routing.

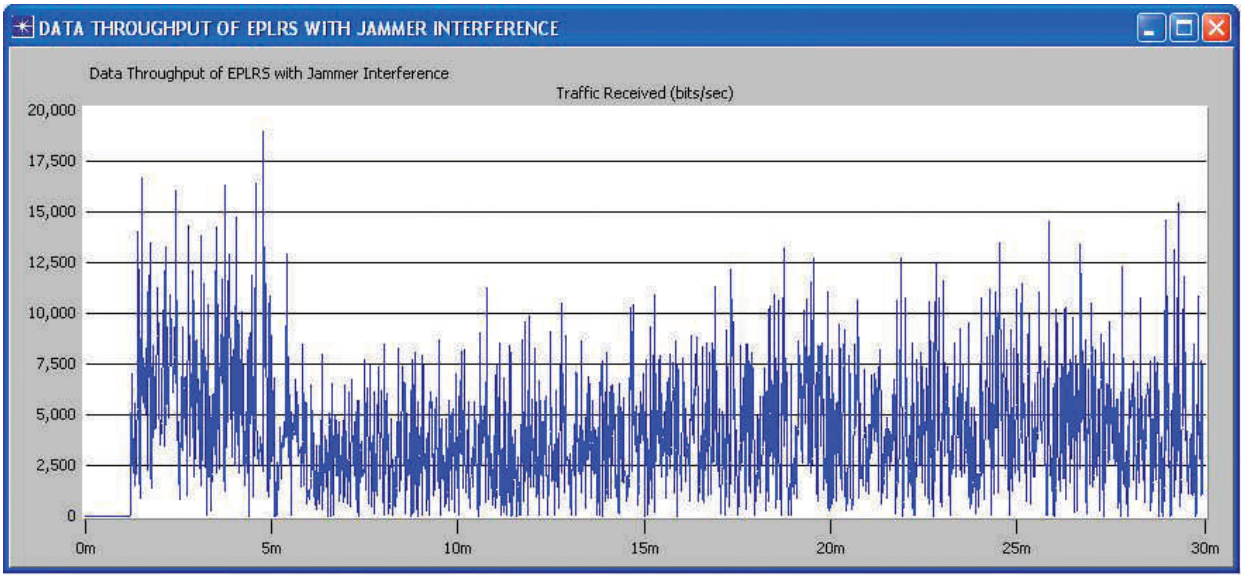


Figure 8. Data throughput of EPLRS.

- provides a means to eliminate issues related to obtaining spectrum frequency allocation and permission for live experiments;
- enables assessing systems with increasing scale;
- provides a means to analyze radio operation in a variety of difficult terrains such as urban and mountainous;
- enables analysis of proposed radio technology before hardware is available.

Experimental results show that with SITL technology, the effects of the LVC communication interactions can be incorporated and measured in an experiment that includes live platforms. The experiments provide measures of performance to support design, development, and deployment of UAS control systems that are also applicable to SoS design.

Work is underway to address the issues with SoS that have a large amount of communications. Experiments of increasing magnitude become a challenge for the simulated domain to run at real time. The research team is examining methods to deploy the simulated domain on more powerful computing platforms along with model abstractions and techniques to reduce computational requirements in the simulated environment.

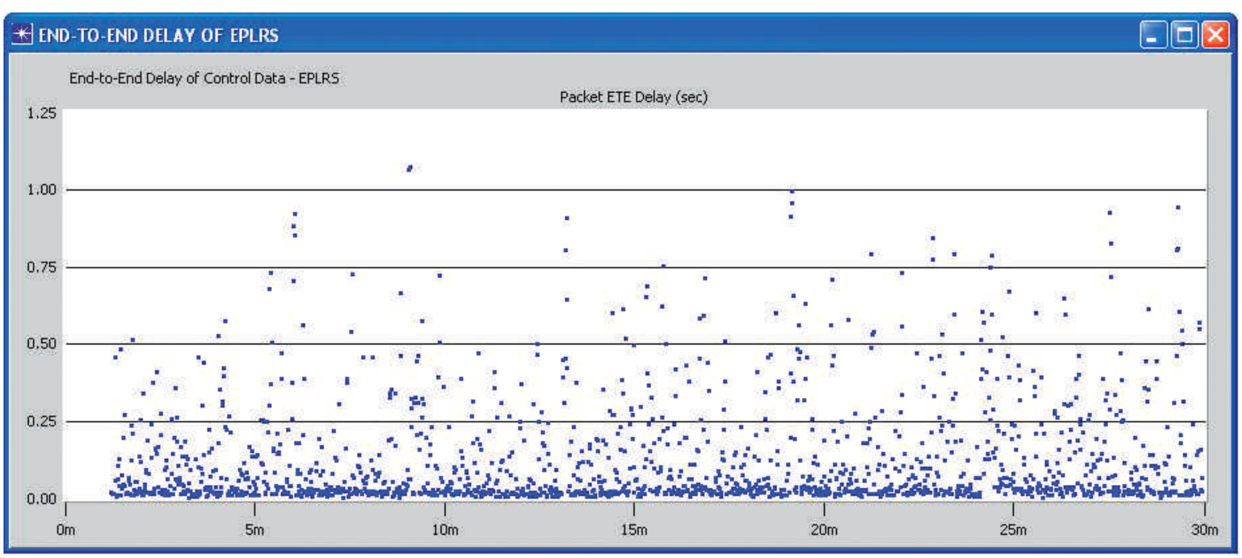


Figure 9. End-to-end latency of EPLRS.

Overall, the LVC framework and the associated test bed facilitate T&E of UAS by providing an experimental environment capable of testing a wide range of communications equipment and UAS. The results discussed here show that with SITL technology the effects of LVC communication interactions can be captured, measured, and categorized. As a result, the possible test-bed scenarios are limitless. Work is underway to expand the experimental scope to determine failure conditions and modes for the UAS being tested. This LVC test bed will be of great value for conducting studies that are beyond the scope of live testing alone. ❑

Eric Paul Parker is a principal member of the technical staff at Sandia National Laboratories in Albuquerque, New Mexico. He obtained his bachelor's of science degree in chemistry from the University of New Mexico in Albuquerque, New Mexico, and his master's of science degree and doctor of philosophy degree in physical chemistry from Cornell University in Ithaca, New York. Since coming to Sandia in 1997, he has worked on many modeling and simulation (M&S) projects and has extensive experience with Army M&S. He leads the Sandia Live/Virtual/Constructive (LVC) Test bed project and an OSD T&E project developing an LVC UAS T&E framework for wireless communications between live and constructive assets. E-mail: epparke@sandia.gov

Brian Van Leeuwen is a principal member of the technical staff at Sandia National Laboratories. For 10 years he has been involved in the analysis of secure information systems using modeling and simulation methods. Before that he played key roles in developing various secure communication systems and communication system modules. He received his bachelor's of sciences degree in electrical engineering in 1987 and his master's of sciences degree in electrical engineering in 1989, both from Arizona State University, Tempe, Arizona. In addition, he received a master's of business from the University of New Mexico, Albuquerque, New Mexico, in 2001. E-mail: bpvanle@sandia.gov

Nadine Elizabeth Miner is a senior member of the technical staff at Sandia National Laboratories. She is part of a team developing system of systems (SoS) M&S tools for analysis of complex, net-centric defense forces. She is the technical lead for the integrated SoS and high-fidelity network modeling analysis capability. Dr. Miner holds a bachelor's degree from the University of New Mexico in Albuquerque, New Mexico (1986), a master's degree from the California Institute of Technology in Los Angeles, California (1989), and a doctor of philosophy degree in computer engineering from the University of New Mexico (1998). E-mail: neminer@sandia.gov

James Brian Rigdon is a senior member of the technical staff at Sandia National Laboratories in Albuquerque, New Mexico. He has been involved in modeling and simulation for 9 years. He worked in the ISRC from 1992–1997 as a contractor, then again from 1999 to the present. From 1997–1999 he worked as a systems engineer for Silicon Graphics Incorporated. Prior to 1992, he developed real-time embedded software for Honeywell, Defense Avionics Systems Division. He graduated with a bachelor's of science degree in computer science in 1985 from the New Mexico Institute of Mining and Technology in Socorro, New Mexico. E-mail: jbrigdo@sandia.gov

References

Hadjichristofi, G. C., Brender, A., Gruteser, M., Mahindra, R., and Seskar, I. 2007. A wired-wireless testbed architecture for network layer experimentation based on ORBIT and VINI. *Proceedings of the 2nd ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization (WinTech 07), September, 2007, Montreal, Canada*.

IEEE 802.11g. 2003. *LAN/MAN Wireless LANS*. http://standards.ieee.org/getieee802/802.11.html (accessed March 15, 2009).

JCSS (Joint Communication Simulation System). 2009. V. http://www.disa.mil/jcss/ (accessed March 15, 2009).

Macias, Filiberto. 2008. The test and evaluation of unmanned and autonomous systems, *ITEA Journal of Test and Evaluation* 29 (4), 388–395.

OPNET Technologies. 2009. V. 15.0. http://www.opnet.com (accessed March 15, 2009).

Parker, Eric P., J. Brian Rigdon, Nadine E. Miner and Brian P. Van Leeuwen. Integrating live and constructive communications for unmanned autonomous systems. Presented at the ITEA LVC Conference, El Paso, Texas, January 2009.

Rigdon, J. Brian, Eric P. Parker and Nadine E. Miner. Distributed, seamless integration of live systems and simulations. Presented at the ITEA Technology Conference in Colorado Springs, Colorado, July 2008.

Scalable Network Technologies. 2009. V 4.5. http://www.scalable-networks.com/ (accessed March 15, 2009).

Zhou, J., Zhengrong, J., and Bagrodia, R. 2006. TWINE: A hybrid emulation test bed for wireless networks and applications. In *IEEE Infocom 2006. 25th IEEE International Conference on Computer Communications Proceedings*, April 2006, Barcelona, Spain. IEEE.

Zhou, J., Zhengrong, J., Takai, M., and Bagrodia, R. 2004 MAYA. Integrating hybrid network modeling to the physical world. *ACM Transactions on Modeling and Computer Simulation*. 14(2), 149–169.

ITEA Journal 2009; 30: 525–530

Use of General Aviation Aircraft as Surrogate for UAS Development, Test, and Integration

Joe Arvai
Aviation Systems Engineering Company (ASEC), Lexington Park, Maryland

There are over 155 unmanned aircraft system designs by 50 institutions in the United States alone. Many manufacturers are small companies lacking sophisticated test facilities. Few have ready access to high-demand restricted area ranges. They must rely upon testing in the National Airspace System, which requires Federal Aviation Administration approval, often with lengthy approval times and significant operational limitations. Using manned general aviation aircraft as surrogate platforms for testing the sensor systems or the air vehicle is a viable option. General aviation aircraft overcome limitations of operating in the National Airspace System and provide an airborne integration laboratory and added flexibility.

Key words: Unmanned aircraft; civil national airspace; safety; civil aviation; Federal Aviation Administration; high-density airspace; special use airspace; surrogate manned aircraft; flight testing; system integration.

Unmanned Aircraft Systems (UAS) are becoming increasingly common, and their manufacturers and operators are attempting to make them an integral part of the U.S. aviation community. UASs are now operated by Homeland Security, the Department of Defense (DoD), police departments, commercial industry, and the Forestry Service to name just a few. Although Unmanned Air Vehicles (UAVs) used to operate at only low altitudes (less than 400 ft), they now operate throughout the entire civil airspace from the surface to greater than 60,000 ft. There are currently over 155 UAS designs produced by 50 different companies and institutions in the United States alone. Many manufacturers are small companies lacking sophisticated test facilities, and few have ready access to restricted areas to test their systems. Although many are tested under U.S. government programs, primarily within the DoD, a large number are developed by private industry. This has led to an increasingly complicated issue of how to test these new systems economically and safely within the civil national airspace system.

National Airspace System environment

The airspace above the United States is designated the National Airspace System (NAS) and is strictly controlled by the Federal Aviation Administration (FAA). On average, about 50,000 flights use the NAS services each day with over 4800 aircraft airborne at any one time. Safety is a primary concern when operating in this high-density airspace.

The NAS had its beginnings with the Air Commerce Act of May 20, 1926. This was the cornerstone of the federal government's regulation of civil aviation and was designed to improve and maintain safety standards. The Act fostered air commerce issues and enforced air traffic rules, licensed pilots, certificated aircraft, established airways, and operated and maintained aids to air navigation. By 1958, the introduction of jet airliners and a series of midair collisions spurred passage of the Federal Aviation Act, which established the FAA.

Although UASs have been in existence for much of this time, only recently have they become a focus of the FAA. The first UASs to enter into the NAS were built for or by the military and tested within special use airspace, specifically, restricted areas. These early systems did not impinge upon the civil use airspace, and the limited amount of UAV testing was easily accommodated. But as UASs have proven their worth on the battlefield, they have grown in size and complexity, far outstripping the capacity of military airspace. Although there are over 1,000,000 sq mi of special use airspace in the United States, the amount usable by UAVs is limited to restricted areas

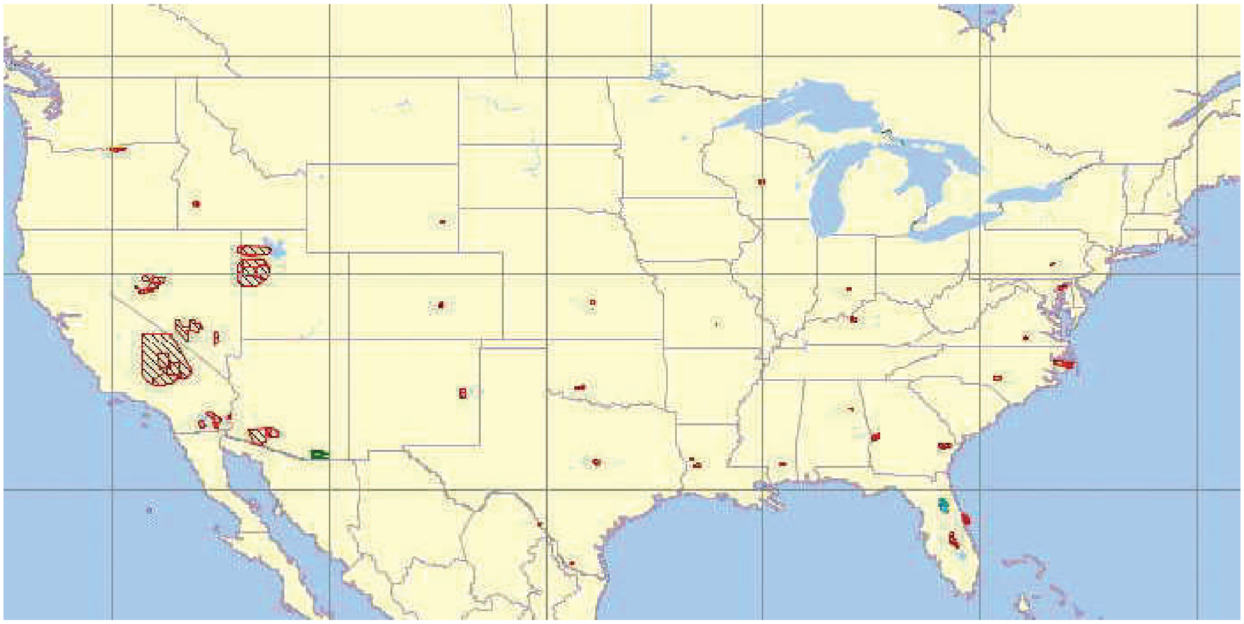


Figure 1. Special use airspace restricted areas.

and is only a small fraction of the existing space *(Figure 1)*.

The military has reduced and consolidated many of these areas over the last 30 years, leading to higher utilization by the military for training and testing. The result is ever-growing competition among agencies for use of the available airspace with the highest priority programs taking the vast majority of the available time. As test ranges become more sophisticated, the cost to use these ranges also increases. Unless a program has sufficient funding and priority, it will not be able to test on existing ranges because of limited range availability and prohibitive costs.

In recent years, several states and universities have been coordinating with the FAA to designate more areas for UAS operations. For example, the University of New Mexico, Physical Science Laboratory, has established a UAV Center of Excellence adjacent to the White Sands Missile Range for UAS flight operations. This offers alternative airspace in which to test, but the proliferation of aircraft vendors and UASs still proves a challenge.

Initially, UAS operators were private model aircraft hobbyists who flew their aircraft in very confined areas in the NAS, typically at less than 400 ft above ground level. The FAA set up guidelines for these operations in 1981 through Advisory Circular (AC) 91-57, Model Aircraft Operating Standards. These operations were intended solely for providing guidance to persons interested in flying model aircraft as a hobby or for recreational use only. This guidance encouraged good judgment on the part of operators so persons on the ground or other aircraft in-flight were not endangered. These guidelines included line-of-sight operation, operations below 400 ft above ground level, and restrictions and limitations on flight in the vicinity of spectators. While this AC was designed for model aircraft, people have misused the guidance for commercial flight operations. AC 91-57 only applies to hobbyists and recreational flyers. It excludes persons or companies operating UASs for business purposes.

The FAA policy for UAS operation is that no person may operate a UAS in the NAS without specific authority. This authority can be obtained by applying to the FAA for an Experimental Airworthiness Certification or a Certificate of Authorization (COA). The first COA for commercial UAS operations was issued in 2005.

A COA is generally based on the following principles.

- The COA authorizes an operator to use defined airspace and includes special provisions unique to each operation. For instance, a COA may include a requirement to operate only under visual flight rules and during daylight hours. Most are issued for a specified period (up to 1 year, in some cases).
- Most, if not all, COAs require coordination with an appropriate air traffic control facility and require the air vehicle to have a transponder able to operate in standard air traffic control mode with automatic altitude reporting.

- To make sure the UAS will not interfere with other aircraft, a ground observer or an accompanying "chase" aircraft must maintain visual contact with the air vehicle.
- A Notice to Airman is usually required to alert general aviation aircraft of impending UAV operations.

The FAA issued 102 COAs in Calendar Year (CY) 2006, 85 in CY 2007, and 164 in CY 2008. As FAA experience with COAs has grown, so has the emphasis on safety; certificates issued today typically have more conditions and limitations, particularly those dealing with a UAS's inability to "detect, see, and avoid" other traffic. There is no guarantee that a UAS vendor will receive a COA, and several have been denied. The time element required to get a COA approved is variable and based on the maturity of the system, FAA inspector availability, and the developer's familiarity with the COA process and requirements. Six months to obtain a COA is not uncommon.

The bottom line is that the FAA has only recently been presented with the issue of flying UASs in the civil NAS and is in the process of establishing procedures for their safe operation. Until this is completed, UAS vendors and testers will have many limits imposed upon them for the testing of their systems. So how does a company or individual test their system in an economical, timely, and efficient manner?

The obvious solution is to minimize the amount of flight testing required by a stand-alone UAS. Where feasible, flight testing on a surrogate manned aircraft is a viable option. The scope of effort can range from evaluating separate components (avionics, sensors, payloads, etc.) to fully mounting the air vehicle on a mother ship platform.

UAS configurations

UASs are composed of several different components. Although generally we think of air vehicle testing, the reality is that a UAS is designed around the mission and payload, just as a manned aircraft is. Integration and operation of the "system" is the most complex and time-consuming part of flight test. A UAS can typically consist of the following subsystems:

- air vehicle,
- command and control,
- payload/sensors,
- data link,
- weapons.

UAVs vary in size, with wingspans ranging from a few inches to larger than a commercial 737, with subcomponents sized to the air vehicle. Although many of these subcomponents may be obtained as government-furnished equipment or are commercial off-the-shelf items, the integration of these components into a UAS may require considerable testing. For example, a government-furnished equipment UAS vehicle could be modified to add an electro-optical–infrared sensor or be controlled via a UHF free-wave radio while downloading sensor information via a tactical data link. Although these systems are all off-the-shelf items, effective hardware and software integration typically includes interfaces and software modifications tailored to the UAS.

A significant portion of this integration can be conducted in a systems integration laboratory, but simulations do not always represent real world conditions. Based on the complexity of the integration and maturity of the systems integration laboratory, an integrator or developer can experience significant changes between a bench test and open-air operation. Although some testing is accomplished in a lab, flight testing remains a significant portion of a system's evaluation.

Most UAVs are used as sensor platforms. With the current conflicts in Iraq and Afghanistan, use of UASs has increased exponentially. They are often employed for remote area surveillance and intelligence gathering but are also used for surveillance of high-density population areas. To provide realistic scenarios for testing, developers should evaluate these systems in environments that range from similar rough mountainous areas to high-density population areas. Unfortunately, FAA limitations prevent testing in these environments. In other words, existing FAA policies and limitations prevent testing UASs in the environment and conditions where they are designed and intended to operate.

General aviation aircraft use regulations

General aviation aircraft have been operating in the civil NAS since the inception of flight. Rules and regulations surrounding their operation are well defined and understood for private use, commercial operation, and research purposes. General aviation aircraft can be operated with either a standard or special airworthiness certificate based on the use and configuration of the aircraft. Many aircraft have been modified over the years to accommodate special equipment or configurations. *Figures 2* and *3* show some of these modified aircraft.

The FAA oversees modifications to aircraft with their primary interest being safety. The type of airworthiness certification requested and the extent of the modification will determine the level of FAA

Figure 2. Twin Otter with nose-mounted sensor pod.

inspection required. If the modification to the aircraft structure has been previously approved through a supplemental type certificate, the certification process can be accomplished in a matter of days *(Figure 4)*. Once the aircraft is certified, it can operate throughout the civil NAS by following either visual or instrument flight rules, thus allowing the system to be flown in the environments for which they were designed.

Flying systems or subsystems on surrogate aircraft is not a new concept. The concept has been used over the years to evaluate systems for most manned platforms. A new sensor will be evaluated on a less expensive test bed aircraft to develop its maturity prior to installation on a more expensive platform for final evaluation. For example, fly-by-wire flight controls were initially installed and evaluated on a mechanical flight control aircraft before testing on a prototype aircraft *(Figure 5)*. In the world of UASs, many of the systems are directly installed in the air vehicle, and flights are flown to evaluate them under the mistaken belief that the system is mature, there are no personnel involved if the system malfunctions, and the potential for collateral damage is low. This belief has led to a high rate of UAV accidents. The military services are only now starting to include UAV accidents in their accident rate recording, and the civilian industry maintains virtually no data on UAV accidents.

