IDA | RESEARCH NOTES

WELCH AWARD 2017

Summer 2018

IDA is the Institute for Defense Analyses, a nonprofit corporation operating in the public interest.

IDA's three Federally Funded Research and Development Centers provide objective analyses of national security issues and related national challenges, particularly those requiring extraordinary scientific and analytic expertise.

The summaries in this edition of *IDA Research Notes*, as well as the original publications on which the summaries are based, were written by researchers in the following IDA research groups. The directors of these groups would be glad to respond to questions on topics related to their work.

Cost Analysis and Research Division (CARD)
Dr. David J. Nicholls, Director, 703.575.4991, dnicholl@ida.org

Information Technology and Systems Division (ITSD)
Dr. Margaret E. Myers, Director, 703.578.2782, mmyers@ida.org

Intelligence Analyses Division (IAD)
RAdm. Richard B. Porterfield, Director, U.S. Navy (retired). 703.578.2812, rporterf@ida.org

Joint Advanced Warfighting Division (JAWD)
Dr. Daniel Y. Chiu, Director, 703.845.2439, dchiu@ida.org

Operational Evaluation Division (OED)
Dr. Robert R. Soule, Director, 703.845.2482, rsoule@ida.org

Science and Technology Division (STD)
Dr. Leonard J. Buckley, 703.578.2800, lbuckley@ida.org

Science and Technology Policy Insitute (STPI)
Dr. Mark J. Lewis, Director, 202.419.5491, mjlewis@ida.org

Strategy, Forces and Resources Division (SFRD)
Adm. John C. Harvey, Director, U.S. Navy (retired), 703.575.4350, jharvey@ida.org

System Evaluation Division (SED)
Dr. Steve Warner, Director, 703.845.2096, swarner@ida.org

We dedicate this issue of *IDA Research Notes* to the memory of Keith B. Meador, who designed the cover and prepared the initial layout.

The Larry D. Welch Award is named in honor of former IDA president and U.S. Air Force (USAF) Chief of Staff, General Larry D. Welch, USAF (retired). The annual award recognizes IDA researchers who exemplify General Welch's high standards of analytic excellence through their external publication in peer-reviewed journals or other professional publications, including books and monographs.

The articles in this issue of *IDA Research Notes* are summaries derived from the winner and finalists in the 2017 Larry D. Welch Award competition. In addition, the Welch Award Selection Committee named six other nominated publications as being noteworthy given the quality of research they reflect.

Names in bold type have current or former affiliations with IDA as researchers, members of division management, or consultants.The original publications that were nominated are cited, along with a link where available.[1]

The 2017 Welch Award winner is an article based on IDA research for the Office of Naval Research, "Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review," by **James A. Kulik** and **J. D. Fletcher** (*Review of Educational Research* 86, no. 1: 42–72, https://doi.org/10.3102/0034654315581420).

Enterprise Level Security, Securing Information Systems in an Uncertain World, by **William R. Simpson** (Boca Raton: CRC Press, Taylor & Francis Group, 2016), derives from 14 years of IDA research for the Chief Information Officers of the U.S. Air Force and the Department of Defense.

"Inflation Adjustments for Defense Acquisition," by **Stanley A. Horowitz**, **Bruce R. Harmon**, and **Daniel B. Levine** (*Defence and Peace Economics* 27, no. 2: 231–57, https://doi.org/10.1080/10242694.2015.1093758), is based on IDA research for Cost Assessment and Program Evaluation in the Office of the Secretary of Defense.

IDA research for the Army and the Joint Staff was the origin of "Mishap Analysis for Brownout Rotorcraft Enhancement System (BORES) Analysis of Alternatives (AoA)," by **William L. Greer** and **Joshua A. Schwartz** (*AHS [American Helicopter Society] International 71st Annual Forum Proceedings*, May 5–7, 2015, Virginia Beach, VA).

"Political Instability in Zimbabwe," by Ambassador (retired) **George F. Ward, Jr.** (Council on Foreign Relations, Center for Preventive Action Contingency Planning Memorandum No. 53, March 2015), was based on research conducted for IDA's Africa program.

[1] IDA assumes no responsibility for the persistence of URLs for external and third-party internet websites referred to in this publication. Further, IDA does not guarantee the accuracy or appropriateness of these websites' content now or in the future.

Multiple IDA projects related to the verification, validation, and accreditation of models and simulations for the Defense Threat Reduction Agency and other government sponsors led to “Prioritization Framework: A Step Toward Cost-Effective VV&A,” by **Susan K. Numrich**, **Robert R. Zirkle**, **James R. Ayers**, **Forrest R. Smith**, and **Anna Vasilyeva** (Interservice/Industry Training, Simulation, and Education Conference 2016, Paper No. 16045).

“Reshaping Space Policies to Meet Global Trends,” by **Bhavya Lal** (*Issues in Science and Technology: Reshaping Space Policies to Meet Global Trends* 32, no. 4, Summer 2016, http://issues.org/32-4/reshaping-space-policies-to-meet-global-trends/), was derived from multiple IDA projects related to space policy issues conducted for different government agencies.

IDA’s technical and statistical analyses for the Director of Operational Test and Evaluation in the Department of Defense formed the basis of “Statistical Approach to Operational Testing of Space Fence,” by **Daniel L. Pechkis**, **Nelson S. Pacheco**, and **Tye W. Botting** (*IEEE Aerospace & Electronics Systems Magazine* 31, no. 1: 30–39, https://doi.org/10.1109/MAES.2016.150176).

“Strategic Material Shortfall Risk Mitigation Optimization Model,” by **James S. Thomason**, **D. Sean Barnett**, **James P. Bell**, **Jerome Bracken**, and **Eleanor L. Schwartz** (*Military Operations Research* 20, no. 4: 5–18, http://www.mors.org/Publications/MOR-Journal), is based on IDA research for the Defense Logistics Agency.

NOTEWORTHY

“A Bayesian Approach to Evaluation of Operational Testing of Land Warfare Systems,” by Lee Dewald, Sr., **Robert Holcomb**, Sam Perry, and **Alyson Wilson** (*Military Operations Research* 21, no. 4: 23–32, http://www.mors.org/publications), was derived from IDA’s technical and statistical analyses for the Department of Defense Director, Operational Test and Evaluation.

“A Trajectory for Homeland Ballistic Missile Defense,” by **Joseph T. Buontempo** (*Defense & Security Analysis* 31, no. 2, 99–109, 2015, https://doi.org/10.1080/14751798.2015.1014157), was based on knowledge gained from multiple IDA projects for the Missile Defense Agency and other government sponsors.

IDA analyses for the U.S. Army Corps of Engineers (USACE) led to “An Integrated Approach for Physical and Cyber Security Risk Assessment: The USACE Common Risk Model for Dams,” by **James D. Morgeson**, **Jason A. Dechant**, and Yazmin Seda-Sanabria (Association of State Dam Safety Officials [ASDSO] Dam Safety 2016 Conference, in Philadelphia, PA, September 11–15, 2016).

Research in the area of computational analyses resulted in “Rapid and Semi-analytical Design and Simulation of a Toroidal Magnet Made With YBCO and MgB_2 Superconductors,” by **Ivo K. Dimitrov**, Xiao Zhang, Vyacheslav Solovyov, Oleg Chubar, and Qiang Li (*IEEE Transactions on Applied Superconductivity* 25, no. 5: 5701208, https://doi.org/10.1109/TASC.2015.2448455).

“Regularization for Continuously Observed Ordinal Response Variables with Piecewise-Constant Functional Covariates,” by Matthew R. Avery, **Laura J. Freeman**, **Mark A. Orndorff**, and **Timothy J. Robinson** (*Quality and Reliability Engineering International*, June 2016), was based on IDA’s operational and statistical analyses for the Director of Operational Test and Evaluation.

IDA research for the Joint Staff was the basis of “Violent Nonstate Actors with Missile Technologies: Threats Beyond the Battlefield,” by **Mark E. Vinson** and **John Caldwell** (*Joint Forces Quarterly* 80, 1st Quarter 2016: 116–123).

Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review

J. D. Fletcher and James A. Kulik

Military operations succeed or fail depending on the knowledge and skill of the soldiers, sailors, airmen, and marines who carry them out. However, the rapidly increasing technical complexity of military operations is raising the level of training needed to perform them. Research has found that one-on-one tutoring adapted to the specific needs, capabilities, and background of individual learners substantially increases learning well beyond that typically provided by classroom instruction. Unfortunately, training of this sort, delivered through the use of one-on-one human tutoring, is, except for rare instances, unaffordable. Nonetheless, it may become practicable through the use of computers employing machine intelligence to provide adaptive, individualized tutorial instruction. This article reviews efforts to build these intelligent computer-based systems and a recent meta-analysis to determine their effectiveness.

Our evaluations found that digital tutors typically raise student performance well beyond the level of conventional classes and even beyond the level achieved by students who receive instruction from other forms of computer tutoring or from human tutors.

Adapting to the Learner

William James, a founder of modern cognitive psychology, stated the following as his First Principle of Perception: "Whilst part of what we perceive comes through our senses from the object before us, another part (and it may be the larger part) always comes out of our mind" (James 1890/1950, 747). Another founder, E. L. Thorndike, concluded that "the practical consequence of the fact of individual differences is that every general law of teaching has to be applied with consideration of the particular person" (Thorndike 1906, 83).

These observations continue to be supported by empirical research. For instance, Gettinger (1984) found a difference in time to learn of about 5:1 among students in elementary school classrooms, which suggests that while some learners in a classroom have fully mastered material being taught, others are struggling to keep up. One primary cause of this difference appears to be prior learning (e.g., Tobias 2003). It is therefore likely for Gettinger's ratio to increase as the ages and experiences of the individuals doing the learning—including military personnel—increase. Corbett (2001) supported this possibility when he reported that the ratio in time for undergraduates to learn elements of programming in LISP was about 7:1. The problems raised by individual differences in background, temperament, and ability can be eased by some classroom practices, but only partially. The use of classroom instruction continues to present unavoidable limits to efficiency and effectiveness in training and education.

These observations are supported by continuing research and theory that emphasize the idiosyncrasy of perception, cognition, memory, and learning. Bloom's frequently cited article (1984) suggested that one instructor tutoring one learner is vastly more effective than classroom instruction. Subsequent research strongly supports this view, but individual instruction is not affordable except for sensitive and critical activities (e.g., brain surgery and fighter piloting). Military training cannot afford an Aristotle for every Alexander or a Mark Hopkins for the rest of us.

But computers *are* affordable. In fact, following the development of writing, which made the content of learning portable, and the development of books, which made learning content both portable and affordable, computers may bring about a third revolution in the teaching-learning process. Full natural language with its use of metaphors, similes, slang, and other peculiarities may remain beyond the reach of computers for some time, but a considerable range of highly adaptable tutorial dialogue is within reach. For the military and elsewhere, this possibility suggests a vision of computer-based devices (e.g., cellular phones) providing training, aiding performance, and supporting decision making via tutorial dialogues any time and practically anywhere. Aside from algorithms for tutoring and private information about the learner, the subject matter data and information needed for tutoring need not be stored locally. It can be collected as needed from the global information grid and tailored to the background, needs, evolving capabilities, and even interests of the individual learner.

In the context of teaching and learning, classroom instruction is a relatively recent technology. For the last 65,000 years or so, most instruction was provided in one-on-one tutorial dialogues. Like many innovations (e.g., wireless telegraph and horseless carriages), computer-assisted instruction began by layering one technology (programmed learning textbooks) onto another (computers) to provide interactive instruction that is somewhat akin to human tutoring.

Programmed learning is based on frames like the one shown in Figure 1. It is easy to write computer code to program these frames and programmed learning is still in common use today. Reviews found it to be moderately superior to classroom learning (Kulik, Cohen, and Ebeling 1980). However, frames require considerable human effort (and expense) to compose. Developers must anticipate and prepare for every possible state of the learner and the instructional system, which was found to be impossible—even for something as rudimentary as second-grade subtraction (Barr and Feigenbaum, 1982). Instead, states of the learner and the system might be determined by the computer—in real time and as needed for tutorial instruction. This possibility was a primary motivation for the Department of Defense to fund research and development of digital tutoring (Fletcher 2009; Fletcher and Rockway 1986).