Figure 4. MX-20 EO/IR turret mounted to nose of a Twin Otter.

Once the systems are initially evaluated and their integration has matured through testing, they can be cross-decked to the UAS for final evaluation. This can be accomplished either through the military, large original equipment manufacturers, or by the individual developers themselves. In the past, the use of a government platform for testing was a viable option, but as budget pressures, downsizing, and real world operational demands increase, availability of resources has become much more limited. Additionally, the cost of operating a military aircraft becomes cost prohibitive for many smaller research and development programs. For these reasons, commercial aircraft use for testing is on the increase. Many of the larger research institutes and vendors have their own research aircraft. Lincoln Laboratories, the National Oceanic and Atmospheric Administration, the National Aeronautics and Space Administration, Lockheed Martin, and Northrop Grumman are just a few of these organizations. They

Figure 3. Radome mounted to bottom of King Air and Research pod mounted out of the door of a Twin Otter.

Figure 5. C-131 Total In-Flight Simulator.

Figure 6. Cabin research stations.

typically use general aviation aircraft that they lease or own or have an aircraft loaned to them from one of the uniformed services. However, what most people don't know is there are many small companies that offer their expertise and aircraft for testing of these systems on general aviation aircraft.

Advantages of surrogate aircraft

Rather than being subject to flight limitations and airspace restrictions imposed upon UASs, a manned aircraft employing payloads is subject to the flight restrictions that are similar to that of any other general aviation aircraft. Either the operator can file under instrument flight rules, and be controlled by air traffic control for operation over sensitive areas, or they can file under visual flight rules. The flexibility of operating under visual flight rules enables aircraft to operate freely and autonomously as long as the aircraft remains outside of special use airspace or air traffic control areas. This has many advantages because the aircraft can operate as required to fully evaluate the systems. The pilots must still maintain see-and-avoid flight principles, but they can be augmented by onboard flight equipment such as traffic collision avoidance systems or by flight following with the local air traffic control service.

Test engineers and operators can "fly along" onboard the aircraft to analyze results during testing and make real-time adjustments to the systems. Many UASs are tested using a "fly-fix-fly" principle of operation wherein the software and hardware is set before a UAV flight and can only be adjusted once the UAV is recovered. On a manned research aircraft, the engineers can fly along with the "system" and accomplish adjustments during flight for those hardware and software components mounted internally in the manned aircraft *(Figure 6)*. This has proven its worth on many test flights where a software glitch is evaluated and modified in real time. This real time analysis and modification turns the airborne surrogate platform into an airborne systems integration laboratory. As an additional benefit, hardware and software can often be changed out in flight so several different configurations can be tested, evaluated, and compared.

The FAA places many restrictions on when and where UAVs can be tested. As previously noted, the process to obtain a COA is lengthy and comes with a host of operational restrictions including where and when they can be operated. Additionally, a Notice to Airman may be required anytime the UAV needs to be flown. The use of a manned surrogate aircraft avoids many of these limitations. The tester is free to fly the system when needed and in the environment desired.

All testers have experienced schedule slips measured in hours, days, or weeks, often based on the maturity of the systems being evaluated, instrumentation issues, support equipment, and weather. If operating a UAS, failing to launch within a certain window of operation may cancel a flight due to air traffic control limits placed on the UAS operation or scheduled range times. The more these restrictions can be minimized, the more efficient testing will become.

The DoD and commercial industry are learning similar lessons and are turning to general aviation aircraft as surrogate platforms for UAS and their subsystems. With the wide variety of general aviation aircraft available, testers and program managers can choose the aircraft that most closely matches the UAS's performance to obtain the most representative data, for example, a King Air for radars to be used in a Predator-size UAS or using a four-place aircraft for mounting an electro-optical sensor for urban surveillance *(Figure 7)*.

The next step is to not only mount the sensors to be tested on the general aviation aircraft, but also mount

Figure 7. Navion modified with E/O sensor.

the UAV itself. Small air vehicles can be easily mounted to a fuselage or a wing station. A small UAV (150 lb or less) could easily be mounted to the wing station of some of the more rugged general aviation aircraft. *Figure 8* shows a Twin Otter aircraft with wing stations capable of holding a 500-lb load and provides four hard points that could be used to mount a UAV.

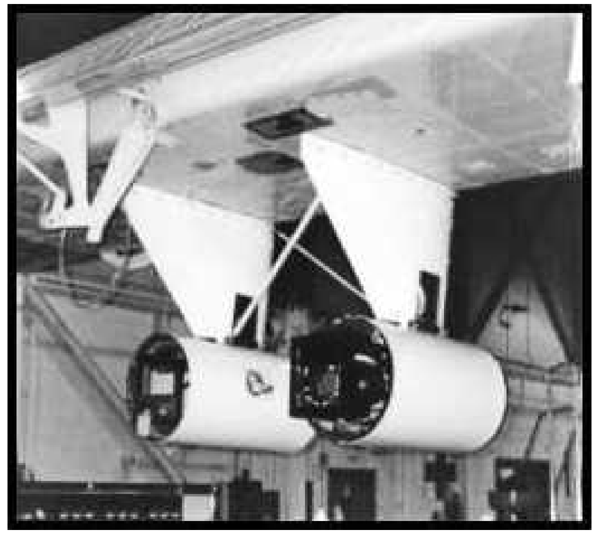


Figure 8. Twin Otter wing stores mounts.

Conclusion

General aviation aircraft are viable, cost effective options for the development and testing of UAS and their payloads or sensors. Types of UASs and their missions are growing exponentially and will continue to do so for the foreseeable future. But the policies governing their operation within the United States, specifically in the civil NAS, are only just beginning to be addressed. Developers and operators will face serious challenges in getting their systems tested and evaluated. In the past, the military could be relied upon for the testing of these systems in special use airspace, but growth in UAS use, rising costs, expanding flight envelopes, and more complex operational mission profiles have required UAVs to be tested outside of special use airspace. Also, too many UAS vendors and testers have limited themselves to using only their actual air vehicles to flight test their systems in the mistaken belief that this will result in a shorter test period. Instead, UAS vendors and testers should take advantage of the lessons learned by manned aircraft testers and use manned surrogate aircraft for their flight testing. General aviation aircraft are a viable, cost effective, and timely means of accomplishing UAS flight testing and are readily available. ❑

JOE ARVAI has over 25 years of test and evaluation experience, and retired from the Air Force in 2006 as the director of operations for the Air Force Operational Test and Evaluation Center. Mr Arvai is a graduate of the U.S. Naval Test Pilot School and has accumulated over 5500 flight hours in more than 50 varieties of aircraft. He now works as a test pilot and senior principal engineer in Lexington Park, Maryland, for Aviations Systems Engineering Company. His duties include conducting advanced technology tests for the DoD. Mr. Arvai holds a bachelor's of science degree in aeronautical engineering from The Ohio State University. E-mail: joe.arvai@asec.aero

ITEA Journal 2009; 30: 531–541

Korean-Dutch Flight Testing for KA-32 Training Simulator Development and Validation

Jasper van der Vorst, Peter J. A. Booij, J. Brugman, and Joost F. Hakkaart
National Aerospace Laboratory, NLR, Amsterdam, The Netherlands
Dae Keun Jeon, Hyoung Sik Choi, and Hyang Sig Jun
Korea Aerospace Research Institute, KARI, Daejeon, Korea

Together with (and contracted by) the Korea Aerospace Research Institute (KARI), the Dutch National Aerospace Laboratory performed a successful flight test campaign with the Kamov KA32T in South Korea during the summer of 2007. These trials were part of the KA-32 Helicopter Training Simulator Development Program managed by KARI. Within this program, National Aerospace Laboratory developed the flight model and executed the flight tests in close cooperation with KARI and the helicopter operator. A very successful flight test campaign was executed from August 1 to 31, 2007, at the Iksan airbase of the Forest Aviation Office. The installation and calibration of the instrumentation was accomplished within 2 weeks. A total of about 30 hours of flight time was performed in 22 flights. This article describes an interesting project with an international touch, including some distinctive logistical challenges: Korean and Dutch engineers working on a Russian helicopter.

Key words: Flight test program; helicopter; Iksan airbase; international collaboration; Jeonju airbase; Kamov KA32T; Korea; simulator development; test data; test plan.

The objective of the KA-32 Helicopter Training Simulator Development Program is to acquire a helicopter simulator that meets level C requirements in accordance with FAA AC 120-63. The Korea Aerospace Research Institute (KARI) managed the development program and was in charge of developing and validating the flight dynamics model based on simulator design data and flight test data. The helicopter chosen for this project was the Kamov KA32T (*Figure 1*) operated by the Korean Forest Aviation Office (FAO), mainly used for fighting forest fires.

KARI was presented with the challenge of finding sufficient data for the development of the flight dynamics model. The Netherlands' National Aerospace Laboratory (NLR) was awarded a contract to develop the flight model and gather flight test data because of its experience with flight simulation development and flight testing for a competitive price. The result was an interesting project with an international touch, including some distinctive logistical challenges: Korean and Dutch engineers working on a Russian helicopter.

The KARI/NLR project consisted of three phases: flight mechanics model development, flight testing, and model tuning. During the flight test phase, the goal was to gather data for flight mechanics model improvement and data for the comparison between model and flight test (qualification test guide). This article presents the preparations and execution of the flight test program and discusses some of its results. At the end of the article, the application of the measured flight test data within the project is discussed briefly.

Helicopter configuration

The Kamov KA32T is an 11-tonne twin engine helicopter with a coaxial rotor system.

All flights were performed with a crew of two pilots and one flight test engineer, complemented during several flights with a flight mechanic. The pilots of the test aircraft were senior pilots within the Forest Aviation Office; however, they had no formal test pilot training. The flight test engineer from KARI was in charge of the in-flight organization of the tests, managing the instrumentation system, and recording of events using event marker and flight test cards.

Figure 1. The Kamov KA32T test helicopter.

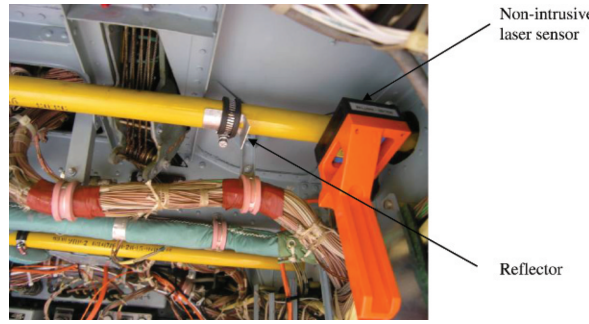


Figure 2. Nonintrusive laser sensor (on orange bracket) and reflector (on yellow push–pull rod) below cockpit floor to measure stick position.

The FAO normally operates the KA32T with a Simplex Model 10900-050 Fire Attack water tank mounted below the fuselage. Because the water tank limits the maximum speed to 150 km/h, as opposed to the normal maximum speed of 230 km/h, it was decided to perform the flight tests without the water tank to enable testing over a larger speed envelope. Both engine inlets are equipped with a dust protection device and an anti-icing system.

It was decided by KARI to vary the helicopter weight using fuel quantity only. Because the external fuel tanks are not available at FAO, only the internal tanks were used. Using this configuration, weights between about 7,300 and 8,700 kg can be achieved. Additionally, the center of gravity range was varied with the position of a flight mechanic in the cabin. The test helicopter was not equipped with an external hoist or air conditioning. The dust protection device and anti-icing system were off for all tests except for those tests measuring the performance impact of these systems.

During normal operation of the helicopter, the autopilot is on, providing rate stabilization and attitude hold. During many of the flight tests, the autopilot had to be switched on. However, some tests specified in the simulator qualification requirements (FAA, 1994) require maneuvers to be performed without the autopilot. The required configuration (autopilot on or off) was indicated on the test cards. "Autopilot" referred only to the yaw, roll, and pitch channels on the center control panel. Other modes such as altitude hold were not used during the test maneuvers.

Instrumentation system

Instrumentation system in helicopter

After several preparatory visits to South Korea, the preliminary design of the instrumentation system was started using NLR's Generic Instrumentation System (GIS) as a basis. The GIS is an advanced airborne measuring and recording system. It is capable of adequately measuring, conditioning, and recording analog signals, discrete signals, digital signals, synchro signals, and manual data entry, e.g., record number.

A constraint for the instrumentation system design from the operator was to install equipment with as little impact on the helicopter as possible, both mechanically and electrically. For both operational and safety reasons, the system had to be nonintrusive. Therefore, the approach for the design of the instrumentation system was to use as many parameters going to the KA32's Flight Data Recorder (FDR) as possible. This required the design of a "breakout box," which enabled recording these parameters by the NLR data acquisition system, while the FDR remained in operation. A second major part in the instrumentation system was a dedicated test inertial reference system, providing ring-laser–based attitudes, rates, and accelerations.

To complement the parameters from the FDR and the NLR inertial reference system, we installed several additional sensors. On the landing light bracket, a probe for outside air temperature was installed. To satisfy concerns about flight safety, we used nonintrusive optical (laser) sensors for longitudinal and lateral cyclic position, with reflectors installed on the longitudinal and lateral push–pull rods below the cockpit floor (*Figure 2*). To measure engine temperature, we installed a breakout connector in the signal from the engine thermocouples. Because it was not possible to measure the cold junction temperature, the measurement varies with cold junction temperature. This deficiency was solved by correcting the measurement with observations of the cockpit instruments from video (for ground tests) and from the flight test engineer (for flight tests). A temporary transducer, for ground test only, was connected to the engine throttles to measure the deflections during engine start-up, (ground) operation, and shutdown. The engine pressure ratio, an indication of engine power, was

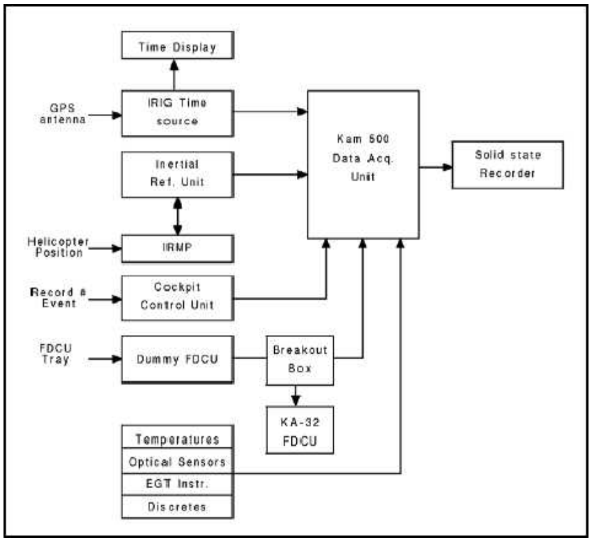


Figure 3. GIS block diagram.

measured by installing a breakout connector in the signal to the cockpit instrument.

To create the breakout connectors for the flight data recorder, and engine pressure and temperature, we had to purchase several Russian connectors, which proved to be a very critical part of the design. A video camera was used to record engine instruments during ground runs in the engine start procedure.

All flight test data were recorded on a solid state data recorder and were processed directly after the flight in the Omega data processing system to enable analysis of the data before the next day. The Omega system contains all the calibration data of the individual parameters and calculates the engineering units from the raw recorder data. The block diagram of the GIS is shown in *Figure 3*, and the system as installed in the helicopter is shown in *Figure 4*. The parameter list can be found in *Table 1*. The instrumentation design concluded with a safety analysis report, showing that the instrumentation design has a high degree of reliability and damage tolerance and that it has provisions to protect the helicopter signals in the event of a failure.

Instrumentation systems on ground

A ground station was located at the FAO base at Iksan. It consisted of a KARI portable office container in which the NLR ground station was installed. The NLR ground station is based on a WYLE Omega processing system in a server–client network environment. The server is operated by the instrumentation engineer, and processes and distributes all available data from helicopter and ground instrumentation. The Omega system contains all the calibration data of the individual parameters and calculates the engineering units from the raw recorder data. The system design allows for quick configuration changes for different test programs. A shared hard-disk unit is used for securely archiving the acquired data. The specialists were provided with client laptop computers, enabling them to analyze the distributed data on- or offline as necessary. The network is completed with a network printer.

Weather data were gathered with a mobile meteorology (meteo) system, consisting of temperature, pressure, humidity, wind speed, and direction sensors. These transducers are mounted on a transportable 10 meter high meteo mast. The system can be powered by a car. The data are logged onto a PC. The meteo system was used during several hover trials at the FAO base at Iksan and the low speed trials at Jeonju airbase.

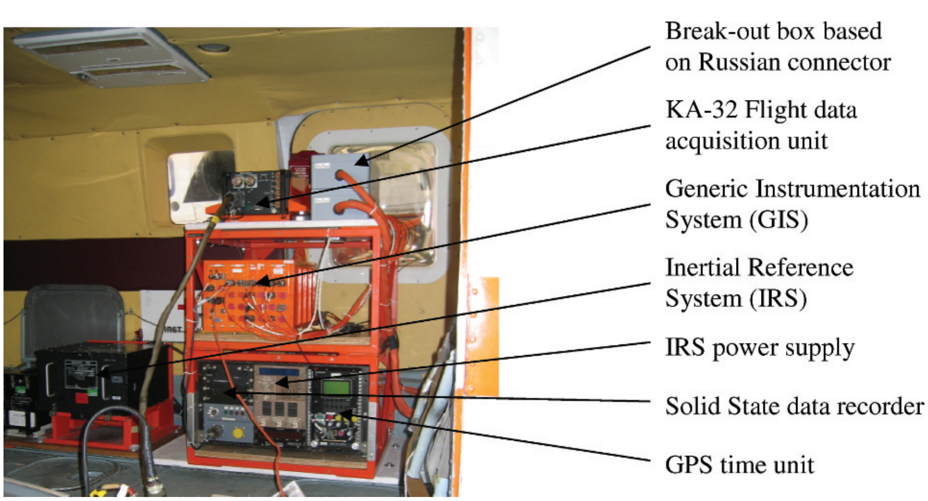


Figure 4. The ring laser gyro and measurement system in the KA32T.

Table 1. Parameter list.

ATA	Description	ATA	Description
	General		**Navigation**
0	Event Marker	34	Roll Attitude_FDR
0	Record number	34	Normal Acceleration
0	Cold Junction Temperature	34	Heading_FDR
	Calibration Tool Arms & legs	34	Lateral Velocity Doppler
0	Time	34	Longitudinal Velocity Doppler
	Air Data	34	Vertical Velocity Doppler
1	Indicated Airspeed	34-28	Pitch Angle
1	Outside Air Temp at heli	34-28	Roll Angle
1	Altitude (baralt)	34-28	Ground Track True
1	Altitude (radalt); upto 300 m	34-28	Body Longitudinal Accel.
	Meteo	34-28	Body Lateral Accel.
15	Wind Direction	34-28	Body Normal Accel.
15	Wind Speed	34-28	Vertical Acceleration
15	Air Pressure Groundstation	34-28	Ground Speed
15	OAT Groundstation	34-28	Magnetic Heading
	Flight Controls	34-28	True Heading
27	Cyclic Lateral Position_FDR	34-28	Present Position Latitude
27	Cyclic Longitudinal Position_FDR	34-28	Present Position Longitude
27	Collective Position	34-28	Body Pitch Rate
27	Cyclic Lateral Position_NLR	34-28	Body Roll Rate
27	Cyclic Longitudinal Position_FDR	34-28	Body Yaw Rate
27	Collective Position	34-28	Velocity N S IRS
27	Cyclic Lateral Position_NLR	34-28	Velocity E W IRS
27	Cyclic Longitudinal Position_NLR		**Engine**
27	Differential Pitch	72	Engine Pressure Ratio 1
	Pedal Position	72	Engine Pressure Ratio 2
	Collective Pitch	72	Gas Generator Speed Engine 1
	Trim button on pilot Cyclic Stick	72	Gas Generator Speed Engine 2
	Landing Gear	72	Rotor Speed
32	Weight-on-wheel signal	72	Total fuel quantity
	Navigation	72	Separate Throttle Control Lever
34	Lateral Acceleration	72	Turbine Gas Temperature Engine 1
34	Longitudinal Acceleration	72	Turbine Gas Temperature Engine 2
34	Pitch Attitude_FDR		

Installation and calibration activities

The flight test campaign in the summer of 2007 started with the installation and calibration of the instrumentation system. Because most of the design work was performed in the Netherlands, some minor adjustments had to be made in Korea to the mechanical interface. After the instrumentation installation, the parameter calibration began. As far as possible, parameters were calibrated on the ground. For example: the fuel gauge was calibrated through a weight and balance procedure at several fuel weights; the airspeed and pressure altitude were calibrated with a pitot-static test set, and the flight control rigging was checked through a ground test with hydraulic power. Other parameters, like the engine temperatures, gas generator speeds, and rotor speed, were calibrated during a ground run. The engine pressure parameters (substitute for engine torque) could only be calibrated in flight.

After the first ground runs for a general instrumentation check and electromagnetic interference and

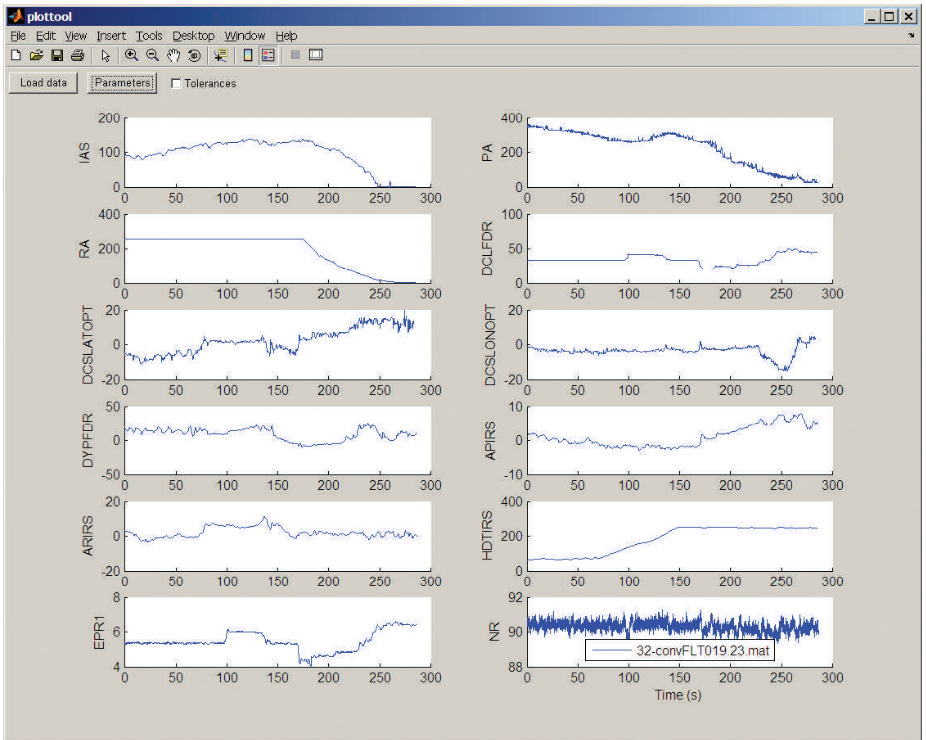


Figure 5. Flight test data plotting tool.

electromagnetic compatibility test, a first test flight took place for instrumentation check and final calibration. Several runs were included to determine the error in the pitot-static system. The activities just described were performed in a 2-week period, ending on July 31, 2007.

Data processing and analysis

Data from the instrumentation system are processed directly after flight and, after calibration in the Omega data processing station, converted to Matlab data files. Several tools have been developed for quick post-processing and analysis of the flight test results.

- A Matlab-based graphical user interface (*Figure 5*) for fast presentation of flight test data. This tool can represent both steady state data (average values and standard deviations) as well as time history data (parameters as a function of time). The appropriate parameters are displayed, depending on the type of test. Additional parameters can easily be added manually. A provision has been made to show AC120-63 tolerances.
- A Matlab-based graphical user interface (*Figure 6*) for the selection of steady state (trim) data. From time history data, selections can be made manually, automatically showing the average value and standard deviation.
- Flight test replay tool: HeliX is a three-dimensional representation of flight path and helicopter motion (*Figure 7)* from either an outside view or a cockpit view with head-up display, including stick positions, enabling the replay of test data. This was found to be a highly valued aid in postflight data analysis.

Flight test plan

The KARI flight test engineer was responsible for the onboard flight test management, briefing, and debriefing, while NLR engineers were responsible for test planning, data processing, and analysis. FAO pilots and mechanics were in charge of safety for the flight and instrumentation. In preparation for the flight test campaign, the test plan was drafted as well as a flight test execution guide. The test plan described in detail which configuration and maneuvers were planned, while the flight test execution guide provided guidelines to the pilots on how to perform the maneuvers. Because of FAO operational limitations, no autorotation or (simulated) single engine flights could be performed.

The majority of the test plan consisted of AC120-63 validation tests. Additional tests were included in the test plan for validation outside the AC120-63 requirements, like accel–decel maneuvers and hover

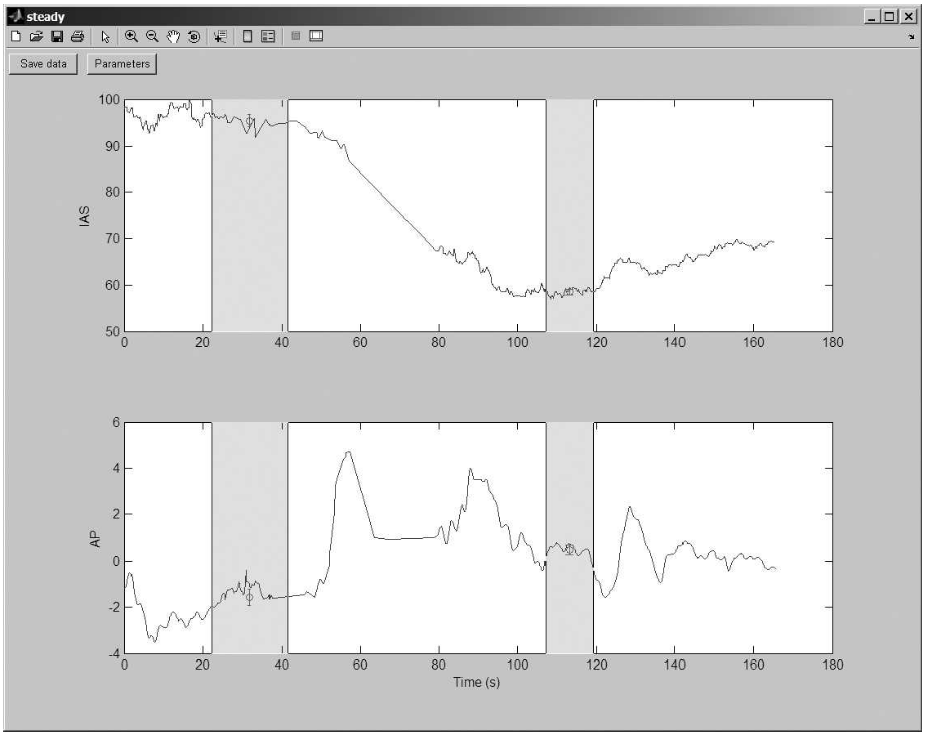


Figure 6. Steady-state data selection tool.

turns. Other tests, like autopilot and engine performance checks, were performed to provide additional data for the simulation model. A total of 143 test cards were prepared, distributed as follows:

- 14% ground,
- 19% hover,
- 8% low speed,
- 5% climb & descent,
- 54% cruise.

The test plan was summarized in an Excel sheet (*Figure 8*), which was the main flight test planning tool. It provides a quick overview of progress and includes test priority and pass/fail indication. Also, from this sheet, test cards are generated automatically,

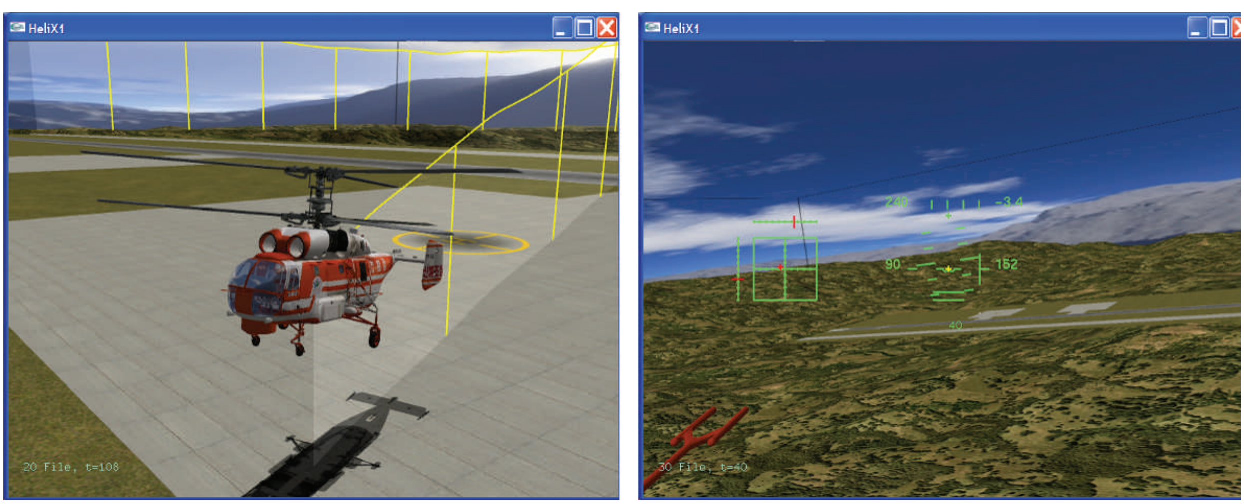


Figure 7. HeliX flight test replay tool.

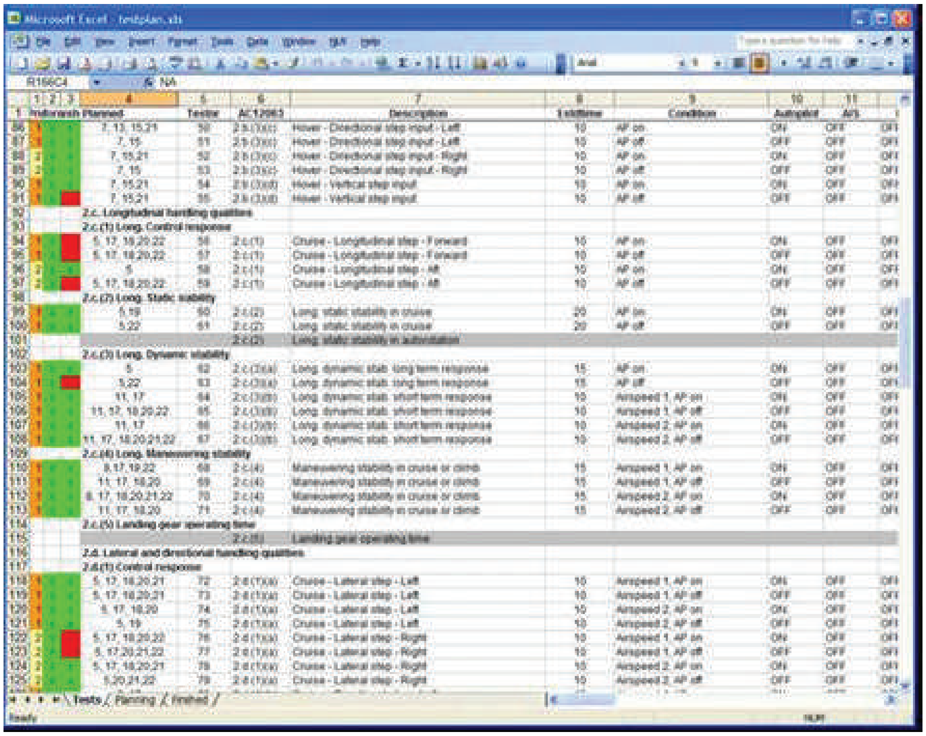


Figure 8. Excel sheet for flight test planning.

21	Level flight performance	KARI - KA-32 Flight Tests FP 619

Configuration	
Autopilot	OFF
AIS	OFF
DPD	OFF
FE	seat 15
Fuel	1800 to 2200 l

Condition	
Date	20070525
Time	
Flight condition	cruise
Target altitude	500 m
Initial IAS	80 km/h
Wind	

- Maintain level flight (altitude, heading, airspeed)
- Accelerate to target airspeed
- Stabilize for 10 seconds
- Record for 20 seconds

Check configuration and condition		
Start recorder – Rec. no.:		

Target airspeed	Check & Comments
80	☐
95	☐
110	☐
125	☐
140	☐
155	☐
170	☐
Fill in kneepad items + see next page	

Figure 9. Example of a test card.

including a short description on how to perform the test, required configuration for the test, and room for remarks by the flight test engineer (*Figure 9*).

Flight test execution

After a 2-week instrumentation installation period, the test campaign started at the Iksan airbase of the Forest Aviation Office on August 1, 2007. Nearly 5 weeks of flight tests followed. The flight tests were performed in a daily schedule of up to two flights a day. After acquiring the actual meteo information, the test cards were selected for each flight based on weather conditions; progress of the test program based on analyzed test results; and an efficient combination of maneuvers with respect to helicopter mass, required altitude and airspeed, pilot's workload, etc.

The test program consisted of the sequence of the selected test cards. The resulting program was briefed to the KARI flight test engineer by NLR in English. Subsequently, the helicopter crew was briefed by the flight test engineer in Korean. Next, the helicopter executed the test flight. During the test flight, previously acquired data were analyzed by NLR on the ground. The main objective of the analysis was approval or rejection of the data as a source for tuning. The approval of data defined the status and progress of the test program. After landing, the acquired data were processed by the NLR instrumentation engineer while

Figure 10. Installation site of meteo mast at Jeonju Air Force Base.

the other NLR engineers were debriefed by the KARI flight test engineer.

Low speed flight tests

Because the FAO base at Iksan has only a helicopter platform, the low-speed flight tests requiring a runway were performed at the Jeonju Air Force Base, which is only 4.5 nautical miles from the FAO base. For these tests a mobile meteo team deployed to Jeonju air force base to set up the 10-m wind measuring mast just outside the base perimeter for security reasons, in close proximity of the runway (*Figure 10*). This team operated from a car with a power supply, laptop, and data acquisition system connected to the measuring mast.

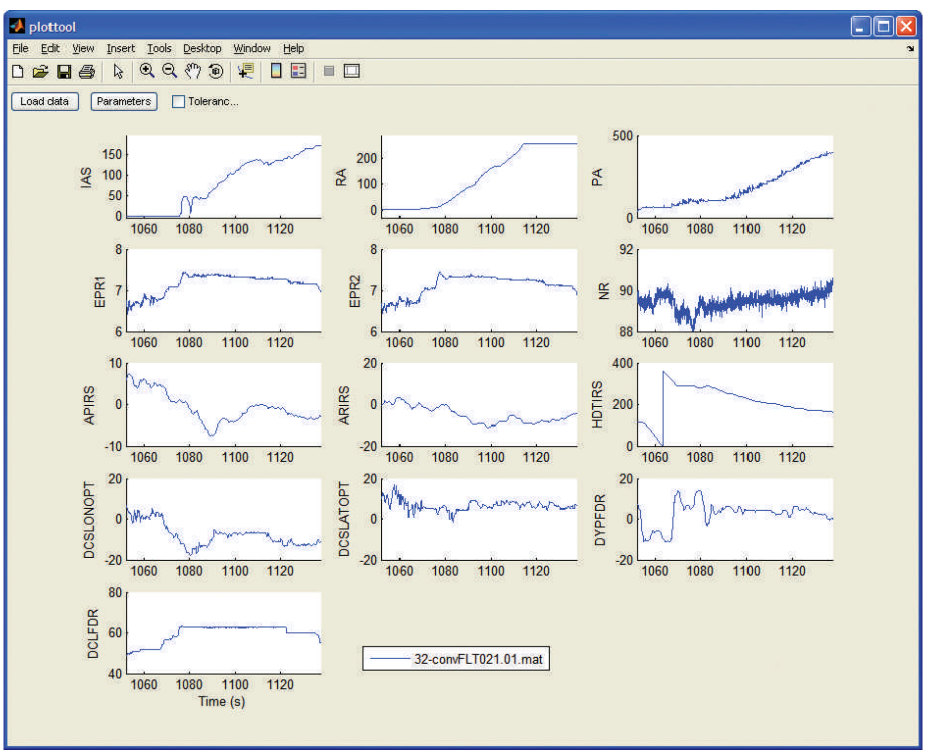


Figure 11. Example of parameter plots used during analysis of the acquired test data.

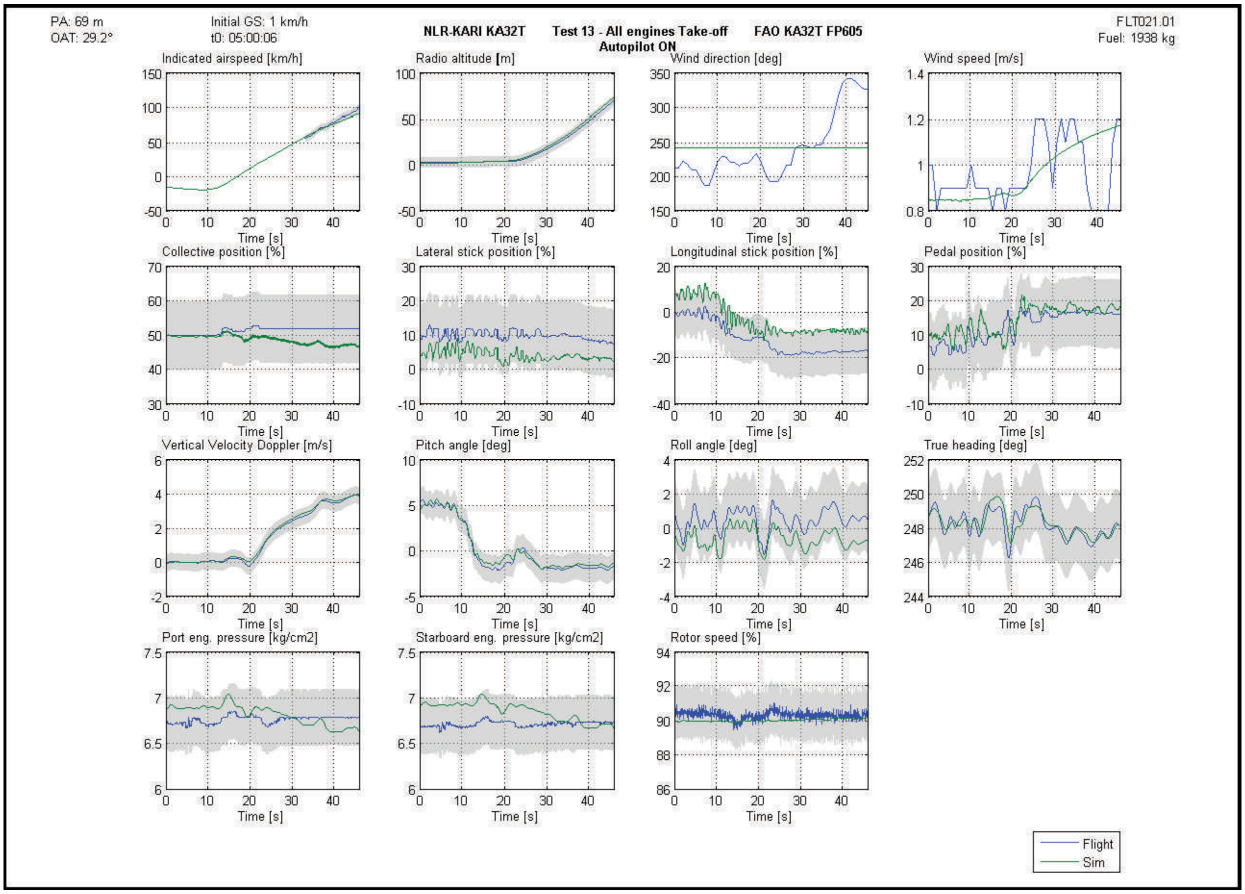


Figure 12. Comparison between model and flight test data for the takeoff maneuver.

Flight test results

In the period from August 1 to 31, 2007, the flight trials at the Iksan airbase of the Forest Aviation Office yielded the following results:

- A total of about 30 hours of flight time has been performed in 22 flights;
- A distinction was made between "performed" tests and "approved" tests: A test was performed once it has been executed during a flight. Only when the data of the test show that the test has been executed satisfactorily and provides sufficient data for model tuning, were the data approved;
- Of the test program, 99% was executed. Of the planned tests, only the engine start and shutdown at altitude was not performed (low priority).

The Matlab-based analysis tools, described earlier, were used to analyze the acquired test data. The analysis consisted of verifications of data quality; steady initial conditions; and steady data, i.e., correctly performed maneuver; the analysis also verified that control inputs applied conformed to the definition required for tuning.

In *Figure 11* an example is given of a graphical presentation of a takeoff. To save space on the screen, we list only the acronyms without engineering units on the vertical axes.

Application of flight test results

During the flight test phase as described in the previous section, the goal was to gather data for flight mechanics model improvement and data for the comparison between model and flight test (qualification test guide). This section presents a brief discussion of how the flight test data were used within the project. The complete results of model development and subsequent tuning process are presented in van der Vorst et al. (2009).

Before starting the tuning phase, the flight mechanics model was updated with data measured during the flight test phase. This included airspeed calibration, flight control rigging, engine performance data, and autopilot performance (gains and limits). The tuning process consisted of an iterative loop. Together with postprocessing the flight test data, an appropriate selection of the flight test data was made, for example, selection of the most successful control inputs or best steady data.

These data were input for the creation of scripts that enabled automatic simulation of all test points. The subsequent data analysis led to changes in the model, or changes in data selection, after which another iteration was performed. An example of the result of the tuning phase is shown in *Figure 12*, the all-engines takeoff. The green and blue lines present the simulation and flight test results, respectively, and the shaded area indicates the tolerance defined in the simulator qualification requirements (FAA, 1994).

During the tuning phase a number of challenges were encountered because of limitations in instrumentation and allowable flight test maneuvers (no single engine or autorotation), lack of wind tunnel data (only computational fluid dynamics), etc. Despite these limitations a very good result was achieved, providing a simulation model that has a high (Level C) fidelity in representing the KA32T and an almost 100% fit to the flight test data.