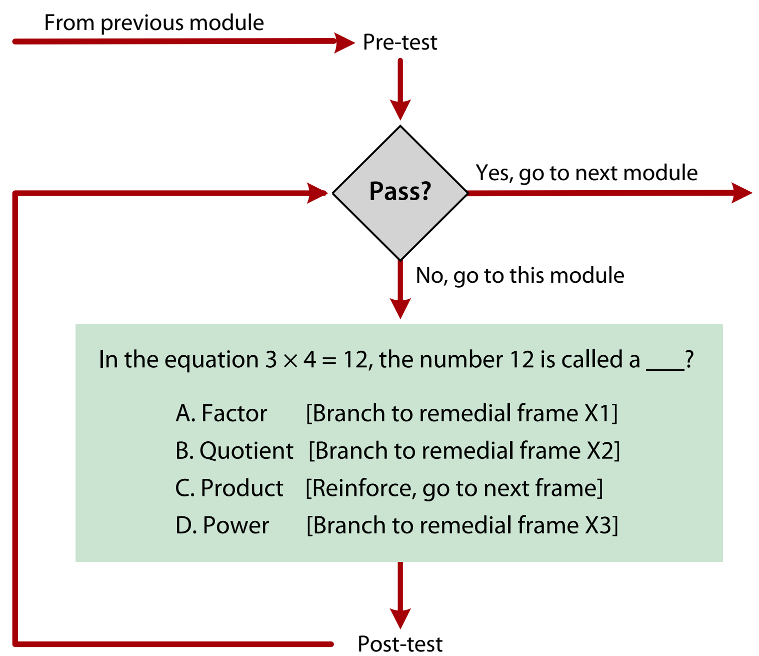


Figure 1. Typical Intrinsic Programming Frame

Digital Tutoring

With support from the Office of Naval Research, Wallace Feurzeig determined that computers could do more than simply mimic programmed textbooks. He developed a computer language and a program (both called Mentor) to prepare learners to perform medical examinations (Feurzeig 1969). It was based on what Carbonell (1970) called information-structure-oriented instruction, using machine intelligence, as opposed to ad hoc frame-oriented instruction based on programmed learning techniques. Further, and in contrast to programmed learning approaches, Mentor allowed mixed initiative dialogues, which meant either the computer or the learner could initiate questions and lines of inquiry. This approach was later the basis for developing the Sophisticated Instructional Environment (SOPHIE) system to train electronic technicians (Brown, Burton, and DeKleer 1982). Example interactions between students and SOPHIE are provided in Figure 2.

Because tutorial dialogues typically include mixed initiative dialogue and information-structure orientation, we treated them as defining characteristics of instructional tutoring systems in our meta-analysis. Instructional tutoring systems typically apply four explicit models: (1) a model of the knowledge and skills to be acquired, (2) a dynamically evolving model of each learner and the learner's acquisition of these knowledge and skills, (3) a model of the instructional techniques that may be used by a specific learner to develop the knowledge and skills derived from the first two models, (4) a model for communication between the learner and the instructional system (Fletcher 1975).

Student: What is the base emitter voltage of the voltage limiting transistor?

SOPHIE: The base emitter voltage of Q5 is -.037 volts.

{{The first student-SOPHIE interaction in this example is a straightforward question and answer. SOPHIE has introduced a fault into the simulated power supply and the student is trying to locate and repair it.}}

Student: What about for the current limiting transistor?

SOPHIE: The base emitter voltage of Q8 is .349 volts.

{{The second student-SOPHIE interaction is a little less straightforward because, in addition to simulating the faulted power supply, SOPHIE must be able to parse the dialogue sufficiently to know that the student is asking for the base emitter voltage of transistor Q8. The meaning of the student's question depends heavily on the context established by the prior question.}}

Student: What should it be?

SOPHIE: In a working circuit, the base emitter voltage of Q8 is .593 volts.

{{The third student-SOPHIE interaction continues its dependence on SOPHIE's understanding of context and additionally requires SOPHIE to extract the answer from both the faulted power supply and one that is operating correctly.}}

Student: Replace Q8.

SOPHIE: I am going to ask you some questions about how Q8 is faulted. Are any junctions shorted?

{{The fourth student-SOPHIE interaction indicates a clear step beyond what Brown et al. [Brown, Burton, and DeKleer 1982] described as a knowledgeable system to what they considered to be an intelligent system. SOPHIE has shadowed the student's solution path, modeled the student's troubleshooting hypotheses, determined that they are incorrect, elected to capture the dialogue initiative back from the student, and is undertaking a series of tutorial interactions intended to lead the student back to a more correct approach to the problem.}}

Source: Foster and Fletcher (2002, 6-15–6-16).

Figure 2. Example of a Digital Tutoring Dialogue from SOPHIE

Meta-Analysis

Meta-analysis is a systematic, statistical technique for reviewing, combining, and summarizing quantitative results from many sources. It is frequently used in medicine and instruction to review the capabilities of a particular technique and provide an overall assessment of its effectiveness. Typically, it calculates statistical probabilities and effect sizes that compare one procedure with another. Statistical results determine the probability that a procedure will be superior in these comparisons (e.g., that a particular medical procedure will cure an ailment or that a particular instructional approach will produce more learning than another). Statistical results follow well-known procedures and identify differences that may be considered probabilistically significant.

However, it is not uncommon for one training procedure to have significant probability of being superior to another, but the difference between the two is so small it has little practical effect. Effect sizes provide a measure, in standard deviations or fractions of standard deviations, of practical significance—how far apart the results from two different approaches are from each other. Effect size is calculated by dividing the difference in results by an estimate of the standard deviation of the population, but discussion over how best to calculate effect size continues. For example, should the estimate be obtained from the standard deviations of all the samples, or should it consider the control group standard deviation alone? Effect sizes reported here are based on pooled standard deviations adjusted for sample size. In the parlance for effect sizes, this measure is known as Hedges's g.

Interpretations of effect sizes vary. A set of interpretations for training and education effect sizes is provided in Table 1. It suggests, in accord with the U.S. Department of Education, that an effect size should exceed 0.25 standard deviations to be worthy of consideration. Bloom (1984) stated that the ultimate goal for effect sizes in education and training research should be 2.00 standard deviations, but researchers in training and education properly celebrate finding an effect size of 0.80.

Results

As in all research, meta-analyses need to leave behind a sufficiently detailed trail of experimental procedures to allow replication. Four steps must be taken and reported clearly in specific detail: (1) identify procedures used to find relevant reports; (2) follow explicit procedures for coding findings from these reports; (3) compile and organize available measures of effectiveness; and (4) use statistical analysis and techniques for combining findings from the reports. Our meta-analysis assembled well over 500 candidate reports and found that 50 of them met the requirements for inclusion that we had established.

Table 1. Overview of Effect Size

Effect Size (ES)	Suggested Designation[a]	50th Percentile (Roughly) Raised To . . .
ES < 0.25	Negligible[b]	60th percentile
0.25 < ES < 0.40	Small	60th–66th percentile
0.40 < ES < 0.60	Moderate	66th–73rd percentile
0.60 < ES < 0.80	Large	73rd–79th percentile
ES > 1.00	Very large	80th percentile and up
ES > 2.00	Bloom's challenge[c]	98th percentile and up

[a] Extended from suggestions by Cohen (1988).

[b] What Works Clearinghouse (2010).

[c] Bloom (1984).

Findings in our meta-analysis of effectiveness of instructional tutoring systems ranged from –0.34 to 3.18. Effect sizes of the larger magnitude were found by Fletcher and Morrison (2014) for the Defense Advanced Research Projects Agency (DARPA) Digital Tutor, which may represent a breakthrough for digital tutoring technology. In 16 weeks, the DARPA Digital Tutor produced U.S. Navy Information System Technicians who scored much higher on tests of both knowledge and troubleshooting skill than other new sailors who had received 35 weeks of classroom training and experienced sailors who averaged 9 years of U.S. fleet experience. The monetary value of avoiding many years of on-the-job training is substantial. The operational value is likely to be larger, but it is more difficult to quantify—the loss of a Navy ship due to information technology failure is conceivable. Results of the DARPA Digital Tutor assessment were outliers for the meta-analysis and were Winsorized—a method to adjust for the statistical effect of extreme data points—by setting the values for its two upper outliers at the 95th percentile and setting the values for its two lowest outliers at the 5th percentile.

With our Winsorized data set, the median effect size was 0.66 overall, and the average effect size was 0.61. Roughly, this suggests an improvement of 50th percentile students to the 75th percentile. These findings are comparable to those of other reviews of digital tutoring techniques (e.g., VanLehn 2011). Our analysis suggests that instructional tutoring systems can provide unusually effective instruction. Students who received intelligent tutoring outperformed students from conventional classes in 46 (92 percent) of the 50 controlled evaluations. The improvement in learning was large enough to be considered statistically significant in 39 (78 percent) of the 50 studies.

Our evaluations found that digital tutors typically raise student performance well beyond the level of conventional classes and even beyond the level achieved by students who receive instruction from other forms of computer tutoring or from human tutors. Kulik and Kulik (1991) found an average effect size of 0.31 in 165 studies of computer-assisted instruction that did not at the time include digital tutoring. Digital tutoring gains are about twice that. Digital tutoring systems may also produce more learning than human tutoring, which typically raise student test scores about 0.40 standard deviations over control level (Cohen, Kulik, and Kulik 1982).

In conclusion, our meta-analytic findings, especially recent results showing effect sizes in excess of 3.00 with the DARPA Digital Tutor, suggest substantial improvements in the ability to provide education and training for military personnel and others. By accelerating learning and the acquisition of expertise, such improvements are likely to yield substantial monetary (Cohn and Fletcher 2010) and operational benefits.

References

Barr, A., and E. Feigenbaum, eds. 1982. "BUGGY." In *Handbook of Artificial Intelligence, Volume 2*, 279–282. Stanford: HeurisTech Press.

Bloom, B. S. 1984. "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring." *Educational Researcher* 13, no. 6: 4–16.

Brown, J. S., R. R. Burton, and J. de Kleer. 1982. "Pedagogical, Natural Language and Knowledge Engineering in SOPHIE I, II, and III." In *Intelligent Tutoring Systems,* edited by D. Sleeman and J. S. Brown, 227–282. New York: Academic Press.

Carbonell, J. R. 1970. "Al in CAI: An Artificial Intelligence Approach to Computer-Assisted Instruction." *IEEE Transactions on Man-Machine Systems* 11, no. 4: 190–202.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences.* 2nd edition. Hillsdale: Lawrence Erlbaum Associates.

Cohen, P. A., J. A. Kulik, and C.-L. C. Kulik. 1982. "Educational Outcomes of Tutoring: A Meta-Analysis of Findings." *American Educational Research Journal* 19: 237–248.

Cohn, J., and J. D. Fletcher, 2010. "What Is a Pound of Training Worth? Frameworks and Practical Examples for Assessing Return on Investment in Training." *Proceedings of the InterService/Industry Training, Simulation and Education Annual Conference.* Arlington, VA: National Training and Simulation Association.

Corbett, A. T. 2001. "Cognitive Computer Tutors: Solving the Two-Sigma Problem." In *User Modeling 2001: 8th International Conference, UM 2001, Sonthofen, Germany, July 13–17, 2001*, *Proceedings*, edited by M. Bauer, P. J. Gmytrasiewicz, and J. Vassileva, 137–147. Berlin: Springer-Verlag.

Feurzeig, W. 1969. *Computer Systems for Teaching Complex Concepts* (BBN Report 1742). Cambridge: Bolt Beranek & Newman, Inc. http://www.dtic.mil/get-tr-doc/pdf?AD=AD0684831.

Fletcher, J. D. 1975. "Modeling the Learner in Computer-Assisted Instruction." *Journal of Computer-Based Instruction* 3, no. 1: 118–126.

———. 2009. "Education and Training Technology in the Military." *Science* 323, no. 5910: 72–75. https://doi.org/10.1126/science.1167778.

Fletcher, J. D., and J. E. Morrison, 2014. *Accelerating Development of Expertise: A Digital Tutor for Navy Technical Training* Alexandria: Institute for Defense Analyses. Draft Final, Document D-5358.

Fletcher, J. D., and M. R. Rockway. 1986. "Computer-Based Training in the Military." In *Military Contributions to Instructional Technology*, edited by J. A. Ellis, 171–222. New York: Praeger Publishers.

Foster, R. E., and J. D. Fletcher. 2002. "Computer-Based Aids for Learning, Job Performance, and Decision Making in Military Applications: Emergent Technology and Challenges." Presented at the RTO HFM Symposium, *The Role of Humans in Intelligent and Automated Systems*, Warsaw, Poland, October 7–9, 2002, 6-1–6-24. RTO-MP-088.

Gettinger, M. 1984. "Individual Differences in Time Needed for Learning: A Review of Literature." *Educational Psychologist* 19, no. 1: 15–29. https://doi.org/10.1080/00461528409529278.

James, W. 1890/1950. *Principles of Psychology: Volume I.* New York: Dover Press.

Kulik, C.-L. C., and J. A. Kulik. 1991. "Effectiveness of Computer-Based Instruction: An Updated Analysis." *Computers in Human Behavior* 7, nos. 1–2: 75–94. https://doi.org/10.1016/0747-5632(91)90030-5.

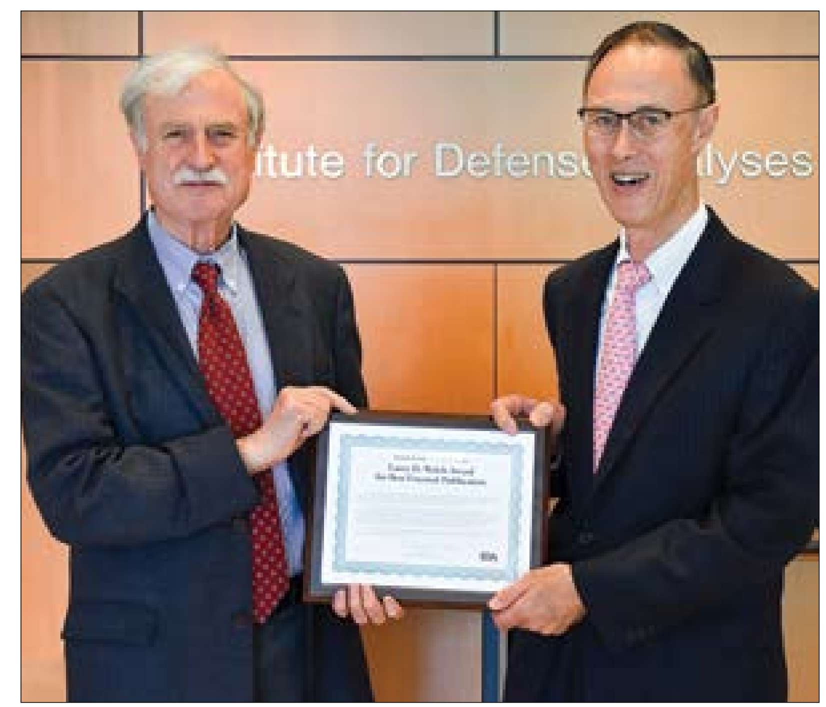
Kulik, J. A., P. A. Cohen, and B. J. Ebeling. 1980. "Effectiveness of Programmed Instruction in Higher Education: A Meta-Analysis of Findings." *Educational Evaluation and Policy Analysis* 2, no. 6: 51–64.

Thorndike, E. L. 1906. *Principles of Teaching.* New York: A. G. Seiler & Company.

Tobias, S. 2003. "Extending Snow's Conceptions of Aptitudes." [Review of the book *Remaking the Concept of Aptitude: Extending the Legacy of Richard E. Snow*, edited by L. J. Cronbach]. *Contemporary Psychology* 48, no. 3: 277–279. https://doi.org/10.1037/00078.

VanLehn, K. 2011. "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems." *Educational Psychologist* 46, no. 4: 197–221. https://doi.org/10.1080/00461520.2011.611369.

What Works Clearinghouse. 2010. *WWC Intervention Report: High School Math, Carnegie Learning Curricula and Cognitive Tutor Software.* Washington, DC: U.S. Department of Education.



J. D. (Dexter) Fletcher (left, with IDA President David S.C. Chu) is a Research Staff Member in the Science and Technology Division of IDA's Systems and Analyses Center. He holds a doctorate in educational psychology from Stanford University.

James Kulik (not pictured), an IDA consultant, holds a doctorate in psychology from the University of California at Berkeley.

The original Welch Award–winning publication was published in *Review of Educational Research* 86, no. 1 (March 2016): 42–78, https://doi.org/10.3102/0034654315581420.

Enterprise-Level Security: Securing Information Systems in an Uncertain World

William R. Simpson

Adversaries continue to penetrate U.S. information technology networks, and in many cases, they have infiltrated the online environment, jeopardizing the confidentiality, integrity, and availability of enterprise information and systems. A multitude of network-related incidents have shown that the fortress model of securing information systems—hard on the outside, soft on the inside—falsely assumes that the boundary between hard and soft can prevent all types of penetration. Given this vulnerability of boundaries, network attacks are pervasive, and nefarious code is present even in the face of system sweeps to discover and clean readily apparent malware.

Early calculations show that ELS-enabled applications can save 90–95 percent of recurring man-hours and eliminate up to 3 weeks of time used for access request processing.

Information Security at the Enterprise Level

Members of all branches of the military must have access to the systems and information they require to execute their missions. The current authorization paradigm requires a cadre of highly privileged administrators to maintain user account permissions for every system and data source required. Human errors, delays in request processing, and credential misuse add to the enormous risks these people face daily. Further aggravating the challenges to successful mission execution and future operations is the determined presence of malicious actors in the contested environment.

Enterprise-level security (ELS) is a web-based security architecture designed to select and incorporate technology into a cohesive set of policies and rules for an enterprise information system. The ELS architecture is based on core security tenets that reflect the enterprise's overall goals and security philosophy. From these tenants, requirements for core security operations are derived to support information sharing within and outside the enterprise.

ELS provides application- and data-level security and is a viable, scalable alternative to current access control management. The initial standup of ELS will cost approximately 75 percent of the annual recurring costs for the current process, and will save thousands of system administration man-hours.

The techniques the architecture employs are resilient, secure, extensible, and scalable. ELS has been tested and is mature in its development. ELS has been named as a potential solution to the identity and access management needs of the Department of Defense's Joint Information Environment, and it is ready to become that solution.

Tenets Guide Decisions and Contribute to Security Principles

ELS is a capability designed to counter adversarial threats by protecting applications and data with a dynamic attribute-based access control solution. ELS helps provide a high-assurance environment in which information can be generated, exchanged, processed, and used. ELS design is based on a set of high-level tenets that are the overarching guidance for every decision made, from protocol selection to product configuration and use (see box). From there, a set of enterprise-level requirements are formulated that conform to the tenets and any high-level guidance, policies, and requirements.

The basic tenets used at the outset of the ELS security model are as follows:

0. Malicious entities are present
1. Simplicity
2. Extensibility
3. Information hiding
4. Accountability
5. Specify minimal detail
6. Service-driven rather than a product-driven solution
7. Lines of authority should be preserved
8. Need-to-share as overriding need-to-know.
9. Separation of function
10. Reliability
11. Trust but verify (and validate)
12. Minimum attack surface
13. Handle exceptions and errors
14. Use proven solutions
15. Do not repeat old mistakes

Current paper-driven access control processes for enterprise operations are plagued with ineffectiveness and inefficiencies. Given that tens of thousands of government and military personnel transfer locations and duties annually, delays and security vulnerabilities are introduced daily into operations. ELS mitigates security risks while eliminating much of the system administration required to manually grant and remove user and group permissions to specific applications/systems. Early calculations show that ELS-enabled applications can save 90–95 percent of recurring man-hours and eliminate up to 3 weeks of time used for access request processing. While perimeter-based architecture assumes that threats are stopped at the front gates, ELS does not accept this precondition and is designed to mitigate many of the primary vulnerability points at the application using distributed security architecture. The ELS design addresses five security principles that are derived from the basic tenets:

- *Know the players* by enforcing bilateral, end-to-end authentication.
- *Maintain confidentiality* through end-to-end unbroken encryption (no in-transit decryption/payload inspection).
- *Separate access and privilege from identity* by means of an authorization credential.
- *Maintain integrity* by ensuring that you receive exactly what was sent.
- *Require explicit accountability* by monitoring and logging transactions.

Know the Players

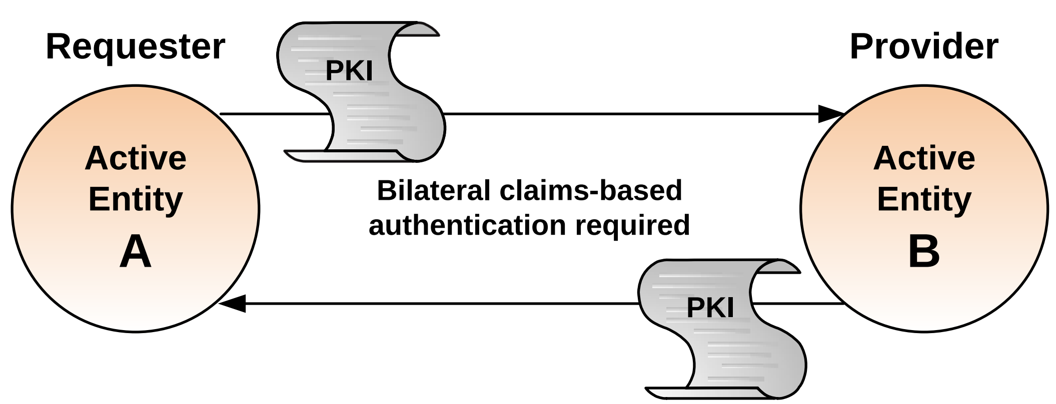
In ELS, the identity certificate is an X.509 public key infrastructure (PKI) certificate.[1] This identity is required for all active entities, both person and non-person (such as a type of service, as shown in Figure 1). PKI certificates are verified and validated. Ownership is verified by a holder-of-key check. Supplemental (in combination with PKI) authentication factors may be required from certain entities, such as identity-confirming information or biometric data.

Maintain Confidentiality

Figure 2 shows that ELS establishes end-to-end Transport Layer Security (TLS) encryption (and never gives away private keys that belong uniquely to the certificate holder).[2]

Separate Access and Privilege from Identity

ELS can accommodate changes in location, assignment, and other attributes by separating the use of associated attributes from the identity. Whenever changes to attributes occur, claims are recomputed based on new associated attributes, allowing immediate access to required mission information. As shown in Figure 3, access control credentials use the Security Assertion Markup Language (SAML).[3] SAML tokens are signed, and the signatures are verified and validated before acceptance. The credentials of the signers also are verified and validated. The credential for access and privilege is bound to the requester by ensuring a match of the distinguished name used in both authentication and authorization credentials.



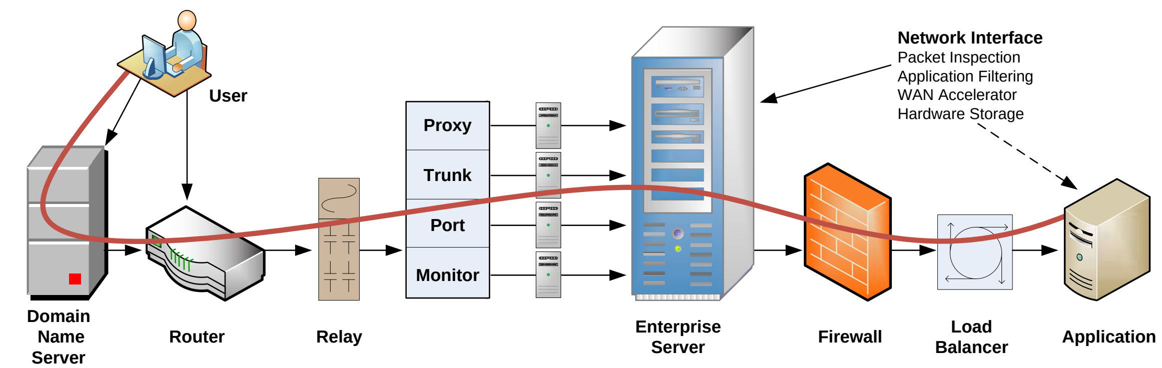
Note: Active Entity A or B may be a user, a web application, a web service, an aggregation service, an exposue service, a token server, or any other entity that can request or provide service.

Figure 1. Bilateral Authentication

[1] The X.509 standard defines the format of public key certificates used in internet protocols. PKI certificates are one of several X.509 certificate types.

[2] The TLS family of Internet Engineering Task Force (IETF) Standards are laid out in a series of Request for Comment (RFC) publications.

[3] The Organization for the Advancement of Structured Information Standards (OASIS) provides an open set of standards for SAML.



Note: WAN stands for Wide Area Network.

Figure 2. End-to-End TLS Encryption

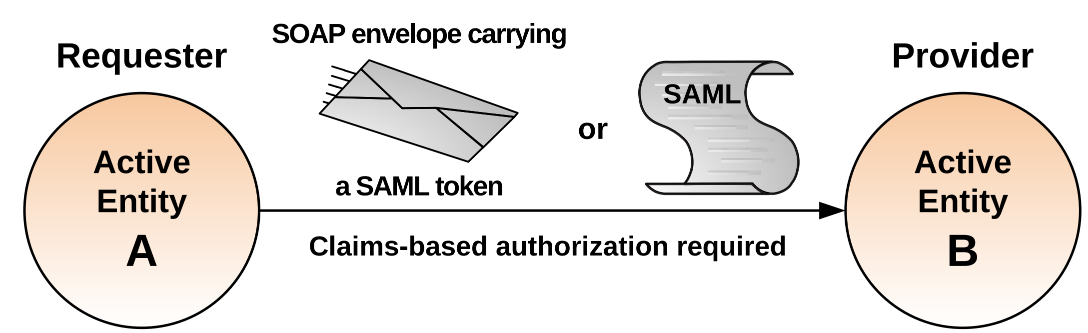
Maintain Integrity

Integrity is implemented at the connection layer by use of end-to-end TLS message authentication codes (MACs) and other integrity measures (Figure 4). Chained integrity, where trust is passed on transitively from one entity to another, is not used since it is not as strong as end-to-end integrity. At the application layer, packages (SAML tokens, etc.) are signed, and signatures are verified and validated.