Conclusions

A very successful flight test campaign was executed from August 1 to 31, 2007 at the Iksan airbase of the Forest Aviation Office in Korea. The installation and calibration of the instrumentation was accomplished within 2 weeks. A total of about 30 hours of flight time was performed in 22 flights. The efficient and flexible setup of the NLR flight testing tools enabled a small test team to quickly analyze the acquired data on-site, resulting in efficient monitoring of the program progress and flexible adaptation of the test program to ambient weather conditions and operational constraints. The flight test campaign provided good quality data for the AC120-63 tuning process, thanks to good cooperation between Korean and Dutch engineers and the Korean helicopter operator. ❑

JASPER VAN DER VORST graduated from Delft University of Technology in 1998, after which he joined NLR as a helicopter specialist. After a 1-year assignment to Eurocopter France on the NH90 flight test team, he returned to the Helicopter Department of NLR as a research engineer, involved with both helicopter flight testing and simulation. He has been involved in several helicopter-ship test campaigns for the Dutch Navy. He is NLR's main source for the simulation software FLIGHTLAB, which was used to set up NLR's Helicopter Pilot Station. Jasper holds a private helicopter pilot license. E-mail: vorst@nlr.nl

PETER BOOIJ graduated from Delft University of Technology in 1988 and joined NLR as a helicopter specialist. He is currently employed as a helicopter flight test engineer at the Department of Helicopters and Aeroacoustics. Since 1991 he has been involved in helicopter ship qualification trials. Recently, he was project leader of the qualification flight trials program with the SH-14D Lynx helicopter on board the new Dutch Landing Platform Dock. He was involved in the flight testing activities with the Korean KA32 helicopter. He is also an active private pilot. E-mail: booy@nlr.nl

HANS BRUGMAN was born August 15, 1956. After obtaining his bachelor degree, he joined NLR in 1980. Since then he has contributed to the development and operation of many flight test instrumentation systems for Fokker aircraft, Royal Netherlands Air Force aircraft, and self-steering parafoils reconnaissance systems. He is also experienced in video techniques, human factors measuring instrumentation, and telemetry systems. E-mail: brugman@nlr.nl

JOOST HAKKAART graduated in 1992 from Delft University of Technology and joined the Dutch National Aerospace Laboratory, NLR, as project manager at the low speed wind tunnel, especially for helicopter development testing. He changed to the Helicopter Department for helicopter testing in general. In 1999 he became the head of Wind Tunnel Projects for the NLR/DNW wind tunnels in Amsterdam. After 5 years, he changed again to the Helicopter and Aeroacoustics department in the position of principal project manager. Currently, he is supervising major (tilt) rotorcraft projects, including flight testing and wind tunnel testing. E-mail:hakkaart@nlr.nl

DAE KEUN JEON graduated at Seoul National University in Korea and was granted a bachelor of science degree and master of science degree in aerospace engineering in 1993 and 1995, respectively. He was involved in the development of the T-50 supersonic aircraft at Samsung Aerospace Industries and Korea Aerospace Industries for five years. Since then he has joined several flight simulator development projects for fixed wing aircrafts including T-50 anf F-4 at Dodaam Systems. The KA-32 Simulator Project was his project at Korea Aerospace Research Institute where he has been working since 2005. Currently he is an engineer who develops surveillance data processing systems for air traffic control. E-mail: bigroot@kari.re.kr

HYOUNG-SIK CHOI earned a bachelor of science degree in aerospace engineering from Ulsan University, Ulsan, Korea in 2000. He also earned a master of science degree in aerospace engineering from Ulsan University, Ulsan, Korea, in 2002. His thesis addressed the design of autolanding guidance and control systems using model inversion. He is presently a doctor of philosophy student in aerospace engineering at KAIST, Daejeon, Korea, and a senior research engineer for the Korea Aerospace Research Institute, Daejeon, Korea. E-mail: chs@kari.re.kr

HYANG SIG JUN received bachelor of science and master of

science degrees in electrical engineering from Pusan National University in Pusan, Korea, in 1988 and 1992, respectively. From 1991 to 1999, he was a researcher at Daewoo Heavy Industries Ltd. From 2000 to 2003, he was a senior researcher at Korea Aerospace Industries Ltd. In 2004, he joined Korea Aerospace Research Institute as a senior researcher. He has been involved in several CNS/ATM and aircraft simulator programs and has contributed to the development of helicopter simulator programs as a project manager. E-mail: hsjun@kari.re.kr

References

FAA (Federal Aviation Administration). 1994. FAA Advisory Circular, "Helicopter Simulator Qualification," FAA AC 120-63, October.

van der Vorst, J., K. D. S. Zeilstra, D. K. Jeon, H. S. Choi, and H. S. Jun, 2009. "Flight mechanics model development for a KA32 training simulator." National Aerospace Laboratory NLR, Korea Aerospace Research Institute, KARI. Presented at European Rotocraft Forum (ERF)-35, September.

ITEA Journal 2009; 30: 542–548

iNET Deployment at Pax River: Identifying and Mitigating the Disruptions

Brian Anderson and Daniel Skelley
Naval Air Systems Command (NAVAIR), Patuxent River, Maryland
Raymond Faulstich
Computer Science Corporation, Lexington Park, Maryland

Recognizing the inability of current telemetry technology to meet emergent needs within the Major Range and Test Facility Base, the Central Test and Evaluation Investment Program launched the integrated Network Enhanced Telemetry (iNET) Project. iNET is taking a systems engineering approach to defining a new architecture for flight test telemetry. The iNET architecture is the first major change to the underlying architecture of flight test telemetry in over 50 years! Changing an architecture that has been in place this long could have unforeseen impacts. Across our ranges, processes, procedures, and systems have characteristics of the traditional telemetry architecture inherent in their design. A careful, defined, and disciplined process is required for assuring the processes, procedures, and systems are ready to accept iNET technology. As one of the initial deployment sites for iNET, NAVAIR Pax River has conducted a continuous process improvement project to study the potential disruptions and mitigations of deploying this revolutionary and potentially disruptive technology. This article describes the study process, the potential disruptions identified, the results of the risk/failure mode effect analysis, and useful end products developed to facilitate the safe deployment of iNET at the Naval Air Warfare Center Aircraft Division at Patuxent River, MD.

Key words: AIRSpeed tool set; continuous process improvement; data flow disruption; flight test; ITD3 process; Lean Six Sigma tool set; network of networks; real-time data telemetry.

Real-time telemetry is an integral component of flight test scenarios executed on Department of Defense (DoD) Major Range and Test Facility Base (MRTFB) ranges. For the last 50 years, virtually all real-time telemetry has been point-to-point one-way transmission of data. Referred to as serial streaming telemetry (SST), data are transmitted one way, from the test article to the remotely located test team. The test team evaluates the data in real-time to ensure safe test execution and to monitor the performance of the test article. Data content and format of the SST data stream are fixed in advance of the test.

As the complexity of weapon systems increased, the amount of data collected onboard test articles began to spiral upwards. However, the rigidity of point-to-point telemetry, coupled with limited spectral resources, severely limited the amount of data that could be transmitted. Increasingly, most of the data collected onboard a test article were recorded vice being transmitted. In some cases less than three percent of the data being collected are transmitted. Test engineers have to wait until the test article returns to base to retrieve the data recorder so that the majority of the data can be downloaded and analyzed. Limited real-time access to all the data being collected has negatively affected the cost and schedule associated with flight test.

Considering the implications of this trend, it became clear that the traditional SST architecture needed to change. The million dollar question was: "Change to what?" The call for change was led by a rogue group of telemetry engineers who advocated scrapping SST and changing to a technology based on wireless networks. However, wholesale replacement of SST telemetry with wireless network technology was also problematic. While telemetry networks, via two-way connectivity, could allow near real-time access to all the data recorded onboard the test article, they were ill suited for time critical, no latency variance delivery of critical safety data. Despite its faults, SST excels at the delivery of time critical, no latency variance data. Slowly emerged the realization that telemetry networks and traditional SST were not mutually exclusive. In fact, they complement each other; each doing well, what the other does poorly. The use of wireless network technology to enhance traditional SST is the basis of the iNET architecture.

The iNET architecture describes how a collection of networks works in harmony with SST to meet emergent needs within the MRTFB. This system can be thought of as a "network of networks" consisting of the following:

- a vehicular network (to be developed by iNET) on the test article that handles all onboard data acquisition functions;
- a wireless network (to be developed by iNET) that provides network communications between test articles and between the test articles and the range infrastructure;
- the existing SST links; and
- the existing network infrastructure on DoD Test and Evaluation (T&E) ranges.

The test and telemetry communities realized that the capabilities enabled by this architecture would have far-reaching consequences. Not only would emergent requirements be more easily satisfied, but *T&E would be revolutionized*. However, safely introducing a revolutionary capability into a high-risk environment, like flight test, could prove problematic. A disciplined process was needed to assure that the deployment of iNET did not negatively affect cost, schedule, or the safety of executing flight test.

Independently, during the same time period, the Department of the Navy was developing a disciplined set of tools to improve productivity. In 1999, the Navy Depots adopted the use of AIRSpeed initiatives to increase productivity through process improvement by using a common set of industry-proven tools. The use of AIRSpeed/Lean, Six Sigma (LSS) tool sets has proven to reduce lead times, remove waste (non-value-added cost), and reduce variation.

In 2004, NAVAIR, recognizing the positive impacts that the Depots were experiencing using AIRSpeed, endorsed the utilization of NAVAIR AIRSpeed initiatives across the Naval Air Warfare Centers (NAWC). Recognizing that the AIRSpeed tool set could be used to design (as well as improve) a process, the "iNET Deployment Process" was launched. The goal of the project was to utilize the AIRSpeed/LSS tool sets to solve the complex problem of deploying iNET technologies without disrupting existing infrastructures, processes, and procedures. This project became a Design for Lean Six Sigma effort, since a process for iNET Deployment did not previously exist.

The INET AIRSpeed Project: "Chartering the Team"

In the fall of 2007, the iNET Chief Architect developed a charter for the NAVAIR AIRSpeed/ Continuous Process Improvement (CPI) Team, with the goal to define a process/concept of operations (CONOPS) for safe deployment of iNET capabilities *(Figure 1)*.

In response to the charter, the "iNET Deployment Process" Team was established to design a process for safe deployment of iNET at NAWC Aircraft Division (NAWCAD). The Team consisted of subject-matter experts (SMEs) representing all aspects of flight test and range operations. This talented group of SMEs possesses a breadth of experience including

- flight test,
- airborne instrumentation,
- range communications,
- data processing and display,
- radio-frequency (rf) systems,
- systems safety,
- time space position information, and
- risk management.

The SME knowledge base was used to analyze the planned deployments of iNET capabilities for potential disruptions leading to negative impacts to cost, schedule, and safety of flight testing at Patuxent River, MD.

Project execution

Recognizing the potential of iNET technology to revolutionize flight test at NAWCAD, this team was extremely motivated and dedicated to the task. In addition to meeting weekly for 9 months, they met in several all-day off-site meetings. Using the Lean Six Sigma tool set, the Team methodically progressed through the five design steps for Lean Six Sigma: Define, Measure, Explore, Develop, Implement (DMEDI). While working through the details of the process, the Team never lost its focus—to create a process for the safe deployment of iNET technology at NAWCAD Patuxent River.

Define

In this phase, the goal is to clearly define the problem. The first step was to validate the project charter. Once the charter was validated, the Team used

Project Name: **INET Deployment process** **Date:** 11-14-07

Competency/PEO: 5.2 National Ranges **Deployment Champion:** Kathy Seals

Project Sponsor: Dan Skelley 5.2.0 **Black/Green Belt:** Brian Anderson

Business Impact ($)

Telemetry systems have not changed in roughly 50 years. The current systems are very static. Many of the processes that govern flight testing have this static nature embedded. INET deployment offers 2 way dynamic connectivity for interfacing with data on the fly however, it may be a disruptive technology that could have unforeseen safety of flight test impacts, resulting in (x) $ loss in program operating cost across the NAE.

Opportunity or Problem Statement

When: Current and ongoing for the past 50 years, the basic telemetry structure has not changed.

- iNET may yield disruption in tech demo deployment FY09 timeframe.
- iNET may yield disruption in the FY11/12 Block (1) deployment timeframe.

What: A process for Safe deployment of iNET for flight testing is required.

Where: Initial Operation Capability (IOC) - Patuxent River Md. & Edwards AFB

- Independent substantiations PAX & Edwards

Extent: 100% percent of the time there is not a documented CONOPS process for deployment of potentially disruptive INET technologies.

- A CONOPS for INET utilization could potentially prevent loss of Test Article or loss of Life.

Goal Statement

Define a Process / CONOPS for Safe Deployment of iNET Capabilities.

Project Scope

In Scope: All Flight Test Processes relevant to Block (1) deployment of iNET.

Out of scope: Non-flight test processes, Block (2) deployment of iNET and beyond.

Project Plan

Team Launch – October 30, 2007
Define – November 19, 2007
Measure – March 4, 2008
Explore –March 04, 2008
Develop – May 14, 2008
Implement – June 25, 2008
Validate – April 29, 2009

Team Selection

1. **Mark Smedley**– (Airborne Instrumentation) SME / Project Green Belt
2. **Sid Jones** – (iNET / RCC) SME
3. **Ray Faulstich** – (iNET) SME
4. **Jonathon Norton /Dennis Normyle / Bob Myers** - (RTPS Telemetry) SME
5. **Jim Pilkerton** /Bob Craft – (RTPS Test Communications) SME
6. **Patricia Khatiblou** /Tom McCaughey – (RTPS Test Management / Control Room) SME
7. **Robert Jacob** / **Jason Stewart** – (Range Safety) SME
8. **Paul Conigliaro** / **Jackie White** – (5.1 Flight Test Engineering) SME
9. **External** – 5.1 & 5.2 West Coast SME's

Figure 1. NAVAIR AIRSpeed Lean Six Sigma Project Charter: iNET Deployment Process.

tools such as swim lane process mapping and voice of the customer/business evaluations to ensure that the problem was understood. Ultimately, the Team thoroughly reviewed the iNET deployment schedule, future iNET capabilities and scenarios, and scenarios relevant to deployments at Patuxent River.

Measure

In the *measure* phase, the Team focused on creating a project schedule, determining the project execution strategy by using a system engineering "V" diagram, and identifying the metrics to decompose. The metrics of choice were 53 preexisting process maps referred to as Operational Sequence Diagrams (OSD). The OSDs were generated by the iNET project during the initial system engineering effort to define the architecture. They are diagrams depicting the flow of events necessary to accomplish each of the 53 use cases the iNET project is chartered to meet.

From these 53, the Team identified the predominant capability to be first used at NAWCAD. The "Fetch Data" scenario describes the details perceived as needed to access previously recorded data from the test article in real-time. The Team then began a 10-week brainstorming session, meeting 3 hours per week, to surface potential disruptions associated with deploying this number-one capability. Using the OSD for "Fetch Data From Aircraft," the Team walked through the multiple process steps asking questions along the way, such as, "If this step were deployed now, what would be the impacts and or disruptions to existing processes, procedures, and infrastructure?"

From this lengthy brainstorming process, the Team categorized 144 disruptions in a Failure Modes and Effects Analysis (FMEA) spreadsheet.

Explore

While this method for analyzing the impacts of deploying a capability was successful, it was extremely labor intensive. As such, it was unsuitable for use as a standard process in deploying iNET. The Team determined that they needed a standard process, defined by a guidebook, which would steer the range in finding and mitigating the potential disruptions in the deployment of iNET. During this phase, the concept of the Disruption Finder Guidebook was established. However, the process this guidebook would describe was still a mystery.

Develop

In the *develop* phase, the Team looked at many different approaches to identifying the potential disruptions. The lengthy brainstorming session that led to the 144 potential disruptions was simply not efficient enough. A particularly creative member of the Team suggested that a process used by industry to uncover threats (or disruptions) to the deployment of complex software systems might be of use. After a bit of research, it was determined that a concept used by Microsoft Corporation, called "Threat Modeling," could be adapted to speed the process of identifying potential disruptions. Central to this concept is the use of data-flow diagrams (similar to the OSDs mentioned above) and an acronym to focus brainstorming disruptions. In the case of Microsoft Corporation, they use the STRIDE acronym which stands for Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege.

The first step is to develop a data flow diagram for the capability being deployed. The Team found that a data-flow diagram could quickly be created using the OSDs as a launching point.

Once a data-flow diagram is documented, the next step is to use the categories identified in the acronym (in the Microsoft case, STRIDE) to guide the Team through a shortened brainstorming exercise to identify the potential disruptions. However, while the STRIDE set of threats might work well for the deployment of a complex software system like Microsoft Vista, it was ill suited for analysis of deploying iNET technology. The Team developed a more appropriate threat description acronym: ITD3. ITD3 stands for Information assurance, Test conduct, Data quality, Data delivery, Displays & human interface. This acronym was derived from the use of affinity analysis to categorize the 144 FMEA disruptions discovered during the 10-week brainstorming session. The Disruption Finder Guidebook outlines the details of each category.

Armed with the ITD3 process, the Team went back to the "Fetch Data From Aircraft" scenario. Using the OSD as a starting point, the data-flow diagram *(Figure 2)* was created.

The data-flow diagram was analyzed for potential disruptions using the ITD3 acronym. In addition to the previously identified 144 potential disruptions, this process uncovered an additional 61. And, this process only took a few days vice 10 weeks! These additional disruptions were also documented in an FMEA. The Team felt they had a winner. The process was not only much quicker, but also more thorough.

Implement

In the *implement* phase, the Team continued to vet the new process. The iNET Disruption Finder Guidebook was perfected. New threat modeling concepts and lessons learned using the threat models were incorporated. In addition, the process was tested against the next three scenarios. A screen shot of the data-flow diagram for one of these scenarios ("Provide Lossless Telemetry Com-

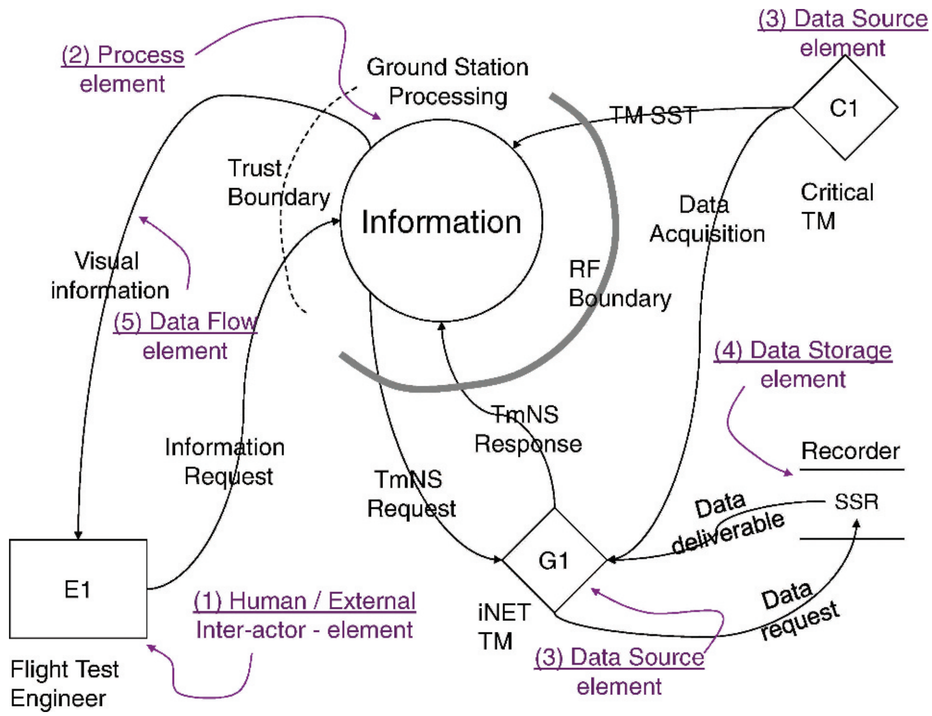


Figure 2. Data-Flow Diagram "Fetch Data From Aircraft."

munications and Detect Telemetry Dropouts") is shown in *Figure 3*. Applying the process to the three additional scenarios yielded 51 more potential disruptions that required mitigation prior to deploying iNET. Upon completion of the threat modeling, the Team had a guidebook vetted through three scenarios, multiple competency SMEs and process owners, yielding potential plans for implementing the top three super-scenarios at NAWCAD.

The Team also designed a process map *(Figure 4)* detailing the new iNET Deployment process for users in the future.

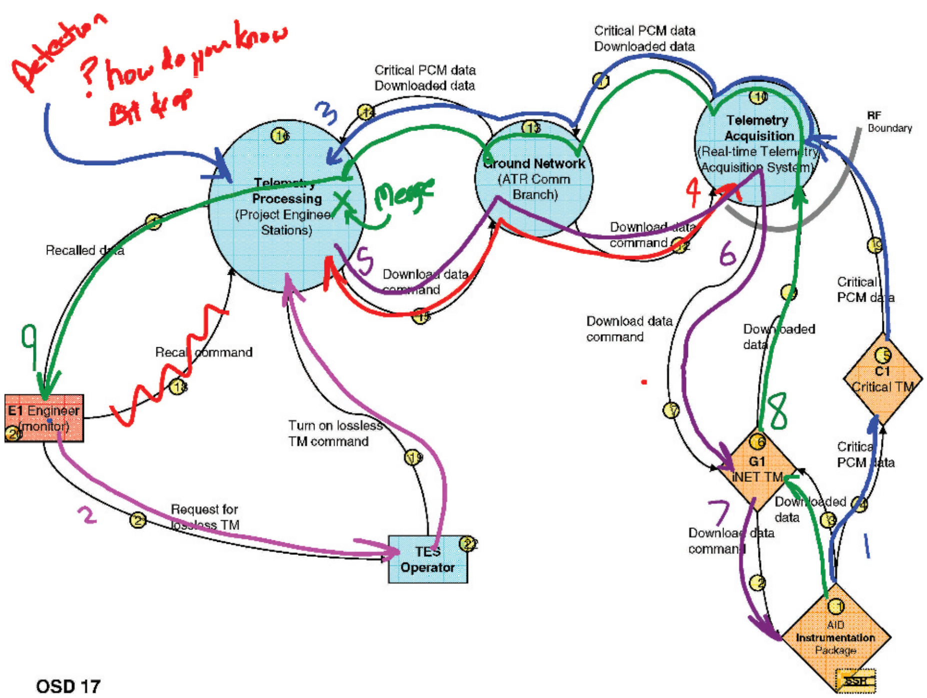


Figure 3. "Provide Lossless Telemetry Communications (aircraft)"/"Detect Telemetry Drop-outs."

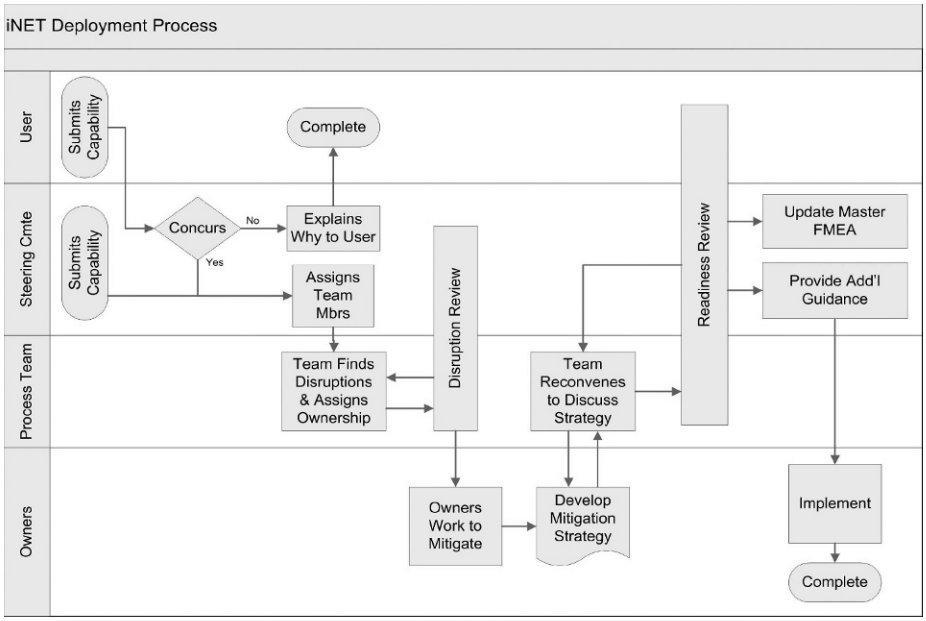


Figure 4. "iNET Deployment Process Map."

Recommendations

Once the AIRSpeed Team had a solid process for deploying iNET at NAWCAD, the question of process administration was addressed. The Team felt that an empowered iNET Steering Committee was required. The iNET Steering Committee would have responsibility for the execution of the process described in the iNET Disruption Finder Guidebook. The Team recommended that the iNET Steering Committee be created and empowered by a Memorandum of Understanding (MOU) signed at the highest levels of the organization (Senior Executive Service (SES) Level leadership of the NAVAIR Ranges, Flight Test, and Laboratories). To aid the functioning of the iNET Steering Committee, the Team created a deployment CONOPS. The iNET Steering Committee is to meet regularly with duties that include the following:

- having overall responsibility to prepare NAWCAD for the deployment of iNET Technology;
- providing guidance to users of the iNET Disruption Finder Guidebook;
- creating ad hoc SME teams to brainstorm disruptions;
- assigning ownership and responsibility for mitigation of disruptions;
- tracking the mitigation status of all potential disruptions to the deployment of iNET.

Technology

At the conclusion of the project, the Team delivered the following solution package to the project sponsor:

1. iNET Disruption Finder Guidebook—*(a detailed guide for creating and using threat models);*
2. iNET CONOPS document—*(how the guidebook and committee work);*
3. iNET Master Disruption List *(FMEA)—(containing 256 disruptions with failure modes and effects);*
4. MOU—*(an MOU binding cross-competencies to participate on the iNET AIR 5.0 Steering Committee);*
5. iNET AIR 5.0 Steering Committee Charter;
6. recommendation for iNET Steering Committee first agenda & schedule; and
7. recommendation for SMEs by name for Steering Committee participation.

Implementation status

The project sponsor is actively pursuing full implementation of all the recommendations from the iNET AIRSpeed Team. The SES Level leadership of the NAVAIR Ranges, Flight Test, and Laboratories have agreed to the MOU. ❑

Brian Anderson is a program analyst for NAVAIR Ranges at Patuxent River, Maryland. He is a former aircrewman with the United States Navy and has logged many hours in rotary wing aircraft as a United States Naval rescue swimmer. Mr. Anderson earned his degree in liberal arts through Navy Campus and is a graduate of the Naval Leadership Development Program. He is also a NAVAIR Certified Lean Six Sigma AIRSPEED Black Belt. His awards include the "Boeing-Vertol Rescue Award," the Patuxent River "Failure Controller Award,"

and the Patuxent River "Citation for Safety." Mr. Anderson has 27.5 years of federal government service, primarily dedicated to designing, developing, and installing airborne instrumentation systems. He is currently working with the NAVAIR Ranges Air Vehicle Modification and Instrumentation Department (AVMI) managing aircraft prototyping programs. E-mail: brian.anderson@navy.mil

DANIEL S. SKELLEY led the iNET Study and is currently the chief architect of the iNET Development Project. He has worked in telemetry for more than 25 years. He has a bachelor of science in engineering degree from the University of Central Florida and a master of science in electrical engineering in communications theory degree from George Washington University. Dan has held positions ranging from instrumentation engineer to the head of Aircraft Instrumentation for the U.S. Navy. He has served as the chairman of the Telemetry Group of the Range Commanders Council and is a NAVAIR associate fellow. In addition, Dan served as the 2000 technical chairman for the International Telemetering Conference. E-mail: daniel.skelley@navy.mil

RAYMOND FAULSTICH is a senior engineer and technical telemetry consultant with the Computer Sciences Corporation, Range and Engineering Services Division in Lexington Park, MD. He augments the integrated Network Enhanced Telemetry project office and execution team, providing management and technical advice and support. Mr. Faulstich has worked in telemetry for over 35 years. He has held positions ranging from an instrumentation engineer to the director of test article preparation for the U.S. Navy. During this time he has worked on both U.S. Army and Navy test ranges. He has extensive experience with the Range Commanders Council, having served as the chairman of the technical representatives, the Telemetry Group, and the Vehicular Instrumentation Committee. He has a bachelor of electrical engineering degree and a master of science in electrical engineering degree from Georgia Tech and a master of science degree in weapons and vehicle systems from the Royal Military College of Science, Cranfield University in England. E-mail: rfaulstich@csc.com

Acknowledgments

The authors would like to recognize and sincerely thank the members of the AIRSpeed Team. Their enthusiasm, dedication, and insight have led to a truly unique solution that will pave the way for seamless integration of iNET technology at NAWCAD Patuxent River.

Last, the authors want to thank the Director of the Central Test and Evaluation Investment Program. Without his support, the iNET Project would simply not exist.

ITEA Journal 2009; 30: 549–556

Power Constrained Distributed Estimation with Cluster-Based Sensor Collaboration

Jun Fang, Ph.D. and Hongbin Li, Ph.D.
Stevens Institute of Technology,
Department of Electrical and Computer Engineering, Hoboken, New Jersey
Joseph Dorleus, Ph.D.
U.S. Army Program Executive Office for Simulation,
Training, and Instrumentation, Orlando, Florida
Hong-Liang Cui, Ph.D.
Stevens Institute of Technology,
Department of Physics, Hoboken, New Jersey

We consider the problem of distributed estimation in a power constrained collaborative wireless sensor network, where the network is divided into a set of sensor clusters, with collaboration allowed among sensors within the same cluster but not across clusters. Specifically, each cluster forms one or multiple local messages via sensor collaboration (in particular, linear operation is considered) and transmits the messages over noisy channels to a fusion center. The final estimate is constructed at the fusion center based on the noisy data received from all clusters. In this collaborative setup, we study the following fundamental problems. Given a total transmit power constraint, shall we transmit the raw data or some low-dimensional local messages for each cluster? What is the optimal collaboration scheme for each cluster? How do we optimally allocate the power among different clusters? These questions are addressed in this article. We will show that the optimum collaboration strategy is to compress the data into one local message that, depending on the channel characteristics, is transmitted using one or multiple available channels to the fusion center. The optimal power allocation among the clusters is also investigated, which yields a water-filling type of scheme.

Key words: Distributed estimation; wireless sensor network; sensor clusters; power allocation; collaboration strategy; data transmission; estimation distortion.

Distributed estimation has attracted much attention recently. One of the network architectures for distributed estimation involves a set of spatially distributed sensors linked with a fusion center (FC). Each sensor makes a noisy observation of the phenomena of interest and transmits its processed information to the FC, where a final estimate is formed. The problem of optimal power allocation among sensors given a total transmit power constraint was considered in Cui et al. (2007), Li and AlRegib (2007), Wu, Huang, and Lee (2008), and Xiao et al. (2006); the goal was to minimize the estimation distortion at the FC. For most of these works, intersensor communication is not considered. Intersensor collaboration can indeed be exploited to enhance transmission energy efficiency and improve system performance.

In this article, we consider distributed estimation in a hierarchical network architecture with localized collaboration. Specifically, we assume that the network is divided into a number of sensor clusters linked with a FC. The sensors within the same cluster have the communication resources to locally collaborate, whereas no collaboration is allowed across clusters. This might be the case for scenarios where multiple sets of sensors are spatially distributed, with each set of sensors within a small neighborhood. Each cluster

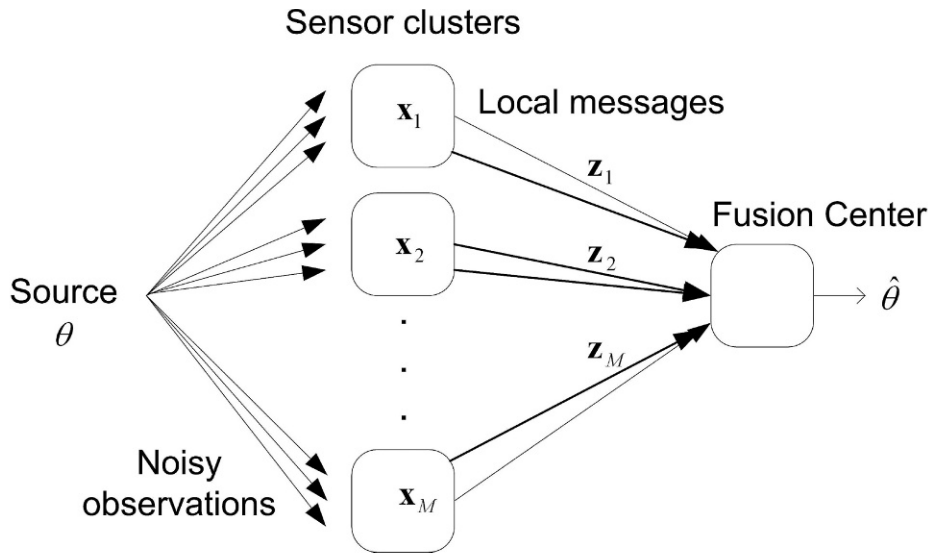


Figure 1. Collaborative setting: the network is divided into a number of sensor clusters. Sensors within each cluster can collaborate to convert their noisy observations {$\mathbf{x}_m$} into some local messages: {$\mathbf{z}_m$}.

then transmits one or multiple one-dimensional messages, which could be the raw data or obtained via sensor collaboration over noisy channels to the FC where a final estimate is formed based on the data received from all clusters. In this context, the following natural questions arise: Given a fixed amount of total transmit power, how should each cluster process its local measurements such that a minimum estimation distortion can be achieved at the FC? How should we allocate the power among the different clusters in an optimal power-distortion fashion? These questions are be addressed in this article, and we develop a fundamental understanding of this important hierarchical collaborative strategy for distributed estimation. Our work is closely related to the distributed compression-estimation approaches in Fang and Li (2008), Luo, Giannakis, and Zhang (2005), Schizas, Giannakis, and Luo (2007), Song, Zhu, and Zhou (2005), Zhang et al. (2003), and Zhu et al. (2005); their objective is to reduce the transmission requirements via dimensionality reduction. While sharing certain similarities with the distributed compression-estimation approaches, our work focuses on the optimal collaboration among sensors in a power constrained scenario.

System model and problem formulation

We consider a wireless sensor network consisting of N spatially distributed sensors, with each sensor making a noisy observation of an unknown random parameter θ: $x = h_n\theta + w_n$, where h_n denotes the observation gain and w_n denotes the additive observation noise. The sensors in the network are divided into M sensor clusters *(Figure 1)*. Each cluster, say cluster m, consists of N_m closely located sensors. The sensors in each cluster are able to collaborate to form local messages that are sent to the FC, whereas no communication is allowed across different clusters. The objective is to obtain an estimate of the unknown parameter at the FC based on the information received from the clusters. In practice, the sensor collaboration can be easily implemented. For each cluster, we choose one sensor to be the cluster head whose task is to collect the data from other sensors within the same cluster and carry out the collaborative processing. The resultant local messages are then transmitted by the cluster head to the FC. We adopt the following assumptions for this collaborative setting.

A1: The links between sensors and the cluster head within each cluster are ideal. Sensor collaboration is confined to linear operations.

A2: An uncoded analog amplify-and-forward scheme is employed to transmit the local messages from the cluster heads to the FC over noisy, wireless channels.

For notational convenience, we use $x_{m,n}$ to denote the sensor measurement of sensor n in cluster m, where $n \in \{1,\ldots, N_m\}$, $m \in \{1,\ldots,M\}$, and

$$x_{m,n} = h_{m,n}\theta + w_{m,n} \tag{1}$$

in which $h_{m,n}$ and $w_{m,n}$ denote the corresponding observation gain and additive observation noise, respectively. To capture the cluster-based collaborative scenario, we write the measurements within a cluster in a vector form: $\mathbf{x}_m=[x_{m,1}\ x_{m,2}\ \ldots\ x_{m,N_m}]^T$, which is given by

$$\mathbf{x}_m=\mathbf{h}_m\theta+\mathbf{w}_m, \tag{2}$$

with $\mathbf{h}_m=[h_{m,1}\ h_{m,2}\ \ldots\ h_{m,N_m}]^T$ and $\mathbf{w}_m=[w_{m,1}\ w_{m,2}\ \ldots\ w_{m,N_m}]^T$. The local messages via sensor collaboration within each cluster can therefore be expressed as

$$\mathbf{z}_m=\mathbf{C}_m\mathbf{x}_m, \tag{3}$$

where $\mathbf{C}_m\in R^{p_m\times N_m}$ denotes the collaboration matrix for cluster m, $p_m\le N_m$ is the dimensionality of the message vector $\mathbf{z}_m$ whose choice is discussed later. The signal received at the FC from the mth cluster is given by

$$\mathbf{y}_m=\mathbf{G}_m\mathbf{A}_m\mathbf{C}_m\mathbf{x}_m+\mathbf{v}_m, \tag{4}$$

where $\mathbf{G}_m\in R^{p_m\times p_m}$ denotes a fading multiplicative channel matrix, which can be diagonal or nondiagonal, depending on the transmission scheme (e.g., orthogonal vs. nonorthogonal channel access); $\mathbf{A}_m=\text{diag}\{a_1,\ \ldots,\ a_{p_m}\}$ is the amplification matrix with a_i denoting the amplification factor used in transmitting the ith message of $\mathbf{z}_m$; $\mathbf{v}_m\in R^{p_m}$ denotes the additive channel noise vector. Without loss of generality, we assume $\mathbf{G}_m=\mathbf{I}$ and $\mathbf{G}_m=\mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix, because the multiplicative effect of the channel matrix can be removed by carrying out a matrix inverse using an estimate of the channel matrix $\mathbf{G}_m$ at the receiver and the amplification matrix A_m can be absorbed into $\mathbf{C}_m$. We have the following assumption regarding observation noise $\{\mathbf{w}_m\}$ and channel noise $\{\mathbf{v}_m\}$.

A3: Noise $\{\mathbf{w}_m\}$ and $\{\mathbf{v}_m\}$ are zero mean with positive definite autocovariance $\{\mathbf{R}_{w,m}\}$ and $\{\mathbf{R}_{v,m}\}$, respectively, which are available at the FC. The noise across different clusters is mutually uncorrelated, i.e., $E[\mathbf{w}_i\mathbf{w}_j{}^T]=\mathbf{0}$ and $E[\mathbf{v}_i\mathbf{v}_j{}^T]=\mathbf{0}\ \forall i\neq j$.

Let $\mathbf{y}=[\mathbf{y}_1\ \mathbf{y}_2\ \ldots\ \mathbf{y}_M]^T$ denote a column vector formed by stacking the data received from all clusters. We have

$$\mathbf{y}=\mathbf{C}\mathbf{x}+\mathbf{v}=\mathbf{C}(\mathbf{h}\theta+\mathbf{w})+\mathbf{v}, \tag{5}$$

where $\mathbf{C}=\text{diag}\{\mathbf{C}_1,\ldots,\mathbf{C}_M\}$ is a block diagonal matrix with its mth block-diagonal element equal to $\mathbf{C}_m$, $\mathbf{x}=[\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_M]^T$, $\mathbf{v}=[\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_M]^T$, $\mathbf{h}=[\mathbf{h}_1\ \mathbf{h}_2\ \ldots\ \mathbf{h}_M]^T$, and $\mathbf{w}=[\mathbf{w}_1\ \mathbf{w}_2\ \ldots\ \mathbf{w}_M]^T$. A natural question arising from this scenario is to find an overall optimal collaboration matrix $\mathbf{C}$, or equivalently, a set of individual collaboration matrices $\{\mathbf{C}_m\}_{m=1}^M$, to achieve a minimum estimation distortion at the FC. Also, because the amplification factors $\{\mathbf{A}_m\}$ are incorporated into the collaboration matrices $\{\mathbf{C}_m\}$, the overall collaboration matrix $\mathbf{C}$ has to satisfy a total transmit power constraint. Specifically, using a Linear Minimum Mean-Square Error (LMMSE) estimator (Kay 1993), it can be readily verified that we are faced with the following optimization problem:

$$\begin{aligned}\min_{\mathbf{C}}\ & E[(\theta-\hat{\theta})^2]=\sigma_\theta^2-\sigma_\theta^4\mathbf{h}^T\mathbf{C}^T(\mathbf{C}\mathbf{R}_x\mathbf{C}^T+\mathbf{R}_v)^{-1}\mathbf{C}\mathbf{h}\\ &\text{s.t. tr}(\mathbf{C}\mathbf{R}_x\mathbf{C}^T)\le P,\end{aligned} \tag{6}$$

where σ_θ^2 denotes the signal variance; $\mathbf{R}_x=E[\mathbf{x}\mathbf{x}^T]$; $\text{tr}(\mathbf{C}\mathbf{R}_x\mathbf{C}^T)$ is the average transmit power required to send the local messages from all clusters to the FC; and P is a prespecified power budget for transmission.

Single cluster case

The development of the optimal collaboration matrix for the single cluster case is quite involved. Because of space limitations, we only present the main results without providing the proof.

THEOREM 1: Consider the optimal collaboration design problem formulated in (6) and described in *Figure 1*, where the sensor measurements $\mathbf{x}_m$, the local messages $\mathbf{z}_m$, and the received messages at the FC $\mathbf{y}_m$ are given by (2), (3), and (4), respectively. When $M=1$, the optimal solution to (6) is

$$\mathbf{C}^*=\gamma\sqrt{P}\mathbf{U}_v[:,1]\mathbf{h}^T\mathbf{R}_x^{-1}, \tag{7}$$

where $\mathbf{U}_v[:,1]$ denotes the first column of $\mathbf{U}_v$; $\mathbf{U}_v$ is an orthonormal matrix obtained by carrying out the eigenvalue decomposition of $\mathbf{R}_v$, i.e., $\mathbf{R}_v=\mathbf{U}_v\mathbf{D}_v\mathbf{U}_v^T$; and $\gamma=1/\sqrt{\mathbf{h}^T\mathbf{R}_x^{-1}\mathbf{h}}$. The associated estimation Mean Square Error (MSE), i.e., the value of the minimum objective function of (6), is given by

$$E\{[\theta-\hat{\theta}(\mathbf{C}^*)]^2\}=\sigma_\theta^2-\sigma_\theta^4\frac{P}{P+\min(\mathbf{d}_v)}\mathbf{h}^T\mathbf{R}_x^{-1}\mathbf{h} \tag{8}$$

PROOF: A rigorous proof is provided in Fang and Li (in press).

The optimal solution (7) has very important implications that we shall explore in the following. Considering the scenario of independent channels, i.e., $\mathbf{R}_v$ is diagonal; $\mathbf{U}_v=\mathbf{I}$ and $\mathbf{U}_v[:,1]=\mathbf{e}_1$, where $\mathbf{e}_i$ denotes the unit column vector with its ith entry equal to 1 and its other entries equal to 0. Therefore the optimal collaboration matrix becomes

$$\mathbf{C}^*=\begin{bmatrix}\gamma\sqrt{P}\mathbf{h}^T\mathbf{R}_x^{-1}\\ \mathbf{0}_{(p-1)\times N}\end{bmatrix}, \tag{9}$$

which is a matrix with its first row equal to

$\gamma\sqrt{P}\mathbf{h}^T\mathbf{R}_x^{-1}$ and all other rows equal to 0. The solution suggests that we should compress the measurements into only one local message and transmit it via the best-quality channel (note that the first row corresponds to the first channel, which has the smallest noise variance because the diagonal elements of $\mathbf{R}_v$ are assumed in an ascending order) to the FC. If the channels have identical qualities, then we can use any of them to send out the local message. Also, by rewriting the collaboration weighting vector $\gamma\sqrt{P}\mathbf{h}^T\mathbf{R}_x^{-1}$ as

$$\begin{aligned}\gamma\sqrt{P}\mathbf{h}^T\mathbf{R}_x^{-1} &= \gamma\sqrt{P}\sigma_\theta^{-2}\sigma_\theta^2\mathbf{h}^T\mathbf{R}_x^{-1}\\ &= \gamma\sqrt{P}\sigma_\theta^{-2}\mathbf{R}_{\theta x}\mathbf{R}_x^{-1},\end{aligned} \tag{10}$$

where $\mathbf{R}_{\theta x}= E[\theta\mathbf{x}^T]$, we can immediately see that the local message is exactly the LMMSE estimate $\mathbf{R}_{\theta x}\mathbf{R}_x^{-1}\mathbf{x}$ multiplied by a scalar $\gamma\sqrt{P}\sigma_\theta^{-2}$. This means that when channels are independent, LMMSE estimation followed by an amplification factor is optimal in a power-distortion sense.

We now investigate the case where the channels are correlated, i.e., $\mathbf{R}_v$ is nondiagonal. Each row of the optimal collaboration matrix can be readily expressed as follows by combining (7) and (10):

$$\mathbf{C}^*[i,:] = \mathbf{U}_v[i,1]\gamma\sqrt{P}\sigma_\theta^{-2}\mathbf{R}_{\theta x}\mathbf{R}_x^{-1}, \tag{11}$$

where $\mathbf{U}_v[i,1]$ is the $(i,1)$th entry of $\mathbf{U}_v$. Therefore the LMMSE estimate is transmitted by multiple channels with different amplification gains that are proportional to $\{\mathbf{U}_v[i,1]\}_{i=1}^p$. The number of local messages to be transmitted, p, $1 \le p \le N$, should be as large as possible because the more channels employed, the more diversity that can be provided. However, system complexity will also increase as more channels are involved.