ELS has been shown to be a viable, scalable alternative to current access control schemas. ELS allows users access without accounts by computing targeted claims for enterprise applications (using enterprise attribute stores and asset-owner-defined claims for access and privilege).

Require Explicit Accountability

As shown in Figure 5, ELS monitors specified activities for accountability and forensics. The monitor files are formatted in a standard way and stored locally. For enterprise files, a monitor sweep agent reads, translates, cleans, and submits to an enterprise relational database for recording log records, periodically or on-demand. Local files are cleaned periodically to reduce overall storage—and to provide a centralized repository for help desk,



Note: SOAP stands for Simple Object Access Protocol.

Figure 3. Claims-Based Authorization

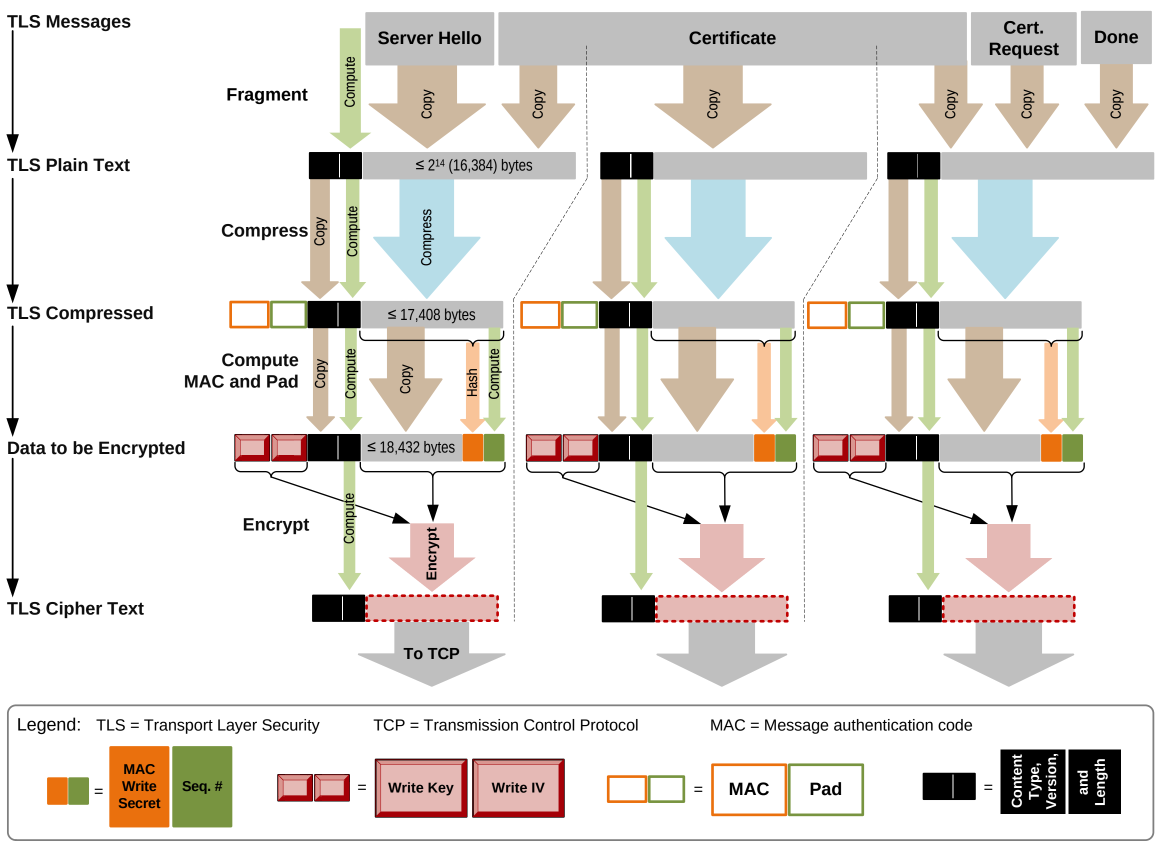


Figure 4. MAC and Other Integrity Measures

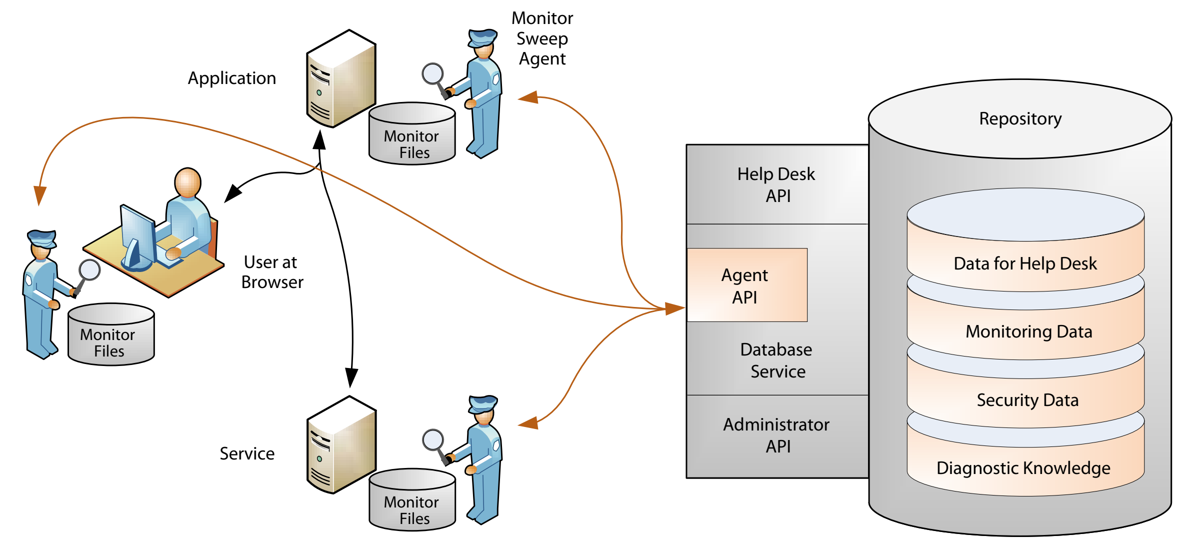
forensics, and other activities. The details of this activity are provided in designated technical profiles (Simpson and Chandersekaran 2010; Chandersekaran and Simpson 2011).

The ELS will reach initial operating capability in a production environment in fiscal year 2018 or fiscal year 2019. Major functionalities have been implemented, and initial penetration testing at the National Cyber Range has found no significant architectural problems. Additional detailed vulnerability testing is planned for future test events.

Authorized users will have immediate access to the application once it is operational. Within the U.S. Air Force alone, system administration requirements will decrease by an estimated 90–95 percent, and user delays for access will dwindle from weeks to hours.

Results of claims-generation tests conducted in late 2013 for 1.2 million unique users show that claims may be generated at 215 million generations per hour. These tests were based on assumptions of 119,614 claims being generated with an average time to generate a claim of 2.0 seconds and an average claim retrieval time (using the ELS process) of 33 milliseconds. These figures are well within the quality of service expected for this user group.

Scaling tests conducted in mid-2012 indicate that a single Secure Token Service (STS) can handle 800



Note: API stands for Application Programming Interface.

Figure 5. Accountability through Centralized Monitoring

SAML tokens per second, which is 50,000 per minute, or 250,000 every 5 minutes. If one STS request every 5 minutes per user is the maximum anticipated (peak sustained rate), then 4 STSs are needed per 1,000,000 users in the enterprise. A planning figure of 10 STSs per 1,000,000 users allows for anticipated redundancy, locality, surges, and load balance latencies.

This is readily achievable and can easily be scaled to larger enterprises. The application handler code to process SAML tokens has been generated for inclusion with .Net and Java applications and services. It has also undergone initial testing.

These test results were documented as the result of a carefully crafted spiral development process that includes:

- Fully encrypted unbroken end-to-end communications (TLS with message authentication codes);
- Bilateral PKI authentication for all enterprise entities;
- SAML-based approaches for access and privilege (the SAML creation and utilization are hardened for vulnerability mitigation);
- Embedded SAML handles for consistency in application;
- Claims-based access and privilege approach, as opposed to attributes and roles;
- Defined federation and delegation processes; and
- Virtualization inspection handlers (in process).

A full implementation began in 2012 with a spiral-based rollout leading to pathfinder applications, testing and evaluation, and application to the Joint Information Environment, which is in process.

Path Ahead

ELS provides a foundation for implementation throughout the Air Force, and the ELS team continues to capture enterprise use cases and define their associated technical solutions. As baselines are established, ELS will be fine-tuned to meet needs identified by evaluation of applications from other military components and environments, such as command and control and tactical.

Development will continue, and with additional testing and feedback, ELS will be hardened and operationalized for enterprise operation. Other elements of ELS, including the handler code installed on servers, will be hardened according to Defense Department policies and provided to developers of new applications and services. Application and service developers will be integrated into the process so that they understand what is expected with ELS, and assistance will be provided through hands-on support and additional documentation of the ELS process.

The ELS web-based security architecture is based on core security tenets and reflects the enterprise's overall goals and security philosophy. The United States must continue to advance its security posture by protecting the applications and data at the source. It is in this vein that ELS was conceived—a superior way to provide secure, scalable access control for the enterprise.

References

Chandersekaran, C., and W. R. Simpson. 2011. "A Multi-Tiered Approach to Enterprise Support Services." Presented at the 1st International Conference on Design, User Experience, and Usability, part of the 14th International Conference on Human-Computer Interaction (HCII 2011), Orlando, Florida, July 2011. https://doi.org/10.1007/978-3-642-21675-6_45.

Simpson, W. R., and C. Chandersekaran. 2010. "An Agent Based Monitoring System for Web Services." Presented at the 16th International Conference on Information Systems Analysis and Synthesis (ISAS 2010), Orlando, Florida, April 6–9, 2010.

William (Randy) Simpson, a Research Staff Member in the Information Technology and Systems Division of IDA's Systems and Analyses Center, holds a doctorate in aerospace engineering from Ohio State University.

This article is derived from *Enterprise Level Security: Securing Information Systems in an Uncertain World* (Boca Raton, FL: CRC Press, Taylor and Francis Group, 2016).

Inflation Adjustments for Defense Acquisition

Stanley A. Horowitz, Bruce R. Harmon, and Daniel B. Levine

Acquisition program managers are required to develop budget projections in terms of then-year dollars. That means they must adjust their future costs for escalating prices. Following a reasonable interpretation of guidance from the Department of Defense Comptroller, program managers have sometimes estimated these costs using a measure of economy-wide inflation, the Gross Domestic Product deflator. But price escalation for a particular kind of defense system may be systematically higher or lower than overall inflation. We used a hedonic cost-estimation approach to develop a price escalation index for fighter aircraft. Applying this index can vastly improve the development of budget requirements compared to using estimates of general inflation.

Because of uncertainty about the validity of existing indexes, we developed a hedonic price index for tactical aircraft that uses data on aircraft characteristics to construct a constant-quality price index.

Uses of Price Indexes in Defense Acquisition

The cost of defense acquisition programs must be adjusted for price increases for two major reasons.

- *Developing budgets.* If the price of a system is expected to rise in the future (escalation), the extent of this rise must be estimated. Using too low an estimate of escalation will lead to budgets that are not adequate to execute the program.
- *Calculating real cost growth for the system.* This requires comparing the actual escalation of system price (relative to the level of general inflation) to the level of escalation that was expected in some base period. Underestimating escalation in the base period will lead to real cost growth. This can subject the program to increased scrutiny and, perhaps, reduction in scope or even termination.

Good estimates of future, program-specific cost escalation require both development of accurate budgets and avoidance of real cost growth.

Comparison of Price Indexes for Aircraft

Several estimating methodologies are in use specifically for aircraft programs. They are as follows:

- The Department of Commerce's Bureau of Economic Analysis (BEA) national defense index for military aircraft tracks the prices the Department of Defense (DoD) pays for military aircraft and major components such as engines and avionics. Costs for systems are obtained from budget exhibits published by the DoD Comptroller supplemented by information from industry literature and general news.

- The Bureau of Labor Statistics (BLS) Producer Price Index, which is published on the BLS website, is calculated for the civilian aircraft production industry from sales price data obtained from commercial producers.
- The Naval Air Systems Command (NAVAIR) index for naval aircraft, which is derived from indexes for airframe, engine, and electronics, is an overall index of flyaway cost for fixed-wing naval aircraft.
- The Gross Domestic Product (GDP) deflator is a chain-weighted price index that BEA calculated as part of the National Income and Product Account (NIPA) from the prices and quantities of the entire U.S. national market basket of goods and services. Published on the BEA website, the GDP deflator is only weakly linked to the growth in prices of military aircraft, since military aircraft are a negligible subset of the entire U.S. market basket.