Multiple cluster case

We now examine a general scenario where the network consists of multiple sensor clusters. In this case, the collaboration matrix $\mathbf{C}$ has a block diagonal structure because intercluster collaboration is not allowed. The approach described in previous subsection, therefore, cannot be directly applied here. To solve (6), we hope to decouple the optimization problem into a set of tractable subtasks. To this goal, we rewrite the estimation MSE as follows (Fang and Li in press).

$$\begin{aligned}E[(\theta-\hat{\theta})^2] &= \sigma_\theta^2 - \sigma_\theta^4\mathbf{h}^T\mathbf{C}^T(\mathbf{C}\mathbf{R}_x\mathbf{C}^T+\mathbf{R}_v)^{-1}\mathbf{C}\mathbf{h}\\ &= \left(\sigma_\theta^{-2} + \sum_{i=1}^{M}\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i\right)^{-1},\end{aligned} \tag{12}$$

where we use the fact that $\mathbf{R}_x = \sigma_\theta^2\mathbf{h}\mathbf{h}^T + \mathbf{R}_w$, along with the block diagonal structures of $\mathbf{C}$, $\mathbf{R}_w$, and $\mathbf{R}_v$. Therefore the optimization problem (6) becomes

$$\begin{aligned}\max_{\{\mathbf{C}_i\}} \quad & \sum_{i=1}^{M}\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i\\ \text{s.t.} \quad & \sum_{i=1}^{M}\mathrm{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T) \le P\end{aligned} \tag{13}$$

in which the power constraint follows from

$$\mathrm{tr}(\mathbf{C}\mathbf{R}_x\mathbf{C}^T) = \sum_{i=1}^{M}\mathrm{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T).$$

To use the theoretical results obtained for $M = 1$, we express the component $\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i$ in (13) as a function of $\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i$, which can be done by resorting to the Woodbury identity:

$$\begin{aligned}&\sigma_\theta^2 - \sigma_\theta^4\mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i\\ &\quad= (\sigma_\theta^{-2} + \mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i)^{-1}.\end{aligned} \tag{14}$$

For notational convenience, let

$$\begin{aligned}\mu_i(\mathbf{C}_i) &= \mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{w,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i\\ \eta_i(\mathbf{C}_i) &= \mathbf{h}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T+\mathbf{R}_{v,i})^{-1}\mathbf{C}_i\mathbf{h}_i.\end{aligned} \tag{15}$$

Therefore (14) can be rewritten as

$$\mu_i(\mathbf{C}_i) = \frac{1}{\sigma_\theta^2}\left(\frac{1}{1-\sigma_\theta^2\eta_i(\mathbf{C}_i)} - 1\right). \tag{16}$$

Substituting (16) into (13), we arrive at the following optimization

$$\begin{aligned}\max_{\{\mathbf{C}_i\}} \quad & \sum_{i=1}^{M}\frac{1}{\sigma_\theta^2}\left(\frac{1}{1-\sigma_\theta^2\eta_i(\mathbf{C}_i)} - 1\right)\\ \text{s.t.} \quad & \sum_{i=1}^{M}\mathrm{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T) \le P.\end{aligned} \tag{17}$$

Clearly, (17) can be decoupled into two sequential subtasks, i.e., a power allocation (among clusters) problem and a set of collaboration matrix design problems that can be solved using the previous results. To see this, suppose $\{P_1^*, P_2^*, \ldots, P_M^*\}$ is an optimum power assignment with

$$\mathrm{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T) \le P_i^* \qquad \forall i \in \{1, \ldots, M\}$$

$$\sum_{i=1}^{M} P_i^* \le P$$

then (17) is simplified into a set of identical problems as

$$\max_{\{\mathbf{C}_i\}} \frac{1}{\sigma_\theta^2}\left(\frac{1}{1-\sigma_\theta^2\eta_i(\mathbf{C}_i)}-1\right)$$
$$\text{s.t. } \operatorname{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T)\le P_i^*. \tag{18}$$

Note that $\sigma_\theta^2\eta_i(\mathbf{C}_i)$ must lie within the interval (0, 1) because we have $\eta(\mathbf{C}_i) > 0$ and $\mu_i(\mathbf{C}_j) > 0$ from their definitions. Hence (18) is equivalent to

$$\max_{\mathbf{C}_i} \ \eta_i(\mathbf{C}_i)$$
$$\text{s.t. } \operatorname{tr}(\mathbf{C}_i\mathbf{R}_{x,i}\mathbf{C}_i^T)\le P_i^*, \tag{19}$$

which is exactly the optimization problem discussed in the previous section. The optimal solution to (19) is given in Theorem 1. The key problem, therefore, is to determine the optimum power assignment $\{P_1^*, P_2^*, \ldots, P_M^*\}$. To meet this goal, we need to find out the relationship between the maximum objective function value $\eta_i(\mathbf{C}_i^*)$ and P_i^*. Recalling Theorem 1, more precisely, (8), we have

$$\eta_i(\mathbf{C}_i^*)=\frac{P_i^*}{P_i^*+\min(\mathbf{d}_{v,i})}\mathbf{h}_i^T\mathbf{R}_{x,i}^{-1}\mathbf{h}_i=\frac{\alpha_i P_i^*}{\beta_i+P_i^*}, \tag{20}$$

where we define $\alpha_i = \mathbf{h}_i^T\mathbf{R}_{x,i}^{-1}\mathbf{h}_i$, $\beta_i = \min(\mathbf{d}_{v,j})$, and $\mathbf{d}_{v,j}$ is a column vector consisting of the eigenvalues of $\mathbf{R}_{v,i}$ (note that $\mathbf{R}_{v,j}$ can be nondiagonal). Substituting (20) into the objective function of (17), we get

$$\sum_{i=1}^{M}\frac{1}{\sigma_\theta^2}\left(\frac{1}{1-\sigma_\theta^2\eta_i(\mathbf{C}_i^*)}-1\right)=\sum_{i=1}^{M}\frac{\alpha_i P_i^*}{(1-\sigma_\theta^2\alpha_i)P_i^*+\beta_i}. \tag{21}$$

Clearly, the optimal power allocation $\{P_1^*, P_2^*, \ldots, P_M^*\}$ must be the one, among all feasible power assignments, that maximizes (21). Therefore, it can be found out by

$$\min_{\{P_1,\ldots,P_M\}} -\sum_{i=1}^{M}\frac{\alpha_i P_i}{(1-\sigma_\theta^2\alpha_i)P_i+\beta_i}$$
$$\text{s.t. } \sum_{i=1}^{M}P_i\le P$$
$$P_i\ge 0 \quad \forall i\in\{1,\ldots,M\}. \tag{22}$$

It is easy to verify that the optimization problem (22) is convex because its Hessian matrix, which is a diagonal matrix in this case, is positive semidefinite on the convex set defined by the linear constraints. Although (22) is efficiently solvable by numerical methods, it can also be solved analytically by resorting to the Lagrangian function and Karush-Kuhn-Tucker conditions, which leads to a water-filling type power allocation scheme. The details are omitted here because of space limitations. Briefly speaking, for a threshold λ, we have

$$P_i=\begin{cases}\dfrac{1}{\varphi_i}\left(\sqrt{\dfrac{\phi_i}{\lambda}}-1\right) & \phi_i\ge\lambda\\ 0 & \text{otherwise}\end{cases} \tag{23}$$

where $\phi_i = \alpha_i/\beta_i$, $\varphi_i=(1-\sigma_\theta^2\alpha_i)/\beta_i$. It is easy to see that each cluster can decide whether to transmit or keep silent by the criterion $\phi_i \ge \lambda$. Note that ϕ_i is the ratio of $\mathbf{h}_i^T\mathbf{R}_{x,i}^{-1}\mathbf{h}_i$ to $\min(\mathbf{d}_{v,i})$, with the former is a measure of the cluster's estimation quality (a larger value indicates a better estimation accuracy) and the latter a measure of the cluster's channel quality (a smaller value indicates a better channel quality).

So far we have developed an analytical approach that leads to an optimal solution to (6). For clarity, we now summarize the steps of our proposed method.

1. Given the prior knowledge of the autocorrelation matrices $\{\mathbf{R}_{v,i}\}_{i=1}^M$, $\{\mathbf{R}_{w,i}\}_{i=1}^M$, and the observation gain vectors $\{\mathbf{h}_i\}_{i=1}^M$, compute $\{\alpha_i\}_{i=1}^M$, and $\{\beta_i\}_{i=1}^M$, where $\alpha_i=\mathbf{h}_i^T\mathbf{R}_{x,i}^{-1}\mathbf{h}_i$ and $\beta_i = \min(\mathbf{d}_{v,\mathrm{i}})$.
2. Given the total power constraint P, find the optimal power allocation among clusters via (22).
3. With the optimal power assignment $\{P_1^*, P_2^*, \ldots, P_M^*\}$ derived in the previous step, determine the optimal collaboration matrices $\{\mathbf{C}_i\}_{i=1}^M$ via (19), whose solution is detailed in Theorem 1.

Simulation results

We consider the single cluster case and carry out a simple performance analysis to corroborate our theoretical results (more analysis and simulation results are available in Fang and Li [in press]). We compare our optimal collaboration strategy with the scheme proposed in Cui et al. (2007), where there is no intersensor collaboration and each sensor transmits its observation to the FC with optimally assigned power. For simplicity, we consider a homogeneous environment with identical observation and channel qualities, where σ_w^2 denotes the observation noise variance and σ_v^2 represents the channel noise variance. Also, all observation and channel gains are assumed to be unitary, i.e., =1, throughout all examples in the article. Clearly, an equal power allocation is optimum for Cui et al. (2007) and the corresponding estimation MSE can be shown to be

$$\mathrm{MSE_{NC}}=\frac{P\sigma_w^2\sigma_\theta^2+N\sigma_v^2\sigma_\theta^4+N\sigma_v^2\sigma_\theta^2\sigma_w^2}{PN\sigma_\theta^2+P\sigma_w^2+N\sigma_v^2\sigma_\theta^2+N\sigma_v^2\sigma_w^2}, \tag{24}$$

where the subscript NC denotes noncollaboration. For our collaboration strategy, the estimation MSE can be

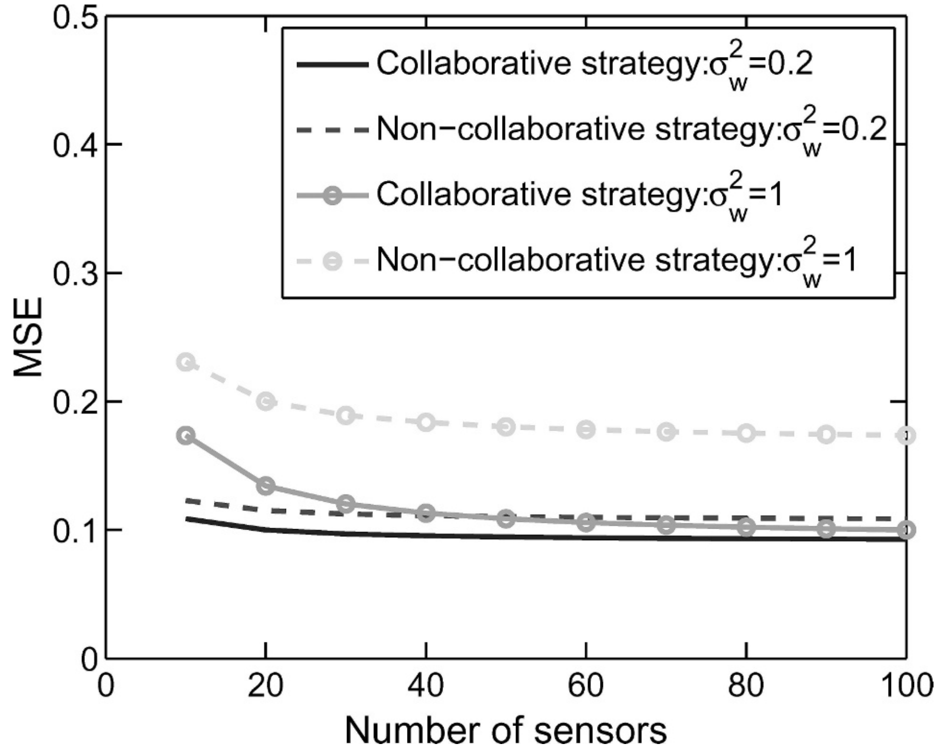


Figure 2. MSEs of collaborative and noncollaborative strategies versus number of sensors $\sigma_v^2=0.1$, $\sigma_\theta^2=1$, and P = 1.

computed by using (8), which reduces to

$$\mathrm{MSE_{OC}}=\frac{P\sigma_w^2\sigma_\theta^2+N\sigma_v^2\sigma_\theta^4+\sigma_v^2\sigma_\theta^2\sigma_w^2}{PN\sigma_\theta^2+P\sigma_w^2+N\sigma_v^2\sigma_\theta^2+\sigma_v^2\sigma_w^2}, \qquad (25)$$

where the subscript OC denotes optimal collaboration. For notational convenience, let $a=P\sigma_w^2\sigma_\theta^2+N\sigma_v^2\sigma_\theta^4$ and $b=PN\sigma_\theta^2+P\sigma_w^2+N\sigma_v^2\sigma_\theta^2$. It can be easily verified that

$$(a+N\sigma_v^2\sigma_\theta^2\sigma_w^2)(b+\sigma_v^2\sigma_w^2) \geq (a+\sigma_v^2\sigma_\theta^2\sigma_w^2)(b+N\sigma_v^2\sigma_w^2), \qquad (26)$$

where (26) becomes an equality only when $N = 1$. Hence as expected, the relationship MSENC $\geq$ MSEOC holds, which means that the optimal collaboration scheme should always outperform the noncollaboration scheme.

Figure 2 depicts the estimation MSEs of the two schemes as a function of N under a total transmit power constraint, with $\sigma_w^2=0.2$ and $\sigma_w^2=1$, respectively. From *Figure 2*, we see that both schemes benefit from an increasing number of sensors; as N increases, the estimation MSEs will asymptotically approach certain values that, however, are nonzero. This observation can be readily verified from (24)–(25). Also, it can be seen that the noncollaborative scheme is sensitive to the value of σ_w^2; as the observation quality deteriorates, its performance degrades considerably. In contrast, the collaborative strategy demonstrates a certain degree of robustness against the observation quality deterioration. In *Figure 3*, we plot the estimation MSE versus the total transmit power. We see that the performance gap between the two strategies shrinks as the transmit power increases. In fact, from (24)–(25) we observe that as the transmit power goes to infinity, these two strategies approach identical performance. This suggests that the collaborative strategy should be preferred especially when the sensor observation qualities are bad and transmit power is severely constrained.

Conclusion

We studied an optimal collaboration and power allocation problem for distributed estimation in a power-constrained collaborative sensor network, where the network consists of a number of sensor clusters, and collaboration is allowed within the same cluster but not across clusters. Our theoretical results showed that, given a specified total transmit power, the power should be assigned among the clusters in a water-filling manner, with each cluster deciding whether to transmit or keep silent by comparing with a threshold the ratio of a measure of the cluster's estimation quality to a measure of the cluster's channel quality. Also, for each cluster, if the channels from this cluster to the FC are independent, then an optimal collaboration yields only one local message, which is sent from the best channel

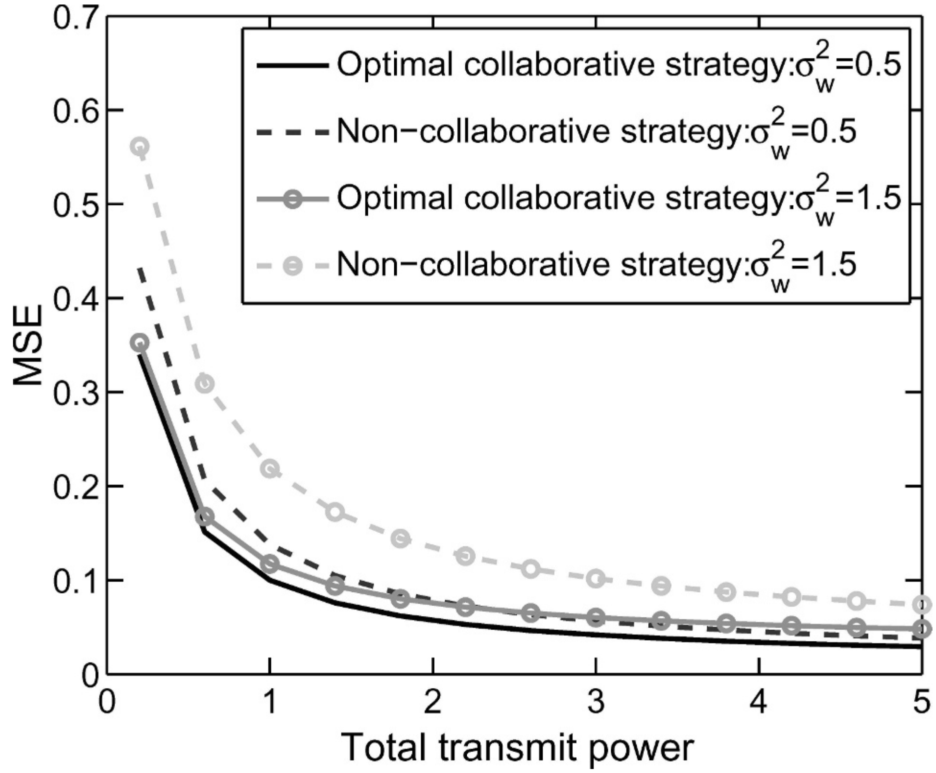


Figure 3. MSEs of collaborative and noncollaborative strategies versus total transmit power. $\sigma_v^2=0.1$, $\sigma_\theta^2=1$, and N = 50.

within the cluster to the FC; otherwise the local message has to be sent across all channels within the cluster at different power levels matched to their channel quality. Specifically, in either case, the compressed local message is exactly the local LMMSE estimate multiplied by an amplification factor. Simulation results have been presented to corroborate our theoretical analysis. ❑

JUN FANG received a bachelor's of science and a master's of science degree in electrical engineering from Xidian University, Xi'an, China, in 1998 and 2001, respectively, and earned a doctor of philosophy degree in electrical engineering from the National University of Singapore, Singapore, in 2006. During 2006 he was with the Department of Electrical and Computer Engineering, Duke University, as a postdoctoral research associate. Currently he is a postdoctoral research associate with the Department of Electrical and Computer Engineering at the Stevens Institute of Technology. His research interests include statistical signal processing, wireless communications, and distributed estimation and detection with their applications on wireless sensor networks. E-mail: Jun.Fang@stevens.edu

JOSEPH DORLEUS is currently a lead telecommunication/systems engineer at PEO STRI, Orlando, Florida. He has worked and held both technical and managerial positions in the private sector as well as in the government. He holds bachelor's of science and master's of science degrees in electrical engineering from Polytechnic University (formerly Polytechnic Institute of New York), Brooklyn, New York, and a doctor of philosophy degree in electrical engineering from Stevens Institute of Technology, Hoboken, New Jersey. His research interests include optical networks, all-optical network management and monitoring, and modeling and simulation of wireless networks. He is a member of the International Test and Evaluation (ITEA), the Institute of Electrical, Electronics Engineering (IEEE), the Defense Technical Information Center (DTIC), the Army Acquisition Corps (AAC), and the International Society for Optical Engineering (SPIE). He has authored, coauthored, and presented numerous technical papers that are published in technical journals, conferences, and proceedings such as IEEE, SPIE, ITSEC, and ITEA. Dr. Dorleus is the recipient of the Army Achievement Medal for Exceptional Civilian Service. He was also the Army Materiel Command's nominee for Black Engineer of Year Award in 2001. E-mail: Joseph.dorleus@us.army.mil

HONGBIN LI received his bachelor's of science and master's of science degrees from the University of Electronic Science and Technology of China, Chengdu, in 1991 and 1994, respectively, and a doctor of philosophy degree from the University of Florida, Gainesville, Florida, in 1999, all in electrical engineering. From July 1996 to May 1999, he was

a research assistant in the Department of Electrical and Computer Engineering at the University of Florida. He was a visiting summer faculty member at the Air Force Research Laboratory, Rome, New York, in the summers of 2003 and 2004. Since July 1999, he has been with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, New Jersey, where he is an associate professor. His current research interests include statistical signal processing, wireless communications, and radars. Dr. Li is a member of Tau Beta Pi and Phi Kappa Phi. He received the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for excellence in research in 2001 from Stevens Institute of Technology, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He is a member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society. He is/has been an editor or associate editor for the IEEE Transactions on Wireless Communications, IEEE Signal Processing Letters, and IEEE Transactions on Signal Processing, and served as a guest editor for EURASIP Journal on Applied Signal Processing, Special Issue on Distributed Signal Processing Techniques for Wireless Sensor Networks. E-mail: Hongbin.Li@stevens.edu

HONG-LIANG CUI is a professor of physics at Stevens Institute of Technology, where he directs the Applied Electronics Laboratory. He received his undergraduate education in applied physics with a concentration in laser optics from the Changchun Institute of Optics and Fine Mechanics in Changchun, China, with a bachelor's of engineering degree. In 1981 he came to the United States for graduate study as one of the first group of Chinese physics students in the CUSPEA program, obtaining a doctor of philosophy degree in theoretical condensed matter physics in 1987 from Stevens Institute of Technology, where he has been on the faculty ever since. His research efforts have been concentrated in the areas of solid-state electronics/nanoelectronics, optical communications and sensing, electromagnetic wave propagation, and interaction with matters such as chemical and bioagents, and high-performance computing approach to modeling of physical devices and phenomena. His work has been funded by NSF, ARO, ONR, and DARPA. He has published more than 190 research papers in peer-reviewed scientific journals, holds 9 U.S. patents, and guided more than 30 doctoral dissertations to completion. He holds membership in the American Physical Society, the Institute of Electrical and Electronics Engineers, the Optical Society of America, and Sigma Xi. E-mail: Hong-Liang.Cui@stevens.edu

References

Cui, S., J.-J. Xiao, A. J. Goldsmith, Z.-Q. Luo, and H. V. Poor. 2007. Estimation diversity and energy efficiency in distributed sensing. *IEEE Transactions on Signal Processing*, no. 9 (55): 4683–4695.