Historical Growth Rates

Figure 1 portrays the growth of these four quality-constant indexes applied to DoD aircraft systems during the 28-year period 1985–2012, inclusive. The rates were normalized to 1985 = 100 for comparison. The slightly negative growth of the BEA index is especially inconsistent with both growth in the BLS index and the general view of Office of the Secretary of Defense budget analysts that aircraft prices have been rising substantially. Military and civilian aircraft are substantially different, of course, but they are similar enough to raise the question of why their growth rates should be so different. We investigated whether differences in data or methods caused the disparities.

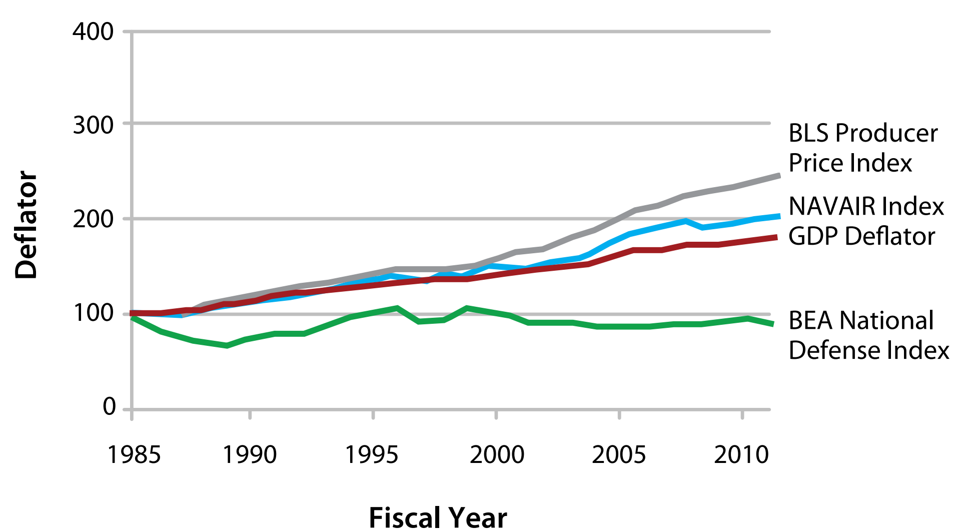


Figure 1. Growth Rate in Price Indexes Often Applied to Aircraft, 1985–2012

Conclusion of Methodological Comparison

The BEA and BLS deflators are generated in mathematically similar ways. The differences in methodologies do not explain the BLS index rising much faster than the BEA index. We suspect that the key difference is the treatment of product improvement.

The price indexes are meant to refer to the price of products of constant quality. In fact, even the same model of aircraft is improved over time. An estimate of the extent to which price increases are due to quality improvements is backed out of raw price data when the index is constructed. We do not have enough information on how these quality adjustments were made to understand their validity or their role in the difference between BEA and BLS price trends for aircraft.

Because of uncertainty about the validity of existing indexes, we developed a hedonic price index for tactical aircraft that uses data on aircraft characteristics to construct a constant-quality price index.

Building a Hedonic Price Index for Tactical Aircraft

Hedonic indexes derive price indexes from regressions that directly relate nominal prices to specific, easily identifiable, quality-related features of the product. In our tactical aircraft case, these features are known with near certainty from legal contracts and from developmental and operational test and evaluation. Table 1 shows the explanatory variables: five quality variables describing the aircraft; two quantity variables describing the number of aircraft produced for use in incorporating the effects of learning and production rate in the procurement process; and a time dummy variable, measured by year of procurement.

Table 1. Explanatory Variables

Quality variables
Empty weight in pounds
Maximum speed in knots
Advanced materials as percentage of structure weight
Dummy variable for 5th-generation aircraft[a]
Dummy variable for short takeoff and vertical landing (STOVL) aircraft[b]
Quantity variables
Cumulative production
Lot size (number of aircraft produced in a year)
Time dummy variable

[a] In our sample, the F–22 and F–35A/B/C fighters are classified as 5th-generation aircraft, which are characterized by stealth, internal weapons carriage, avionics with information fusion, and support of net-centric operations.

[b] In our sample, the AV–8B attack and F–35C fighter are both aircraft with STOVL capability, which is needed for operations from small aircraft carriers and short, unimproved airfields.

Our regression analysis used pooled cross-section and time-series data. The time-series covers the 40 fiscal years from 1973 to 2012, inclusive. Each year other than the base year, 2012, is given a different time dummy in order to calculate a different price index for that year. The cross-sections are the 22 aircraft programs shown in Table 2, consisting of 11 original designs plus 11 derivatives of these original designs from series or block changes.

Table 2. Aircraft Programs

Original Designs	Derivatives (Series or Block Changes)
F–14A	F–14A+, F–14B
F–15A	F–15C, F–15C MSIP, F–15E
F–16A	F–16C Blocks 25/30/50
F/A–18A	F/A–18C Night Attack
A/V–8B	A/V–8B Night Attack, A/V–8B Radar
F/A–18E	–
F–22A	–
F–35C	–
EA–18G	–
F–35A	–
F–35B	–

Result of Hedonic Estimation

Our regression equation from which the hedonic price index was derived had high explanatory power and predictive variables. All had the correct signs and high levels of statistical significance. The R^2 of the equation was 0.97.[1]

Figure 2 compares the trend in the hedonic index with the trends shown in Figure 1. The hedonic index shows a relatively high growth rate that agrees with the perception in the DoD acquisition community that (1) the GDP deflator understates annual quality-constant price increases and (2) the BEA index greatly understates them. This implies that real program growth in the area of tactical aircraft procurement has been less than is generally calculated.

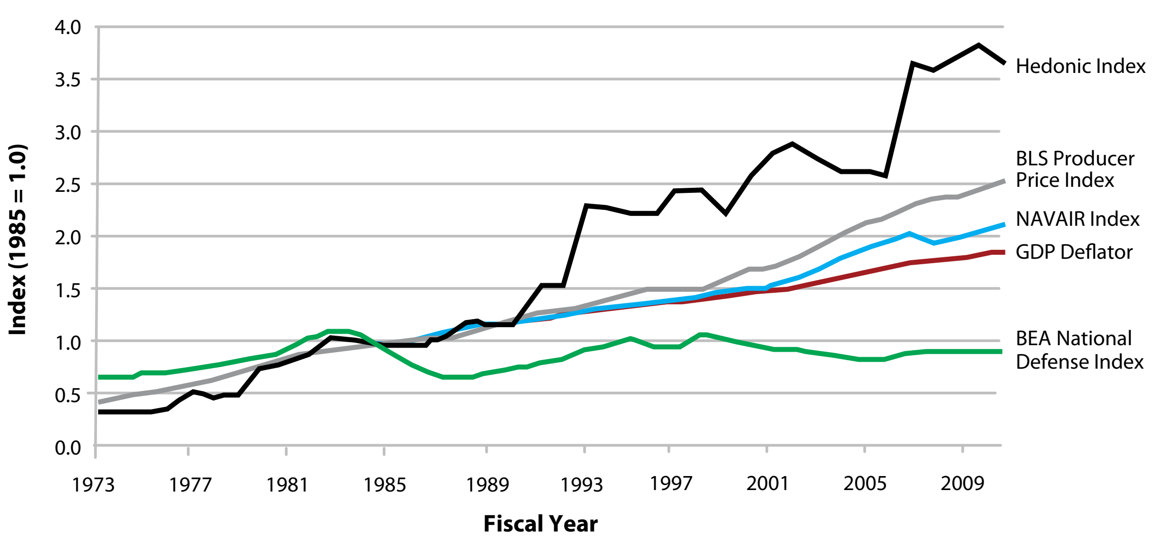
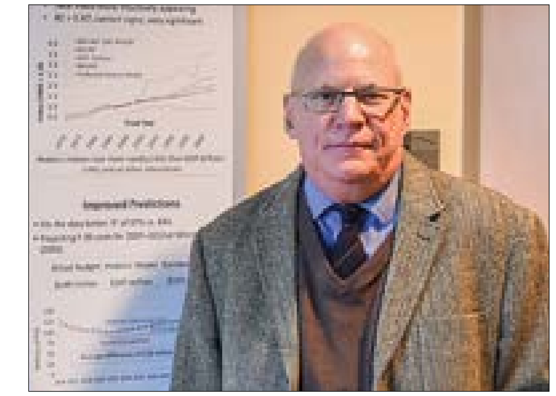


Figure 2. Hedonic and Other Aircraft Price Deflators

[1] R^2, or the coefficient of determination, is a statistical measure of how close data are to the fitted regression line.



Stanley Horowitz (left), an Assistant Director in the Cost Analysis and Research Division (CARD) of the IDA Systems and Analyses Center, holds a master's degree in economics from the University of Chicago.

Bruce Harmon (right), an Adjunct Research Staff Member in CARD, holds a master's degree in economics from the University of Cincinnati.

Daniel Levine (not pictured), a former Adjunct Research Staff Member in CARD, holds a doctorate in physics from the Catholic University of America.

The original article, "Inflation Adjustments for Defense Acquisition," was published in *Defence and Peace Economics* 27, no. 2 (2016): 231–257, https://doi.org/10.1080/10242694.2015.1093758.

Assessment of Brownout Mishaps in Military Rotorcraft

Joshua A. Schwartz and William L. Greer

A degraded visual environment called a brownout occurs when dust, sand, and debris envelop rotorcraft operating close to the ground and aircrews experience spatial disorientation and loss of situational awareness. Collision, crash landing, or dynamic rollover of the affected rotorcraft can result. In 2014, in response to brownout-induced mishaps that occurred during Operation Iraqi Freedom in Iraq and Operation Enduring Freedom in Afghanistan, the U.S. Army conducted an analysis of alternatives for the Brownout Rotorcraft Enhancement System (BORES). To inform the Army's analysis, IDA estimated the costs of future losses of the Army's H–47 and H–60 helicopters based on costs of all U.S. military rotorcraft mishaps for 2000–2013. The overall investment in BORES versus the savings that could be realized by avoiding helicopter losses was one metric the Army used to assess the value of BORES. This article explains the method used to estimate the costs of potential future losses.

The value of the 26 destroyed rotorcraft between 2000 and 2013 due to brownout was approximately $533 million based on military cost documentation. About half of these aircraft losses and three-fourths of the costs were borne by Army helicopters.

Background

Rotorcraft such as helicopters and vertical takeoff and landing aircraft create what is known as downwash, which is the force of air equal to and in the opposite direction of the force the aircraft exerts on the rotor to produce lift. For rotorcraft operating close to the ground over arid desert terrain, downwash can cause dust, sand, and other debris to circulate upwards and envelop the rotorcraft. The resulting degraded visual environment is called a brownout.

Rotorcraft landings and near-ground hovers are particularly vulnerable to brownouts (Figure 1). The rotorcraft's aircrew can experience spatial disorientation and loss of situational awareness, resulting in collision with an obstacle, crash landing, or dynamic rollover.

Approach

Aircraft mishaps are grouped into discrete classes based on a combination of property loss and personnel casualty levels involved. Those involving the highest costs or loss of life are Class A mishaps, which is our focus here. These involve one or more of the following: total rotorcraft loss through destruction, total cost of damages in excess of $2 million, or one fatality or permanent total disability. Class B mishaps are the next most costly and serious mishaps, and Class C and D mishaps are progressively less costly and less serious.

Figure 1. Military Helicopter Brownout

Historical Mishaps

An IDA research team investigated rotorcraft mishaps occurring between 2000 and 2013, a period including many brownout incidents in both Operation Iraqi Freedom (OIF) and Operation Enduring Freedom (OEF). By direction from the Army, we focused only on brownout-induced mishaps and not those caused by other degraded visual environments such as snow, rain, smoke, darkness, fog, smog, flat light, or clouds. Previously published literature (e.g., U.S. Army Program Executive Office, Aviation [2011]) was ambiguous about the numbers of degraded visual environment mishaps caused solely by brownout, so we set out to clarify those distinctions.

We used records maintained by the U.S. Army, Air Force, and Navy Safety Centers. Brownout mishaps were organized into categories that included the extent of property damage, fatalities and injuries, year, location (OIF/OEF or Rest of World [ROW]), branch of the military, aircraft type and model, day or night conditions, and flight phases. Flight-hour data were then used to calculate the average numbers of brownout-induced incidents per 100,000 flight hours, which is the standard rate metric within the aviation safety community. This rate was crucial to projecting future mishaps. We also estimated costs associated with these losses.

Future Mishaps

Next, we examined the possibility of avoiding future brownout mishaps for the Army H–47 and H–60 rotorcraft, the two helicopters in the U.S. Army's BORES analysis of alternatives. Using future inventory projections, we estimated the number of future mishaps that would occur without BORES. Flight-hour assumptions were bracketed by two operational tempos: programmed flight hours and the higher numbers of flight hours experienced in fiscal years (FYs) 2000–2013. The flight-hour environments were also bracketed by ROW exclusively and a combination of ROW plus OIF/OEF.

We also estimated costs for the projected numbers of rotorcraft lost or subject to Class A repairs. Using inflation indices, we calculated all costs in FY 2014 dollars. From this, we determined the approximate break-even cost, defined as the dollar amount where the full unit life-cycle cost of implementing BORES on these aircraft equals the estimated cost of the mishaps prevented.