Fang, J., and H. Li. 2008. Joint dimension assignment and compression for distributed multisensor estimation. *IEEE Signal Processing Letters*. 15, 174–177.

Fang, J., and H. Li. 2009. Power constrained distributed estimation with cluster-based sensor collaboration. *IEEE Transactions on Wireless Communications*. 8: 3822–3832.

Kay, S. M. 1993. *Fundamentals of statistical signal processing: estimation theory*. Upper Saddle River, New Jersey: Prentice Hall.

Li, J., and G. AlRegib. 2007. Rate-constrained distributed estimation in wireless sensor networks. *IEEE Transactions on Signal Processing*. 55 (5): 1634–1643.

Luo, Z.-Q., G. B. Giannakis, and S. Zhang. 2005. Optimal linear decentralized estimation in a bandwidth constrained sensor network. In *Proceedings of the International Symposium on Information Theory*. Adelaide, Australia.

Schizas, I. D., G. B. Giannakis, and Z.-Q. Luo. 2007. Distributed estimation using reduced dimensionality sensor observations. *IEEE Transactions on Signal Processing*, no. 8, 55: 4284–4299.

Song, E., Y. Zhu, and J. Zhou. 2005. Sensors' optimal dimensionality compression matrix in estimation fusion. *Automatica*. 41, 2131–2139.

Wu, J.-Y., Q.-Z. Huang, and T.-S. Lee. 2008. Minimal energy decentralized estimation via exploiting the statistical knowledge of sensor noise variance. *IEEE Transactions on Signal Processing*. 56 (5): 2171–2176.

Xiao, J.-J., S. Cui, Z.-Q. Luo, and A. J. Goldsmith. 2006. Power scheduling of universal decentralized estimation in sensor networks. *IEEE Transactions on Signal Processing*. 54 (2): 413–422.

Zhang, K., X. R. Li, P. Zhang, and H. Li. 2003. Optimal linear estimation fusion—Part VI: sensor data compression. In *Proceedings of the International Conference on Information Fusion*. Queensland, Australia.

Zhu, Y., E. Song, J. Zhou, and Z. You. 2005. Optimal dimensionality reduction of sensor data in multisensor estimation fusion. *IEEE Transactions on Signal Processing*. 53 (5): 1631–1639.

ITEA Journal 2009; 30: 557–561

PASGT Helmet Test: An Example of Effective Intra-Government Testing Collaboration

Sara Campbell
U.S. Army Evaluation Center, Aberdeen Proving Ground, Maryland

The Office of the Director of Operational Test and Evaluation requested Army Test and Evaluation Command and Army Research Laboratory Survivability/Lethality Analysis Directorate support to plan, execute, and report on a test program to assess ballistic performance of Kevlar helmets. The effort was a result of a Department of Justice investigation into a primary contractor for Kevlar helmets. That investigation identified alleged noncompliance with manufacturing requirements specified in the Kevlar helmet military specification, which could result in decreased personnel protection. Helmets tested during this effort were selected from a population of items manufactured during the spanned Department of Justice investigation (1987–2006). An initial test strategy was developed with Office of the Director of Operational Test and Evaluation, Army Test and Evaluation Command, and Army Research Laboratory Survivability/Lethality Analysis Directorate to characterize the helmet performance using ballistic v_{50} test procedures.

Key words: Ballistic limit (v_{50}); intra-government collaboration; Kevlar helmet performance; military specifications; testing.

This effort required extensive collaboration and cooperation from several agencies and organizations within the Department of Defense (DoD). Additional complexity arose due to the potential legal implications stemming from the Department of Justice (DOJ) investigation. Within the DoD, the following agencies and organizations were involved in the effort.

Office Secretary of Defense (OSD)

- Executive Secretariat,
- Under Secretary of Defense for Acquisition, Technology & Logistics,
- Deputy Under Secretary of Defense for Logistics and Material Readiness (DUSD [L&MR]),
- Director, Operational Test & Evaluation (DOT&E),
- Defense Logistics Agency (DLA),
- Defense Supply Center–Philadelphia (DSC-P).

U.S. Army

- Army Test and Evaluation Command (ATEC),
- Army Evaluation Center,
- Developmental Test Command,
- Aberdeen Test Center (ATC),
- Army Research Labs–Survivability/Lethality Analysis Directorate (ARL-SLAD),
- Program Executive Office Soldier (PEO Soldier).

U.S. Marine Corps

- Marine Corps System Command (MARCORSYSCOM).

The effort was initiated by the DOJ with an informal request to DoD addressed “To Whom It May Concern.” Because the DOJ letter did not identify an agency, the request was circulated throughout DoD in an effort to identify the most appropriate agency to execute a response. The DOJ request was first given to PEO Soldier. PEO Soldier gave the letter to DOT&E, who then turned the DOJ letter over to the OSD General Counsel. OSD General Counsel contacted DOJ and requested the letter be rewritten and sent to the Secretary of the Army.

The Secretary of the Army then handed the DOJ request up to the OSD Executive Secretariat, who in turn named DUSD (L&MR) as the lead organization. DUSD (L&MR) tasking from OSD was to investigate the quantity of fielded Personnel Armor System for

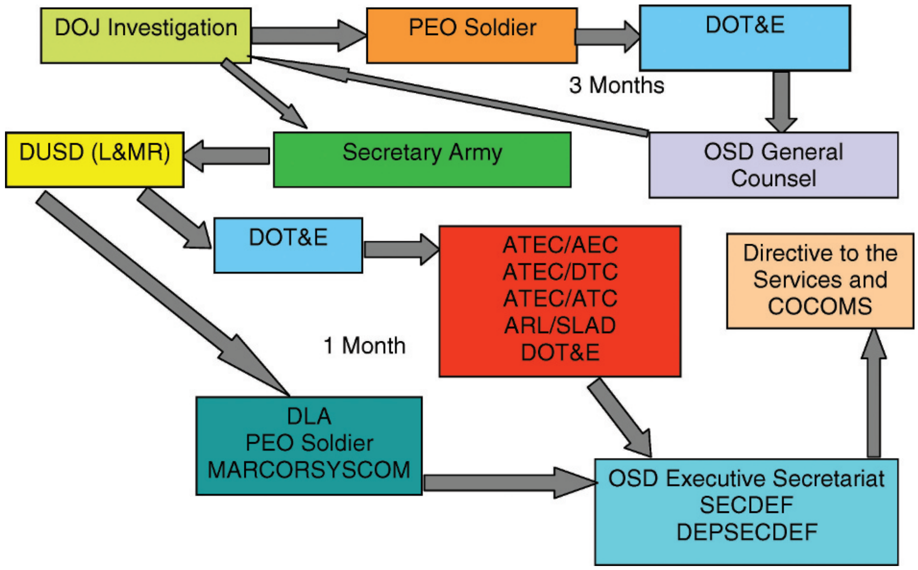


Figure 1. Intra-government testing and collaboration.

Ground Troops (PASGT) helmets, determine what services the PASGT helmet had, and develop a path forward (including the cost and logistics of replacing the PASGT helmets). The Honorable Jack Bell (DUSD [L&MR]) sent a letter to the Honorable Dr. Charles McQueary requesting DOT&E perform testing. DOT&E tasked ATEC and ARL-SLAD to execute a test and report back to DOT&E on the results. DOT&E then reported to the Secretary of Defense.

DUSD (L&MR) directed DLA to determine the PASGT helmet worldwide distribution and stockpile. DUSD (L&MR) requested that the information be prepared in the event that the PASGT helmet needed to be replaced with the Advance Combat Helmets (ACH) and Light Weight Helmets (LWH) in the event of an unfavorable test outcome. DLA designated DSC-P as lead logistics organization for that effort.

DLA contacted PEO Soldier and MARCORSYSCOM to determine the quantity of ACH and LWH in inventory and available for fielding. DSC-P also led the effort to acquire the necessary helmet test assets in various sizes and manufacturing years to represent the helmet population in question.

Figure 1 illustrates the time line and coordination that crossed agency boundaries to execute an effective and timely test. The coordination of the DOJ request took approximately 3 months, while the test was planned, executed, and reported on in 1 month.

Impact

The outcome of the test had the potential to impact the survivability of millions of U.S. and foreign service men and women. The purpose of the test was to determine if the alleged noncompliance with Kevlar manufacturing requirements resulted in a decrease in ballistic protection. The PASGT helmet was widely fielded to millions of U.S. and foreign service men and women. *Figure 2* illustrates the breakdown of PASGT use by service and foreign military.

In 2007, the PASGT helmet use by U.S. military Services in theater was as follows:

- Army: none,
- Navy: 2,568 in Iraq,
- Air Force: 6,768 in Iraq,
- Marines: none,
- approximately 29,000 issued to noncombat personnel.

At the time of the test in 2007, the PASGT helmet was protecting service men and women in formations alongside those wearing the ACH and LWH. The ACH and LWH are 21st-century helmets with superior technological advantage and were designed to complement the body armor now being used.

Test objective

The objective of the test program was to determine if the empirical estimates of the ballistic limit (v_{50}) met the requirement of 2,000 feet/second (MIL-H-44099A, para. 3.5.2) with a 95 percent confidence level for the PASGT helmet population in question.

Test

ARL-SLAD and ATC conducted testing simultaneously in order to increase the speed of the test and

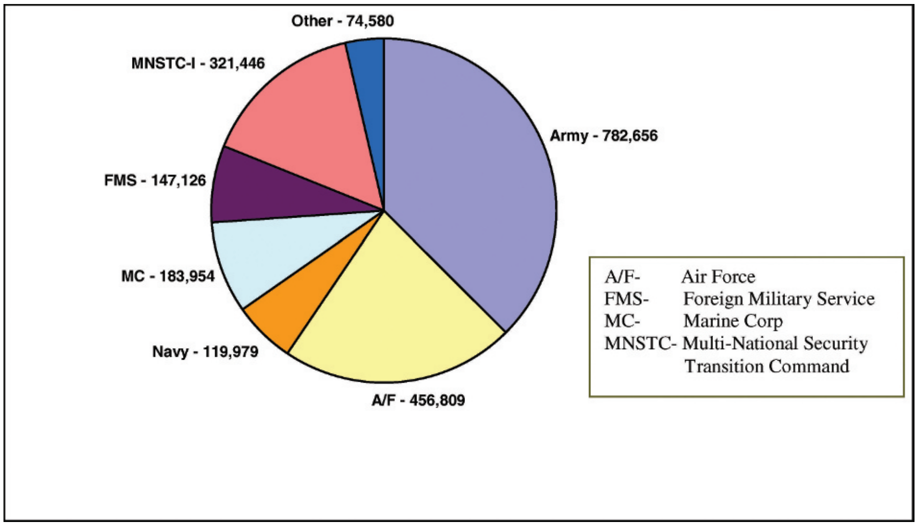


Figure 2. PASGT—total quantity demanded, FY89 to present.

reduce any potential biasing. Both facilities followed test protocols identified in military specifications MIL-H-44099A Helmet, Ground Troops and Parachutists dated 22 December 1986 and MIL-STD-662F, "V50 Ballistic Test for Armor" dated 18 December 1997. An industry standard gun barrel was used to launch the 17-grain fragment simulating projectile (FSP) where the striking velocity was recorded and yaw measured at or near target impact. Yaw card data was compared, post-shot, to a predetermined maximum 5-degree yaw template to ensure that the penetrator impacted the helmet at less than the 5-degree yaw requirement. The helmet was rigidly mounted to a stationary target fixture. Fair hit impact points were a minimum distance of 1.5 inches from each other, from holes, crease locations, or the edge of helmet. A complete penetration was one where any part of the FSP perforated a witness plate positioned inside the helmet, 2 inches from the impact point. *Figure 3* shows exterior and interior views of the impact locations for a tested helmet.

Helmets are manufactured in five sizes: X-Small, Small, Medium, Large, and X-Large. Test sample dates of manufacture are from 1987 to 2006, spanning the period of helmet production in question. Helmets were obtained from current inventory levels at the DLA and tested against criteria set forth in Military Specification MIL-H-44099A. The test samples included two manufacturing age groups (1987–1996 and 1997–2006) and five sizes (XS, S, M, L, XL). Based upon inventory received from DLA, ATEC and ARL determined the randomization and distribution of test samples. ATEC and ARL also made the assumption that the random sample received from inventory was representative of the PASGT helmet population in question.

For each helmet group, the v_{50} ballistic limit was determined using four series of the modified Langlie sequential firing procedure. For each helmet group, four series (10–15 shots per series) of the modified Langlie firing procedure were conducted. Size Small was not available for the older helmets; therefore, there were a total of 36 series. *Table 1* shows the total number of shots conducted by helmet size and year.

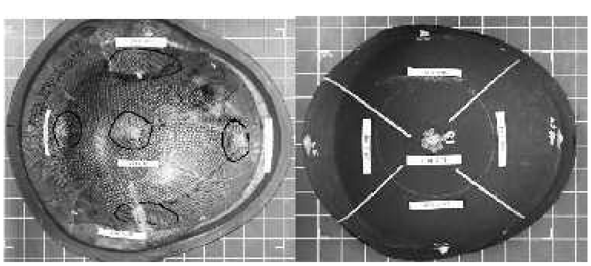


Figure 3. Exterior and interior views of a helmet after testing.

Table 1. Total number of shots conducted by helmet size and manufactured group.

	Manufactured date		
Helmet size	**1987–1991**	**2003**	**Shots per size**
X-Small	50	50	100
Small	N/A	58	58
Medium	43	52	95
Large	48	54	102
X-Large	60	50	110
Shots per manufactured group	201	264	465

Data collection

The modified Langlie sequential firing procedure is based on the Langlie method (DARCOM Pamphlet 706-103, 1983 and TOP 2-2-710) and was used to select velocities for obtaining estimates of the v_{50} ballistic limit. Several modifications were made to obtain velocities away from the mean to better estimate the entire response curve. A computer program was created to automate the procedure for each testing facility (ARL and ATC). The modified Langlie sequential firing procedure is listed below. The upper and lower projectile velocity limits (gates) were set at 200 feet/second from the postulated mean. All shots were conducted at zero degrees obliquity.

For each helmet, one shot was fired in each subdivision of the helmet as specified in MIL-H-44099A: the crown and four circumferential quadrants—front, back, left, and right. Therefore, to complete a v_{50} series, at least two or three helmets were required. In addition, four series of the sequential procedure were conducted for each helmet group. Each series acted independently, as if starting over.

Testing contained the following for each series, with a minimum of 10 shots and maximum of 15:

- A zone of mixed results (at least one partial penetration has a higher velocity than a complete penetration). The size of the zone of mixed results is defined as the difference in velocity between the highest partial penetration and the lowest complete penetration.
- The average of the complete penetrations is larger than the average of the partial penetrations.
- The spread of the three highest partial penetrations and the three lowest complete penetrations is within 125 feet/second.
- Ensure that the data set contains values approximately ± 66 feet/second from the v_{50} that is estimated from the three highest partial penetrations and three lowest complete penetrations.

Since all the data from four series were combined, there was always a zone of mixed results from which to get parameter estimates.

Analysis

Statistical hypothesis tests were used to analyze and compare the v_{50} data. Statistical hypothesis testing is a procedure that involves stating something to be tested, collecting evidence, and then making a decision as to whether the statement (null hypothesis) should be accepted as true, or rejected. To determine failure to accept or reject the null hypothesis (H_0), the probability values (*P* values) were evaluated.

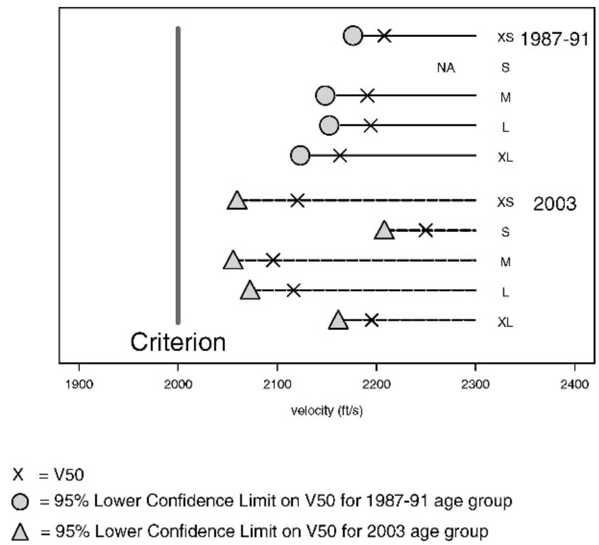


Figure 4. v_{50} and 95 percent lower confidence limits on the v_{50} compared with criterion of 2,000 feet/second.

If the *P* value is less than .05, the null hypothesis would be rejected with a statement that there is a significant difference between or among the v_{50}s. However, if the *P* value is greater than or equal to .05, the null hypothesis would fail to be rejected with a statement that there is not enough evidence in the data to determine that v_{50}s differ.

Results

Analysis of the test data showed that there was no significant statistical difference between the ARL-SLAD and ATC test facility data and between the year groups. Analysis did show a small statistical difference among the sizes in the 2003-year group. *Figure 4* shows the v_{50} for each size within each year group along with the 95 percent lower confidence limit (LCL). Each LCL is above the criterion of 2,000 feet/second. The sizes were combined (pooled) into one group with a v_{50} of 2,157 feet/second. The 1987–1991 sizes were not significantly different and their combined v_{50} is 2,187 feet/second.

Since the manufactured years were not significantly different, the data were combined to produce one v_{50} of 2,170 feet/second, and the 95 percent LCL on the v_{50} is 2,154 feet/second as shown in *Figure 5*. Therefore, if we repeatedly conducted this experiment (same sample size from the same population), and a confidence interval was calculated from each repetition, then 95 percent of these intervals should contain the population mean (v_{50}). There is a 95 percent confidence that our interval with a lower bound at 2,154 feet/second contains the true v_{50}.

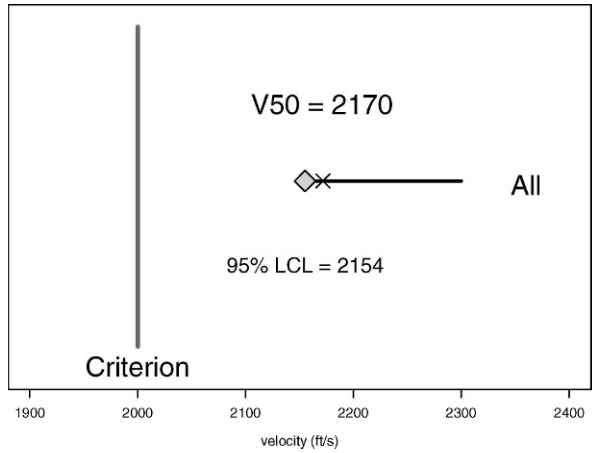


Figure 5. Combined v_{50} for all groups and sizes and the 95 percent lower confidence limit on v_{50}.

Test conclusion

Testing of PASGT helmets in five different sizes (XS, S, M, L, XL) from production years 1987–2003 produced an aggregate v_{50} at 2,170 feet/second and an LCL at 2,154 feet/second. The robustness of testing was sufficient to produce an aggregate v_{50} with 95 percent level of confidence for the population. There was some statistical variation among sizes for the 2003-year group; however, variations between groups were not significantly different, and v_{50}s could be aggregated into one summary result, still maintaining a 95 percent confidence level. As stand-alone size and year groups, all combinations exceeded the v_{50} requirement at the 95 percent confidence level. The helmets tested are in compliance with the MIL Specification MIL-H-44099A Helmet, Ground Troops and Parachutists.

Conclusion

Although the test results indicated the PASGT helmets were within specifications, Deputy Secretary of Defense recalled all PASGTs and directed the Services to immediately cross-level all personnel participating in operations in South West Asia to the ACH or LWH, thereby increasing the survivability of our forces. Consequently, the U.S. Services now have one uniform protection standard for troops operation in the same battle space. ❑

SARA CAMPBELL currently works for the Army Test and Evaluation Command (ATEC) as the lead evaluator on all Army Personal Protective Equipment at Aberdeen Proving Ground, Maryland. She has been active in the personal Blunt Trauma and Non-Lethal Munitions communities. She received a bachelor of science degree in electrical engineering from Virginia Tech and a masters of science degree in engineering management from the University of Maryland Baltimore Campus. E-mail: sara.campbell1@us.army.mil

References

ATEC. *System Evaluation Report for Personnel Armor System Ground Troops (PASGT) Military Helmets.* Alexandria, VA: U. S. Army Test and Evaluation Command (ATEC), August 2007.

ATEC. *Event Design Plan for the Ballistic Vulnerability Test of the Personnel Armor System Ground Troops (PASGT) Military Helmets.* Alexandria, VA: ATEC, July 2007.

ITEA Journal 2009; 30: 562–567

Testers and Managers Need to Examine More Than Their Normally Recognized Test Standards

George Jackelen
Software Consultants, Inc. (SCI), Lanham, Maryland

Many standards, methodology, and model producing groups have processes to plan and execute various testing levels, and to produce related test documentation. Many testers and managers are unaware of what are currently considered non-test standards or test standards not normally used in their test environments. These standards should be evaluated for an enhanced testing perspective to assure testing meets overall business, customer, and end user needs and concerns, not just documented requirements. This article addresses some test and non-test standards, and how these standards help people improve the quality of testing, test artifacts, and overall reliability of delivered products. The objective is to show the need to expand testing knowledge beyond finding product defects or validating that a product works as documented. Testing as viewed by business managers, customers, and end users is critical to assure products are accepted and have operational business benefits, rather than just technical benefits.