Selected Findings

Historical Mishaps

The numbers of Class A mishaps attributed to brownouts from 2000 through 2013 are provided in Figure 2. All branches of the U.S. military are included. The stacked bars use the scale on the left and comprise lost aircraft and other (repairable) Class A incidents. The cumulative total numbers of Class A mishaps are also displayed with the overlaid lines using the scale on the right. A spike in mishaps during the initial OIF/OEF operational buildup is clearly shown. For the ROW, by comparison, relatively few brownout mishaps occurred over the same period and, in many years, none at all. Overall, 26 rotorcraft losses occurred among the 53 Class A brownout mishaps recorded between 2000 and 2013. Of these 26 losses, OIF/OEF accounted for 22 of them.

Over two-thirds of the brownout-induced Class A mishaps involved Army helicopters—virtually all in OIF/OEF, as shown in Figure 3 (OIF/OEF on the left bar chart, and the ROW on the right bar chart). The table below each bar chart shows the types of rotorcraft involved. Within each cell, the first number is the number of aircraft destroyed; the second is the number of other Class A mishaps.

Brownout is not the only cause for rotorcraft mishaps. Military records show that of all rotorcraft lost in OIF/OEF between 2000 and 2013, only 17 percent are attributable to brownout. For the ROW, the brownout Class A mishaps represent fewer than 4 percent of all other mishap causes. The total number of Class A brownout mishaps in OIF/OEF and the ROW combined account for about 12 percent of all mishap events.

These brownout mishaps resulted in 6 fatalities (all in OIF/OEF) and 175 injuries (147 in OIF/OEF) across all branches of the military. Of these, U.S. Army rotorcraft were involved in 3 fatalities (all in OIF/OEF) and 107 injuries (88 in OIF/OEF).

The value of the 26 destroyed rotorcraft between 2000 and 2013 due to brownout was approximately

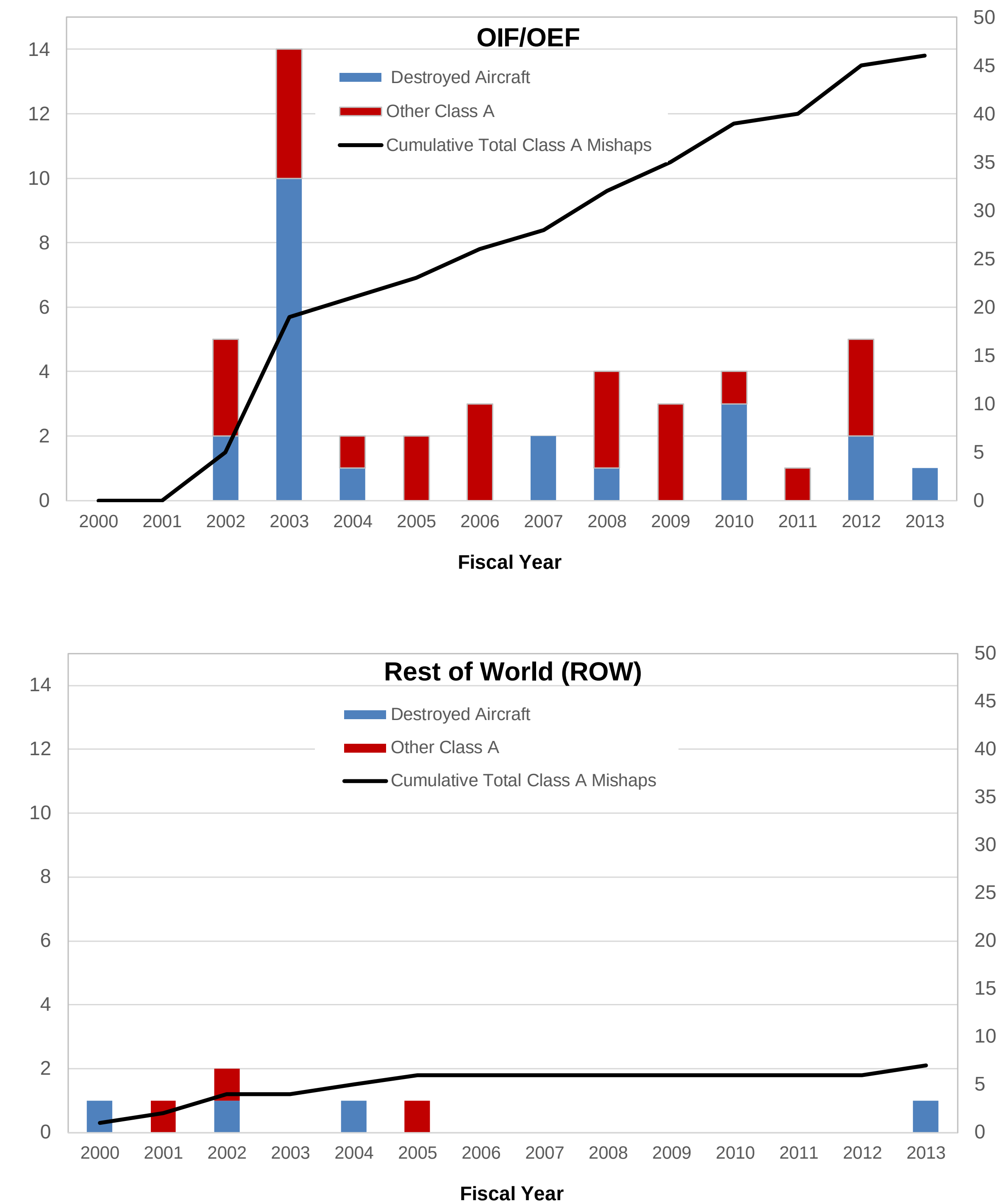
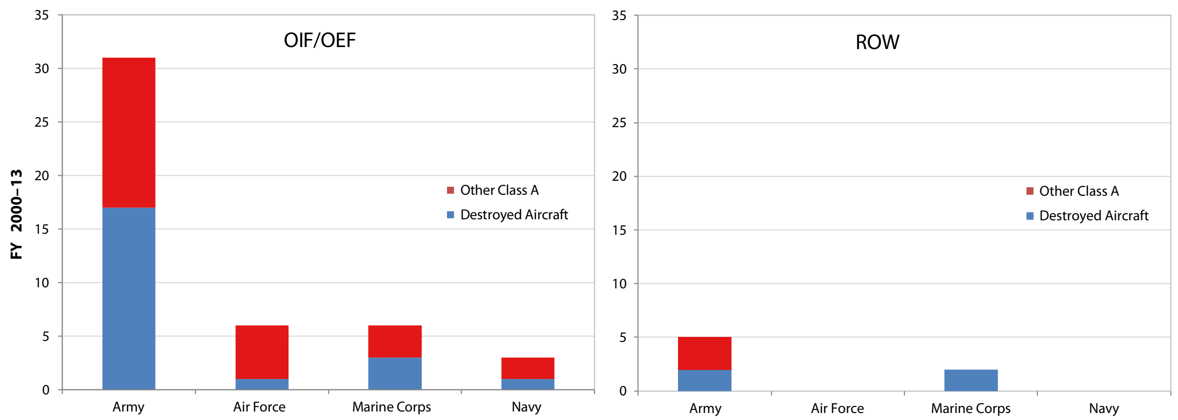


Figure 2. Class A Brownout Summary for OIF/OEF (Top) and ROW (Bottom): All Military Branches by Year



# Destroyed Aircraft + # Other Class A				
Rotorcraft Type	Army	Air Force	Marine Corps	Navy
H–60	2 + 11	0 + 5	0	1 + 2
H–47	7 + 3	0	0	0
H–58	4 + 0	0	0	0
H–64	4 + 0	0	0	0
H–1	0	0	2 + 1	0
H–53	0	1 + 0	1 + 2	0
H–46	0	0	0	0
Total	17 + 14	1 + 5	3 + 3	1 + 2

# Destroyed Aircraft + # Other Class A				
Rotorcraft Type	Army	Air Force	Marine Corps	Navy
H–60	2 + 2	0	0	0
H–47	0	0	0	0
H–58	0 + 1	0	0	0
H–64	0	0	0	0
H–1	0	0	1 + 0	0
H–53	0	0	1 + 0	0
H–46	0	0	0	0
Total	2 + 3	0	2 + 0	0

Note: Figures in red were the focus of our research—Army H–60 and H–47 Class A mishaps.

Figure 3. Total by Military Branch and Rotorcraft Type

$533 million based on military cost documentation. About half of these aircraft losses and three-fourths of the costs were borne by Army helicopters.

Future Mishaps

To estimate the numbers and types of future incidents in the absence of BORES, the researchers combined flight hours, inventories, and mishap rates.

Flight Hours

Between 2000 and 2013, records show that the average number of flight hours per aircraft per year was 224 for the Army's H–60 and 181 for its H–47. By contrast, the programmed flight hours per aircraft is lower: 163 for the H–60 and 128 for the H–47. Therefore, two bounding cases for annual flight-hour rates were considered: a lower bound using programmed values and an upper bound using the more demanding FY 2000–2013 experience.

Inventories

The inventory projections for the H–60 and H–47 through the end of their service lives are shown in Figure 4, which indicates that essentially all H–47s and H–60s will be retired by 2050. Figure 5 shows the number of remaining flight hours each year based on the inventory projections and under the two flight-hour bounding assumptions.

Mishap Rates

Average Class A brownout mishap rates based on historical analyses for the Army H–60 and H–47 combined are displayed in Table 1.

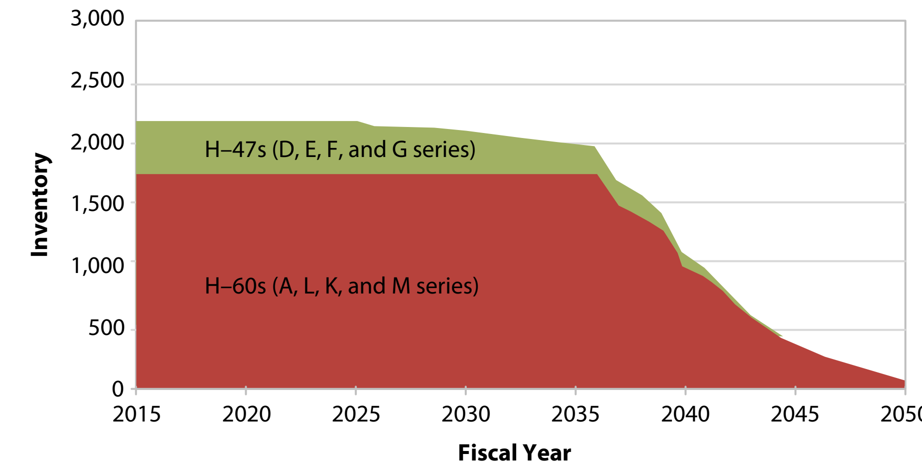


Figure 4. Projected Inventories of Army H–47 and H–60 Helicopters through 2050

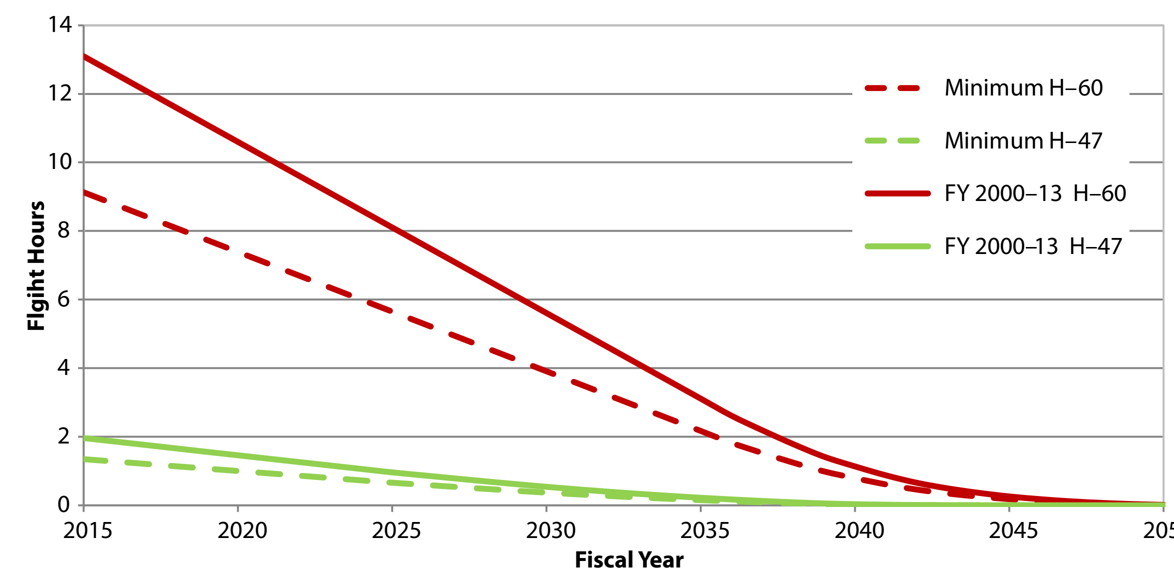


Figure 5. Remaining Flight Hours for Army H–47 and H–60 Helicopters through 2050

Table 1. Brownout Class A and Destroyed Rates, 2000–2013

Location (Category)	Selected Parameter for 2000–13 Timeframe	H–60 + H–47 Combined Value
OIF/OEF	Class A Brownout Rate	0.89
	Loss from Brownout Rate	0.35
ROW	Class A Brownout Rate	0.11
	Loss from Brownout Rate	0.06

Note: All rates represent numbers of mishaps per 100,000 flight hours.

Costs

The Class A brownout mishap costs included rotorcraft losses as well as repairs. Mishap costs were based on the value at the time the mishap occurred, so they were time dependent. Army cost data were used for the H–47 and H–60 rotorcraft acquisition and modifications, the aircraft were given diminishing value over a 35-year period as they aged, and all costs were reported in FY 2014 dollars no matter when the mishaps were expected to take place. The mishap costs did not include casualties because the casualty rates from OIF/OEF and ROW experiences were low and official indemnity estimates are varied and undetermined.[1]

Cases

The baseline case is a hypothetical one for comparison in which BORES is implemented in a single year (2015) for the entire fleet of approximately 2,669 rotorcraft (534 H–47s and 2,135 H–60s). More realistic cases would implement BORES on only a portion of the fleet each year and potentially not even on all fleet aircraft. Given this, we explored three cases that would begin introducing BORES in 2017:

- *Case 1:* BORES introduced into the full H–47 and H–60 fleet to include installation in new-build/remanufactured aircraft as they are delivered and retrofits on existing H–47s and H–60s at a rate of 200 per year.
- *Case 2:* BORES introduced for only one-third of the H–47 and H–60 fleet with the newest/most valuable aircraft covered and proportional effectiveness (one-third of projected losses prevented) at an installation rate of 200 per year. One-third represents approximately the total fleet fraction deployed to OIF/OEF during 2000–2013.
- *Case 3:* Same as Case 2, but with all projected losses prevented through judicious deployment to theater of only BORES-outfitted aircraft.

In Case 3, only the aircraft fitted with BORES would be deployed where brownout conditions might be encountered. In Case 2, aircraft are randomly used where needed, so only one-third would be properly protected against extreme brownout conditions.

Figure 6 shows the break-even unit life-cycle costs for the three cases mentioned above along with the hypothetical 2015 baseline case for comparison. Case 1 had lower break-even values than the baseline, since implementation would take place over a longer span of time, and fewer mishaps could be prevented within the remaining lifetime of the H–47s and H–60s. Case 2 had higher values because only the newer and more valuable aircraft would be affected. In Case 3, fewer aircraft would need to be upgraded to prevent all projected brownout losses, so its break-even value was the highest.

[1] The original IDA paper on which this article is based shows that the inclusion of casualty costs would increase total costs by less than 10 percent and, in some cases, by far less than 10 percent (Greer et al. 2014) for different government value systems.

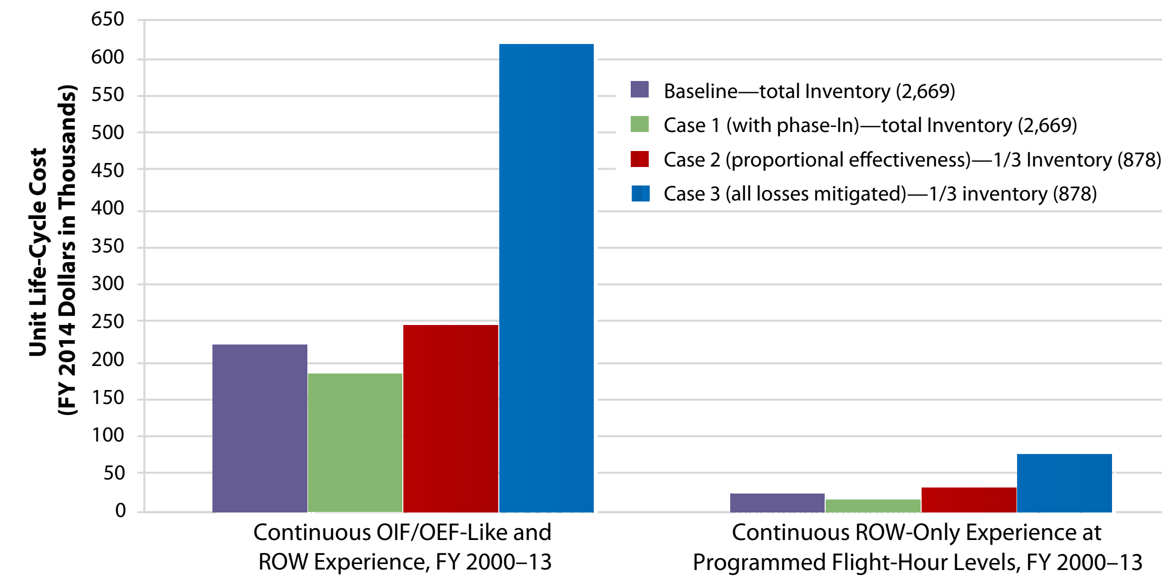


Figure 6. Break-Even Cost Excursions

Summary

Historical records of brownout-induced rotorcraft mishaps for all branches of the military were investigated from 2000 through 2013. An assessment of the cost of aircraft destroyed from brownout in the same 14-year period indicated a cost of around $533 million total in FY 2014 dollars from 26 losses. About half of the losses were Army rotorcraft.

We used these historical data to make projections of numbers and costs of future brownout Class A mishaps involving Army H–47 and H–60 rotorcraft. The costs were expressed in terms of break-even costs, the values at which the cost of BORES equaled the cost of rotorcraft saved. If BORES were introduced starting in FY 2017 for inclusion in new-build or remanufactured aircraft and as retrofits to existing H–47s and H–60s at a rate of 200 per year, a cost-effective BORES life-cycle cost for each of the cases examined should be as follows:

- *Case 1:* less than ~$30,000 per unit if 2000–2013 ROW-only environment prevails and programmed flight hours continue indefinitely (i.e., for lifetime of the fleet).
- *Case 2:* less than ~$245,000 per unit if only one-third of the fleet is outfitted with BORES and if the 2000–2013 operating situation (including more demanding OIF/OEF-like conditions plus ROW) prevails and continues indefinitely. Aircraft are randomly selected for deployments, so one-third would be protected in OIF/OEF-like brownout conditions.
- *Case 3:* ~$600,000 per unit if only one-third of the fleet is outfitted with BORES, and to minimize losses to brownouts, these are the only rotorcraft used in 2000–2013 OIF/OEF-like conditions indefinitely.

Clearly, BORES would prove most cost-effective when only a third of the fleet is outfitted with the new systems and then deployed selectively in OIF/OEF conditions to minimize brownout losses. This approach minimized the numbers of BORES units acquired while maximizing brownout mishap reduction. It is least cost-effective if ROW conditions with minimal brownout environments dominate the future.

Finally, although the focus for this analysis was brownout mishaps, brownouts accounted for only 12 percent of all Class A rotorcraft mishaps between 2000 and 2013. Extension of the analyses to include additional types of rotorcraft mishaps that could be mitigated by a BORES alternative might thereby raise the break-even cost values, making BORES a more attractive alternative. The Army has subsequently expanded its examination of technologies to encompass solutions to mitigate the wider problem of degraded visual environments.

References

Greer, W. L. , A. O. Gallo, B. R. Harmon, B. C. Prindle, J. A. Schwartz, J. Silva, J. W. Stahl, and R. V. Uy. 2014. *Mishap Analysis for Brownout Rotorcraft Enhancement System (BORES) Analysis of Alternatives (AoA).* Alexandria, VA: Institute for Defense Analyses. Paper P-5206, December.

Milluzzo, J., and J. G. Leishman. 2010. "Assessment of Rotorcraft Brownout Severity in Terms of Rotor Design Parameters." *Journal of the American Helicopter Society* 55, no. 3 (July): 032009-1–9. https//doi.org/10.4050/JAHS.55.032009.

U.S. Army Program Executive Office, Aviation. 2011. "Report and Recommendations on Terrain Awareness Aspects of Rotorcraft Mishaps in Degraded Visual Environments (DVE)." August.

Joshua Schwartz (left), a Research Staff Member in the System Evaluation Division (SED) of IDA's Systems and Analyses Center, holds a master of science in aeronautical engineering from Rensselaer Polytechnic Institute.

William (Bill) Greer (right), an Assistant Division Director in SED, holds a doctorate in chemistry from the University of Chicago.

The full article nominated for the Welch Award was published in *AHS [American Helicopter Society] International 71st Annual Forum Proceedings*, May 5–7, 2015, Virginia Beach, Virginia. For details of the work IDA undertook with co-sponsorship from the Joint Chiefs of Staff, see IDA Paper P-5206, *Mishap Analysis for Brownout Rotorcraft Enhancement System (BORES) Analysis of Alternatives (AoA)*," W. L. Greer, A. O. Gallo, B. R. Harmon, B. C. Prindle, J. A. Schwartz, J. Silva, J. W. Stahl, and R. V. Uy.

Political Instability in Zimbabwe: Planning for Succession Contingencies

George F. Ward, Jr.

Political instability and potential violence are ever-present threats in Zimbabwe. The country's nonagenarian president, Robert Mugabe, born on February 21, 1924, has not established a clear succession plan. The nation's economy is perennially weak and vulnerable to shocks. The government suppresses the exercise of fundamental freedoms. Instability in Zimbabwe would be a threat to the region and especially to South Africa, which would lose trade revenue and gain the burden of additional refugees. It would also be a blow to U.S. interests in southern Africa, which are focused on support for good governance, trade, and investment. At the same time, post-Mugabe transition scenarios provide some opportunities that the United States could take advantage of by working with others, notably South Africa and the other countries of the southern African region.

Zimbabwe is richly endowed with human and natural resources that could give it a leading role in shaping the future of the African continent.

Editor's Note: Ambassador (retired) Ward's article was originally written in 2015, and the political situation in Zimbabwe has changed dramatically since then. Robert Mugabe resigned as Zimbabwe's president and former Vice President Emmerson Mnangagwa was inaugurated in November 2017. Interestingly, many of Ambassador Ward's predictions about how events might evolve in Zimbabwe have proven accurate.

Potential Contingencies, Warning Indicators, and Possible Effects on U.S. Interests

President Mugabe has retained a tight grip on the levers of power within both the government and the ruling party, Zimbabwe African National Union-Patriotic Front (ZANU-PF), but potential successors are jockeying for position. In 2014, Mugabe removed one potential successor, Joice Majuru, from her positions as vice president of both the government and party, and installed former Justice Minister Emmerson Mnangagwa in her place. Since then, Mnangagwa has not been able to solidify his status as successor to Mugabe. As political infighting has increased, so have the risks of instability and violence, which could play out along one or more of the following lines:

- *Mugabe dies or becomes incapacitated before installing a chosen successor.* Mugabe's health is clearly deteriorating. True to past form, Mugabe treats his current vice president as a figurehead rather than as a successor. Mnangagwa has not been able to cement the loyalties he would need to smoothly assume power.
- *Mugabe's control is challenged and undermined by growing factionalism.* The political opposition party in Zimbabwe is demoralized, discredited by electoral losses, and divided into factions. Nonetheless, the ruling party is also

divided, with factions crystalizing around Vice President Mnangagwa and First Lady Grace Mugabe.

- *An economic crisis triggers demands for political change.* Misguided economic policies, including land confiscation and forced indigenization of businesses, continue to depress productivity. Economic growth fell to 0.5 percent annually in 2016. No apparent plan is in place for resolving Zimbabwe's debt problem and returning the country to the economic mainstream. Renewed economic decline could lead to both civil unrest and new flows of refugees.

The following are developments that could warn of impending instability in Zimbabwe:

- *Further decline in Mugabe's health.* Observers should be alert for repeated absences by Mugabe from important state or party functions.
- *Increasing dissent, infighting, and factionalism within the ZANU-PF.* Party factionalism may be the most likely source of political violence. Observers should be alert for evidence of choosing of sides within the party by the leaders of state security organizations.
- *Public unrest.* Divisions within the ZANU-PF and increased activity by opposition parties have already stimulated an uptick in civil unrest. Observers should track trends and watch for changes in reactions to unrest by the major military and police commands.

Should a serious political crisis develop in Zimbabwe as a result of such developments, U.S. interests could face one or more of the following negative effects:

- A humanitarian crisis generated by refugee flows or food shortages would likely require an expensive U.S. aid commitment.
- Hopes for a productive bilateral trade and economic relationship would fade, and U.S. trade with Zimbabwe would remain minimal.
- U.S. military forces might be needed to evacuate the small U.S. citizen population in the country, estimated in 2010 at less than one thousand.
- Frictions could arise between the United States and member states of the Southern African Development Community (SADC) on how to respond to human rights violations by the Zimbabwean government.