Key words: Best practices; business objectives; customer acceptance; end user needs; learning; non-test standards; technical requirements; testing knowledge; verification & validation.

Some pure-test standards (standards only addressing currently defined test issues) are heavily involved with test case development, i.e., documents providing step-by-step instructions about how to execute specific tests to verify the known requirements and validate the functional specifications through traceability from system requirements to software and hardware requirements. Other pure-test standards deal with development of test result reports or test plans. In many situations there is a test-plan standard or template for each test level, e.g., overall test process, user acceptance, system, unit, integration, and regression testing.

Many products are tested to prove the satisfaction of the technical requirements rather than business (including the test organization) objectives and goals. Part of the problem lies in test preparation. For instance, many testers assume the requirements analysis and design phases assure the requirements are documented and known for the testers to comprehensively verify, and validate the end products. Unless the requirements have been developed to clearly incorporate the business (e.g., corporation and division) goals and objectives as well as the actual end users (not just the initial target set) needs and expectations, then the end product may successfully meet the known requirements without providing any return on investment (ROI) to the business or being adopted by end users. Frequently, systems are developed in a business vacuum, only meeting the requirements of a particular domain, or stovepipe, operational area. Today, the government and industry need to address the entire needs of the enterprise, whether agency or corporate wide, to enhance ROI, reduce redundancy, and improve enterprise effectiveness and efficiency.

Another concern is when testers and managers look at national and international standards (including models and methodologies) dealing with testing, their domain is normally limited to standards having a variation of the term "test" in the title. Thus, many people are not aware of other standards related to or concerned with testing or risk management.

The author recommends looking at non-pure test standards for insight into areas to improve the testing process and knowledge of testers and managers. To this end, the article examines the following:

- four International Organization of Standardization (ISO) standards,
- two Institute for Electrical and Electronics Engineers, Inc. (IEEE) standards,

- a Software Engineering Institute (SEI) methodology/model known as: Capability Maturity Model—Integration (CMMI),
- The Project Management Institute's "*A Guide to the Project Management Body of Knowledge—PMBOK® Guide*,"
- two United States Department of Defense (DoD) standards.

These standards were chosen based on the author's familiarity with these standards and the availability of material often overlooked by test organizations. This is not a comprehensive list nor are these standards restrictive in their use.

ISO, IEEE, The Project Management Institute (PMI), etc., have many standards that people should look at to receive a broader view of project, software, system, etc., processes. This broader view helps testers and managers understand how others operate and test within our global environment. To maintain currency, these groups periodically (e.g., every 5 years) perform reviews to determine if the standards need to be updated or deleted. Thus, there is a continuous effort to maintain up-to-date standards.

The following sections discuss some standards about which testers and managers may not be aware.

ISO

ISO 9000

ISO 9000, *Quality Management System*, is a set of Quality Management System (QMS) standards (International Organization for Standardization, 2005).

Testing (e.g., Verification and Validation [V&V]) isn't the main topic. But, because testing is a quality measuring tool, this standard should be examined to relate testing to organizational quality issues (e.g., measurement, analysis, and improvement). ISO 9000 describes the need to verify vendor and customer products prior to implementation; a form of testing often overlooked by testers and managers.

ISO 12207

ISO 12207, *Standard for Information Technology—Software Life Cycle Processes*, establishes a software life cycle process (International Organization for Standardization, 2008). Each of the 23 processes has a set of outcomes (224 in total) associated with it. ISO 12207 is fully harmonized with ISO/IEC 15288, *System Life Cycle Processes*, which also describes testing. The standard's main objective is to supply a common structure, so buyers, suppliers, developers, maintainers, operators, managers, and technicians involved with software development use a common language established in the form of well-defined processes. The standard's structure is intended to be conceived in a flexible, modular way, adaptable to the necessities of whoever uses it. ISO 12207 is based on two basic principles: (a) modularity—processes with minimum coupling and maximum cohesion, and (b) responsibility, establish a responsibility for each process, facilitating the application of the standard in projects where many people can be legally involved. This set of processes, activities, and tasks can be adapted according to any software project.

ISO 12207 presents testing as support processes (V&V) and describes how these processes fit within a total software life cycle. For testers and managers, ISO 12207 can be a bridge between the groups to help with communications and development of strategies and techniques. Inadequate communication between business management and technical management can result from not speaking the same "language," which easily results in not properly capturing the business requirements and the translation of these requirements into test cases.

ISO 20000

ISO 20000, *Information Technology Infrastructure Library (ITIL)*, ensures systems are in place to manage customer-identified incidents of operational systems (International Organization for Standardization, 2009). ISO 20000 consists of two standards and promotes the adoption of an integrated process approach to effectively deliver managed services to meet business and customer requirements.

For the purpose of illustration, think of ISO 20000 supporting a help desk operation: incidents are received and problem reports are prepared by the help desk for problem determination and repair by others (e.g., developers or maintainers). When a problem's root cause is identified and an approved solution is prepared, the solution is tested prior to incorporating a fix into the defective system and preparation for deployment.

ISO 20000 doesn't specifically address how testing should be implemented but instead addresses how testing must be involved with other activities to assure a thorough understanding and coordination of work on defects, defect resolution, and communications to the help desk and system users. ISO 20000 shows how testing needs to be fully integrated into an organization-customer relationship. ISO 20000 also describes the need to test contingency plans in case there is an organizational disaster (e.g., the lost of key facility, hardware or software due to floods, tornadoes, human-caused damage, etc).

ISO 27000

ISO 27000, *Information Technology—Security Techniques—Information Security Management Systems*, is part of a family of ISO Information Security Management (ISM) System (ISMS) standards (six in all) and provides uniformity and consistency of fundamental terms and definitions (vocabulary) used throughout the ISMS standards (International Organization for Standardization, 2009). The series provides ISM best practice recommendations, and risks and controls within the context of an overall ISMS. ISO 27000 is similar in design to management systems for Quality Assurance (QA) (ISO 9000).

ISO 27000's objective is to provide a model for establishing, implementing, operating, monitoring, reviewing, maintaining, and improving an ISMS. Further, the design and implementation of an organization's ISMS is influenced by their needs and objectives, security requirements, the process employed, and the size and structure of the organization.

ISO 27000 is broad in scope and not limited to privacy, confidentiality, and information technology or technical security issues. It is applicable to organizations of all shapes and sizes. Organizations are encouraged to assess their information security risks, then implement appropriate information security controls according to their needs, using ISO 27000 guidance and suggestions where relevant. Given the dynamic nature of information security, ISMS incorporates continuous feedback and improvement activities, seeking to address changes in the threats, vulnerabilities, or impacts of information security incidents.

A key ISO 27000 aspect is the coverage of the security Certification and Accreditation (C&A), which involves systems verification/testing to formally indicate customer product acceptance, from a security perspective. Showing how well security testing was performed and that the test results accurately indicate that the requirements were met is part of C&A. This may sound like acceptance testing, and in a way it is. However, the main difference is the structured C&A approach with its emphasis on security—typically, a customer's Chief of Information Office (CIO) or Security Officer signs the C&A document rather than having a Project Manager or Contracting Officer make formal acceptance. Thus, for some contracts, there may be two levels of acceptance (i.e., acceptance by a Contracting Officer and acceptance by a CIO or Security Officer).

IEEE

IEEE 1012, *IEEE Standard for Software Verification and Validation*, addresses projects from an oversight perspective (International Organization for Standardization, 2004). The main difference between V&V and QA is that V&V emphasizes engineering and risk management, whereas QA normally emphasizes processes and occasionally products. V&V is *focused on building the right thing right*—where the *thing* can be either a process or work product. Specifically, verification addresses whether the thing was built right—the "how," while validation addresses whether the right thing was built—the "what." Thus, V&V takes on a more holistic view. (*Note:* for some organizations, the V&V and QA functions could be limited to, or include, independent testing.)

IEEE 1012 provides a list, based on project complexity, of expected input and output products for each project life cycle phase. For testing, IEEE 1012 shows the suggested life cycle for test products and indicates what an independent auditor should look for within the products. Thus, testers and managers have another perspective of what should be in test products.

SEI, Carnegie Mellon University

SEI developed a software development methodology/model known as Capability Maturity Model—Integration (CMMI®) (Chrissie, Konrad, and Shrum 2007), which provides opportunities to avoid or eliminate stovepipes (systems for particular domains rather than systems able to cross domains) and barriers through integrated models transcending disciplines. CMMI for development consists of best practices (covering a product's life cycle from conception through delivery and maintenance) addressing development and maintenance activities applied to products and processes. Emphasis is on the work necessary to build and maintain the total product. CMMI is one of many national and international models that measure the capability of projects and organizations to be successful.

An implemented CMMI provides a full life cycle process showing process area relationships and how testing is an integral part of a project. Of CMMI's 22 process areas, the V&V process areas deal directly with testing. Included in CMMI's V&V are some tips and hints (Chrissie, Konrad, and Shrum 2007).

Validation

Validation demonstrates a product or product component fulfills its intended use when placed in its intended environment.

Tips.

- Requirements development must address requirements validation—i.e., how will each requirement be validated.

- Validation can be an expensive activity, requiring good judgment to select and limit what needs to be validated.
- Requirements validation determines requirements adequacy and completeness.
- Integration tests can address validation-type activities—with an end user present to evaluate the integrated product under different scenarios. Thus, for some product components, product integration and V&V activities may be addressed together. This saves time for testing and helps eliminate test duplication.
- When seeking to validate a yet-to-be-built product component, developing a prototype based on the component's requirement and design should be considered.
- Validation procedures and criteria should address operations, maintenance, and services.
- It is important early in a project to identify and communicate with stakeholders (e.g., customers and end users) about validation methods, so appropriate preparations can begin.

Hints or considerations.

- Only a well-verified work product should be used in validation; otherwise, valuable time may be lost with disruption or rediscovering requirements you already knew.
- Validation helps uncover missing needs that can be incorporated into requirements.
- Any product or product component can benefit from validation. What is selected depends on the issues relating to user needs that pose the highest risk to project success and available resources.
- It is important early in a project to develop requirements to interface with the validation environment.
- Items selected for validation might be shown in a table with columns identifying items to be validated, issues to be investigated, related requirements and constraints, and validation methods. The table might also list the work products to be verified and the verification methods to be used. This table may lead to discoveries about opportunities to combine V&V.
- Preparing for and conducting validation requires coordination and commitment to support the planned validation effort.
- Validation planning can be a challenge (e.g., there is a need for fidelity with the operational environment).
- Validation activities are expensive, thus it is important to maximize learning.
- If there are problems with the validation methods, environment, procedures, or criteria, the planning for validation activity should be reexamined.
- Waiting until the acceptance test to find issues may result in troublesome complications.
- Some CMMI process areas linked to validation are Technical Solution and Project Monitoring and Control.

Verification

Tips.

- Verification is often confused with "testing"; however, verification methods may also include review, auditing, analyses, simulations, demonstrations, and formal methods.
- Requirements volatility can greatly increase verification activities cost and can lead to schedule slips.
- Technical performance parameters help to monitor key characteristics of the emerging product.
- Testing and peer reviews are verification methods (peer reviews focus on getting the product "right" and obtaining data necessary to prevent defects and to improve the process).
- Methods used for each work product may be shown in a table with columns identifying the work product to be verified, requirements to be satisfied, and verification methods to be used.
- Some work products and verification methods may require special facilities and tools, which should be identified and obtained in advance. Some facilities and tools may be "long-lead items" (i.e., they may require months or years to acquire, set up, train on, and test for compliance.
- Re-verification should be considered when planning verification activities.
- Since participants often skip preparation when under schedule pressure, incorporating time for peer review into the development plans increases likelihood of compliance. Peer reviews should be postponed if participants are not ready.

Hints or considerations.

- Verification can be thought of as a "tool" to use throughout the project to help ensure the delivery of quality products to customers.
- Verification methods, environments, and procedures should be determined as early as practical.
- A work product's importance to other life cycle activities (e.g., maintenance, documentation, installation, and operator training) should to be verified.

- Work products that put the project or product at the highest risk should to be identified.
- Comprehensive verification procedures and criteria should be established for nondevelopmental items (i.e., off-the-shelf or re-use products), which are subjecting the project to moderate or high risk.
- Since there is often no one "right" result, the amount of variability that will be allowed from the expected value should be specified.
- Peer reviews should be used to address the artifacts of development, project management (e.g., plans), process management (e.g., process descriptions), and support (e.g., measure definitions).

PMI

PMI's international standard *PMBOK® Guide* (Project Management Institute 2008) allows organizations and project managers to plan, execute, and close projects. Although testing is not a main topic, the *PMBOK® Guide* does address project risk management, which should be part of test planning and test execution. As with ISO 9001, the *PMBOK® Guide* recognizes the need to test vendor products to ensure the received products satisfy agreed-to requirements.

U.S. DoD

The DoD, and other U.S. and world government agencies, has many standards, some of which were developed by other groups (e.g., IEEE). DoD has standards, for example, dealing with the following:

- environmental engineering considerations and laboratory tests, MIL-STD-810F, and
- test procedures for packaging material, MIL-STD-3010.

These DoD standards are not normally considered by systems or software testers, but these standards address key subjects that could be contractual requirements or could address some test concerns. Managers and testers may need to become involved with standards such as these.

Each level of government throughout the world has standards that they recognize and require. As with other groups (e.g., ISO and IEEE), there may be government-recognized standards that have easy-to-find test standards. Governments and corporations also have test requirements and standards that may need to be followed or at least examined. Testers and managers should be aware of this to assure products are properly tested and documented for quick, formal customer acceptance.

Recommendations

Besides looking at other standards to help improve how to perform testing, people need a broad view of what is available to maintain their employability. Many people find a niche in their line of work and limit themselves to this niche because they enjoy the work, the pay, the people they work with, etc. In today's fast-expanding world, this concept may result in loss of employment because people have limited their scope of work. The author witnessed the elimination of job positions when a company lost a contract rebid and the people had not kept up with technology, standards, or corporate thinking. These people (e.g., testers) so specialized in a particular programming language and methodology that they were outdated and no longer employable by the company. It was cheaper to hire college graduates with timely knowledge, but no experience, than to train current employees on new computer languages, methodologies, and technologies.

People and organizations say there isn't enough time in the day to work, have a private life, and to read or train in new areas. Time must be found! One thing testers and managers must find time for is to learn about best practices and new standards being developed or already in place. For example, working on IEEE, ITEA, and ISO committees to develop new standards can provide testers with insight into the future of testing and the better integration of testing into other life cycle phases and ways of doing business. Learning how testing can improve the way an organization operates and deals with current and future customers expands a person's skill set and areas of knowledge and can improve a person's future employment opportunities.

Learning includes—if a person wants to stay in the same field—looking at non-test standards and standards not normally examined, while thinking about how these standards can be applied to the testing and business fields. Looking at an organization's non-test standards could also better relate testing to the bigger organizational picture. If a tester wants to grow within an organization, this helps the person and the organization. For example, testers should be able to communicate better with non-testers by better understanding how an organization operates and speaks. Testers should consider the following questions:

- Do you know your organization's goals and objectives?
- Do you know how your group helps implement these organizational goals and objectives?
- Are you aware of opportunities your organization is looking at and how testing may fit?

Conclusion

The author's intent could have been for all possible standards to be examined to help test organizations and people. But, there are too many standards for this to be fully accomplished. Rather, the goal of this article is to make testers, managers, organizations, and projects aware that more information is available than many people realize. In many situations the mentioned standards help organizations and projects to better pull various life cycle phases and activities together to provide a much better global way of doing business. The author hopes this improves test planning, implementation, and reporting. For example, too often the test schedule is reduced or, because of no test criteria, nobody has defined when to terminate testing. Looking at a broader range of standards will enhance test planning and help ensure effective and efficient testing resulting in improved ROI and stakeholder acceptance. Having a customer say, "Oops, we did not want that," is not what people or organizations want to hear, especially late in the test system life cycle.

Assuming that a group (including functional groups such as an automobile association) already has a complete test process in place, then spending time to examine other outside test standards could improve existing or future test processes and test organizations. This review could result in opening the test doors to other options or ways of thinking (e.g., how life cycle phases and activities should interact to provide successful organizations and projects). A test organization should also look at expanding its operation to include addressing risks through testing of contingency plans, evaluation of test environments, and other nonconventional testing opportunities.

This article provides examples of standards to improve the way organizations and people prepare for and implement test processes. The article also illustrates how testing can and does impact other life cycle phases and activities. This impact must be carefully considered while looking at test standards when an organization plans or develops its tests methodology, test tools, test environment, test resources, test staff, and communications.

Examining non-test standards and other test standards could be part of a tester's training curriculum during bench or down time. Expanded awareness and incorporation of standards into the tester's tool suite will benefit the corporation and their clients through reduced risks, increased efficiencies, and work products that receive broader adoption and acceptance. In return, this should help testers broaden their knowledge base and provide an advancement opportunity both inside and outside the testing environment. ❑

George Jackelen has worked for several industries in support of the U.S. Federal Government after spending 20 years in the U.S. Air Force performing various computer-related activities. Since 1996, he has been working as a manager and engineer in the field of independent verification and validation (IV&V). He has performed work with the Institute of Electronic and Electrical Engineers, Inc. (IEEE) and the Organization for International Standardization (ISO) to develop and review standards. George has helped companies achieve ISO 9001 and 20000 certification and become cognizant of the Software Engineering Institute's Capability Model—Integration (SEI CMMI) level 3. He has also worked on the 4th edition of the Project Management Institute's (PMI's) Guide to the Project Management Body of Knowledge. George has written 10 published articles dealing with computers and speaks at different conferences and symposia. E-mail:georgej@scigrp.com

References

Chrissie, M. B., M. Konrad, and S. Shrum. 2007. *CMMI, Guidelines for Process Integration and Product Improvement, Version 1.2.* 2nd ed. Boston: Addison-Wesley.

Institute for Electrical and Electronics Engineers, Inc. *IEEE 1012, Software Verification and Validation*. Geneva, Switzerland: International Organization for Standardization.

International Organization for Standardization. *ISO 9000 Series, Quality Management System*. Geneva, Switzerland: International Organization for Standardization.

International Organization for Standardization. *ISO 12207, Standard for Information Technology—Software Life Cycle Processes*. Geneva, Switzerland: International Organization for Standardization.

International Organization for Standardization. *ISO 20000 Series, Information Technology Infrastructure Library (ITIL)*. Geneva, Switzerland: International Organization for Standardization.

International Organization for Standardization. *ISO 27000 Series, Information Technology—Security Techniques—Information Security Management Systems—Overview and Vocabulary*. Geneva, Switzerland: International Organization for Standardization.

Project Management Institute. *A Guide to the Project Management Body of Knowledge—PMBOK® Guide*. Newton Square, PA: Project Management Institute.

The
ITEA Journal
of Test and Evaluation

Rate Card #3
Effective November 2009

Rates

Black & White	Base Rate	4X
Full Page	$840	$788
½ Page	$541	$504
¼ Page	$441	$410
2 Page Spread	$2,011	$1,806

4 Color Rate	Base Rate	4X
Full Page	$1,680	$1,575
½ Page	$1,208	$1,123
¼ Page	N/A	N/A
2 Page Spread	$3,455	$3,150

Premium Position

Cover 2 (ISFC)	$1,920
Cover 3 (ISBC)	$1,920
Cover 4 (Back Cover)	$2,300

2010 Production Schedule

Issue	Space Reservation	Ad Material Deadline
March	1/13/10	01/27/10
June	4/14/10	4/28/10
September	7/16/10	8/2/10
December	10/13/10	10/27/10

Mechanical Specifications

Ad Sizes	Non Bleed W x H	Bleed W x H
Full Page	6 ½ x 9	8 ¾ x 11 ¼
Half Page (H)	6 ½ x 4 ½	
Half Page (V)	3 x 9	
Qtr Page	3 x 4 ½	
2 Page Spread	15 x 9	17 ¼ x 11 ¼

Trim Size: Journal Trims ⅛" off top, bottom, and outside edge. Final trim size of publication is 8 ½" x 11". Live matter should be a minimum of ½" inside trimmed edges, and a minimum of ½" should be allowed for the bind.

Binding: Perfect

Paper Stock: Inside Pages: 60 lb gloss Covers: 100 lb gloss

Insert Requirements: All inserts must be furnished printed, folded, ready for binding. Inserts must be folded to 8 ¾" by 11 ⅛". Folded inserts will be placed between signatures.

Electronic Files: If you are supplying the images on disk, please include a directory of files contained on the disk and a hard copy printout. Supplied hard copy must match the file. Digital color files must be submitted on CMYK mode. Digital color proofs supplied with the file must be output in CMYK.

Halftone Screen: 133–175 line screens are acceptable.

Resolution: Graphics should be a minimum of 350 dots per inch (dpi) or higher.

Advertising Sales
Onkar Sandal
1-800-627-0326 x 218
osandal@allenpress.com
810 East 10th Street
Lawrence, KS 66044
Fax: 785-843-1853

Format: We accept TIFF or EPS format for both Macintosh and PC platforms. We also accept files in the following Native Application File Formats:

Adobe Acrobat (.pdf) (use Press settings under Job Options)
Adobe Photoshop (.psd) Illustrator (.ai)
Macromedia FreeHand (.fh) Corel Draw (.cdr)
Canvas (.cvs) Pagemaker (.pmd)
InDesign (.indd)
QuarkXPress (.qxd)

Fonts: Include the screen and printer font files for any text. PC or Macintosh versions of Adobe PostScript fonts should be used. Do NOT use True Type fonts. Do NOT use system "bitmap" fonts.

Data Transfer: We can accept the following media:

Macintosh/PC Platforms
Zip Disks
CD_ROM

We can also accept or retrieve files via File Transfer Protocol (FTP). If you will be posting to the FTP site please follow the following directions:

1. The user should FTP to: ftp.allenpress.com
2. Login with the username of anonymous
3. Use your email address as the password
4. Place files in the inbox

Compression: Large files should be compressed with Stuffit or Winzip if possible.

Disclaimer: All claims for errors in advertisements must be made in writing and received within ten days of publication and will be considered only for the first insertion of the advertisement.