On the other hand, a stable and prosperous Zimbabwe would benefit U.S. interests. Zimbabwe is richly endowed with human and natural resources that could give it a leading role in shaping the future of the African continent. Bilateral trade and investment would likely build over time. Revival of Zimbabwe's agricultural sector would obviate continued humanitarian food aid. Eventually, Zimbabwe's security forces might play constructive roles in peace operations by the SADC and the African Union.

Ways of Preventing Violence

The United States possesses few policy instruments for directly influencing developments in Zimbabwe. Diplomatic relations are, at best, formal. President Mugabe continues to characterize the United States as a hostile force. U.S. assistance to Zimbabwe provides little political leverage because it is channeled through civil society groups. The government in Zimbabwe's capital city of Harare likely assumes correctly that U.S. humanitarian assistance would continue despite political instability. Targeted economic sanctions are widely seen as having little impact, and Mugabe uses the sanctions to stir up rancor against the United States.

Two types of preventive strategies are available to the United States. First, the United States could use both positive and negative incentives to attempt to shape a political outcome. The limited policy instruments available to the United States, mentioned above, would constrain the effectiveness of this strategy. Further neither Zimbabwe's neighbors nor the European allies of the United States would be likely to join such a strategy.

Second, the United States could choose a less ambitious but more realistic strategy by seeking to minimize the risk of political violence and economic turmoil, while positioning itself to support post-succession opportunities for political and economic reform. Under this strategy, the United States could seek a relatively swift and uncontested succession in order to enable a new government in Harare to begin to attend to Zimbabwe's economic and social challenges.

No single outside actor has the capacity to directly influence President Mugabe's choices regarding succession, but a well-orchestrated multilateral strategy could help him and others understand the potential negative consequences of decisions that would increase repression, deepen the country's economic problems, and lead to social instability. In pursuing such a strategy, the United States would maintain support for civil society in Zimbabwe and continue a frank and direct dialogue with the Mugabe government. It would also seek to persuade others to act as follows to prevent violence or, failing that, reduce the consequences of violence:

- South Africa and other SADC member states could remind President Mugabe of his responsibilities under the organization's statute to maintain peace and stability in his own country.
- China is Zimbabwe's most important economic partner with bilateral trade of over $1 billion annually. In the interest of protecting its sizable investments in Zimbabwe, China might be motivated to privately indicate to Mugabe its concerns over the possibility of instability.
- The European Union (EU) countries could increase their involvement with civil society organizations in Zimbabwe and indicate clearly to the Mugabe government that they would consider reimposing sanctions in response to greater repression.

Mitigating the Consequences of Potential Violence

In addition to working to prevent violence, the United States could seek to reduce the consequences of any potential violence:

- With South Africa and other SADC partners, there could be quiet, advance consultations on the response to violence in Zimbabwe. Advisory warnings to Mugabe by SADC countries could lay the basis for more meaningful measures in the future.

- In addition, South Africa and other states bordering Zimbabwe could ensure that they are prepared to deal with additional refugees.
- Even though China would be unlikely to agree to consult in advance on actions that it might take in the event of violence in Zimbabwe, the subject should be given a prominent place on the agenda for U.S.-China consultations on Africa.
- With the EU countries, the United States could coordinate contingency planning for humanitarian assistance, including food aid.

If, despite these efforts, significant political violence occurs in Zimbabwe, U.S. policy options could include support for the following:

- *Mediation by SADC.* The choice of mediator would be crucial. South African Vice President Cyril Ramaphosa was effective as a mediator in Lesotho.
- *Action by the United Nations Security Council.* China might block decisive action by the Council, but might agree to the creation of a United Nations special envoy for Zimbabwe.
- *Coordinated increases in economic sanctions.* Reimposition of EU sanctions could affect the calculations of the Zimbabwe government.
- *Intensified official U.S. and Western dialogue with moderates in the ZANU-PF.* Senior figures in the ZANU-PF with extensive business interests might be interested in limiting violence.
- *Increased U.S. humanitarian assistance.* Additional humanitarian assistance, especially food aid, might be essential in order to assist large numbers of refugees and internally displaced persons.

Recommendations

In crafting its approach to a post-Mugabe Zimbabwe, the United States should keep its broad regional interests in mind. Even as bilateral relationships with East and West Africa have grown stronger, U.S. ties with the SADC countries have tended to stagnate. This has been particularly true in the case of South Africa. Zimbabwe presents an opportunity to begin strengthening the U.S.-South African security partnership. Early contacts with South Africa, other SADC countries, European allies, and China should be pursued as follows:

- *Intensified interagency efforts to define U.S. interests and options in Zimbabwe.* In the context of a formal interagency contingency planning effort on Zimbabwe, the U.S. government should forge a consensus on the goal of limiting violence and economic turmoil in Zimbabwe, and define the incentives and disincentives available. The objective should be an integrated approach that focuses on practical, measurable steps that the government of Zimbabwe could take to permit greater political expression and liberalize the economy.
- *Open a consultative channel on Zimbabwe with the U.S. Congress.* The purpose would be to help members of Congress understand that positive change in Zimbabwe is likely to take place incrementally if at all, and to build a basis of trust for executive branch actions, whether carrots or sticks, further down the road.
- *Pursue understandings on Zimbabwe with South Africa and other SADC countries.* Conversations with members of South Africa's government under President Jacob Zuma should focus on achieving South African agreement to urge President Mugabe and other ZANU-PF leaders to avoid violence.
- *Consult regularly on Zimbabwe with senior African affairs officials in EU countries.* The objective of these contacts would be to work toward consensus on positive and negative incentives for Zimbabwe, including sanctions.
- *Seek to influence China on Zimbabwe.* The United States should propose regular, in-depth conversations on Zimbabwe, focused on persuading the Chinese government to support a peaceful political transition in Zimbabwe.
- *Seek senior-level dialogue with the Zimbabwean government in multiple venues.* To supplement contacts in Harare, the United States should seek to strengthen parallel communications channels in Washington, D.C., and at the United Nations.
- *Expand youth and student exchanges.* Consideration should be given to further expanding access by young Zimbabweans to the Young African Leader Initiative and similar programs.
- *Ensure the security of the U.S. mission in Zimbabwe.* Plans for ensuring the security of the U.S. embassy and its personnel and for conducting an evacuation of those personnel if necessary should be updated regularly.

Looking toward the longer term, the United States should test the waters for expanding the bilateral dialogue. Given the economic plight of Zimbabwe and its humanitarian needs, there is potential for cooperation with a new government on trade and commercial issues. It would be possible to begin to unfreeze the bilateral relationship with steps such as trade and investment missions. The United States should then pursue political dialogue in close coordination with its Western allies and South Africa.

Conclusion

Zimbabwe's problems, which have been created by decades of authoritarian misrule and poor economic management, will not be quickly solved. Any successor to Mugabe will have to deal with a bitter political legacy and difficult economic conditions. The alternatives open to the United States are limited by strained political relationships and minimal economic ties. The scarcity of options should not be a rationale for doing little or nothing. Rather, it should be seen as a call for the United States to focus on what is essential—reducing the possibility of political instability and civil violence during the post-Mugabe succession—while laying the groundwork for a better relationship with an eventual successor government.

George Ward, a Research Staff Member in the Intelligence Analyses Division of IDA's Systems and Analyses Center and a former U.S. Ambassador to Namibia, holds a master of public administration degree in systems analysis from Harvard University.

This article is based on "Political Instability in Zimbabwe," Contingency Planning Memorandum No. 53 (New York: Council on Foreign Relations, Center for Preventive Action, March 2015), https://www.cfr.org/report/political-instability-zimbabwe. A follow-up piece, "The Day After in Zimbabwe: Contingency Planning Memorandum Update," was published online on June 15, 2017, and is available at https://www.cfr.org/report/day-after-zimbabwe.

Prioritization Framework: A Step Toward Cost-Effective Verification, Validation, and Accreditation

S. K. Numrich, Robert A. Zirkle, James R. Ayers, Forrest R. Smith, and Anna Vasilyeva

Verification, validation, and accreditation (VV&A) is too often done as an afterthought to the development process of a simulation at a time when most of a project's resources have been exhausted. VV&A may also be done when some higher authority threatens the existence of a program because the simulation tools used have not been subjected to VV&A. Under such circumstances, the process becomes a tax on already burdened programs and a headache for the program manager. The method discussed here was developed to assist a program manager in performing verification and validation (V&V) on an existing simulation within budget and without sacrificing application performance. Prioritization based on the intended use of the simulation seemed the most beneficial route to performing a cost-effective V&V process (Department of Defense 2007).

The framework gives the user the ability to provide rational, repeatable, and documented evidence for decisions concerning where to focus V&V resources.

Terms

The purpose of verification is to determine whether the equations or computational models used to represent the entities in a simulation are encoded properly so that the software accomplishes what the model developer intended it to accomplish. In validation, the essential question is whether or not the encoded *representation* corresponds to the measure of the physical (or real) world it is supposed to represent. The measure of the real world, called a *referent*, encapsulates an understanding of the segment of the real world to be captured in a simulation. Evaluation of how well the representation corresponds to the referent could be accomplished using comparisons to measured data or to a commonly accepted mathematical relationship. Lacking either of those means, developers may seek the considered opinion of subject matter experts.

To accomplish V&V of a simulation, the *V&V agent*—the organization, group, or person performing V&V activities—must have a viable set of *requirements* describing what the software is supposed to do, along with referents or acceptable standards of representation for those requirements. These requirements and referents should have guided the development of the model or simulation. Without the guidance provided by requirements and in the absence of good referents, any degree of performance could constitute acceptable correspondence to the real world, potentially leaving the user with software that is inappropriate, inadequate, or unusable.

Nonetheless, upon initiating the V&V effort, the V&V agent often discovers that the list of requirements is incomplete, thus failing to cover the user's requirement space and lacking in requisite specificity. Furthermore, referents frequently are not specified for any of the required representations. In a quest for supporting documentation, the V&V agent is likely to find that there is no conceptual model (and thus, no record of agreements and compromises between the user and developer), and the documentation on any verification and validation done during development is incomplete at best.

The V&V agent is often left with the task of refining a weak set of requirements and engaging the user in determining a potential set of referents. We developed a prioritization framework to help users determine the relative importance of representations and the *behaviors* contained in the refined requirements in light of the type of conditions under which a simulation was intended to be used. The goal was to allow users to make judicious choices as to which representations and behaviors had to undergo V&V now and which could be left to a future date, based on cost and priority.

Prioritization Framework

Asking a user about requirements and referents often ends in consternation on the part of both the user who never thinks in those terms and the V&V agent who works with them daily. The prioritization framework employs a scenario-based approach that allows the user to express requirements by examining the intended uses of the simulation and determining their relative importance. The importance could be based on frequency of use or on components of the scenario that cannot be tested in real-world exercises. The evaluations by the user are expressed in a probability tree and result in a weighting factor for each intended use. The V&V agent uses the same scenario descriptions, but focuses on each representation (an environmental factor, weapon, sensor, etc.) to determine how critical each specific representation is to the execution of the scenario and intended use. The final evaluation of priority uses both the user's weighting factors and the determination of criticality to compute the final priority. The computation is done using a spreadsheet. The user can easily change any of the probabilities and recompute the prioritization in minutes. The result is a simple system readily explained in terms the user understands and can defend. The following sections illustrate the use of the framework.

Scenarios

Our user had three significant missions for which the simulation was to be used: protection of a warehouse facility, defense of a convoy, and protection of a distribution center. By walking through each of these scenarios, the user was able to identify essential representations in the scenario and conditions under which the scenario would take place (time of day and weather conditions). While all three missions were assessed in the application of the prioritization framework, the following explanations show only two: protection of the warehouse and defense of the convoy. Additional missions can be added as branches on the probability tree.

Weighting Factors

The prioritization framework uses a tree structure virtually identical to a probability tree where the starting point is the set of problem scenarios, which are represented as mission areas in our example (Figure 1). Each mission area is assigned a percentage based on its importance or frequency of use (according to the user's preference). The percentages are used to assign a weighting factor between 0 and 1.

The second tier uses time of day. For the example shown, only day and night were used; however, it would be possible to use day, night, dawn, and dusk as each of these times presents unique lighting conditions that affect sensors. The third tier refers to weather condition: clear, rain, or snow. At each tier, the sum of the weighting factors assigned within that tier must sum to 1.

Criticality Rating

Assisting the user in setting the weighting factors is the first step. The second step consists of examining the representations and behaviors and determining whether they are important or useful for any scenario at the specified time of day and weather condition. For simplicity, the following numerical assignments were made: 2 for critical to use at that time and under that condition, 1 for occasionally used at that time and under that condition, and 0 for not needed at that time and under that condition.

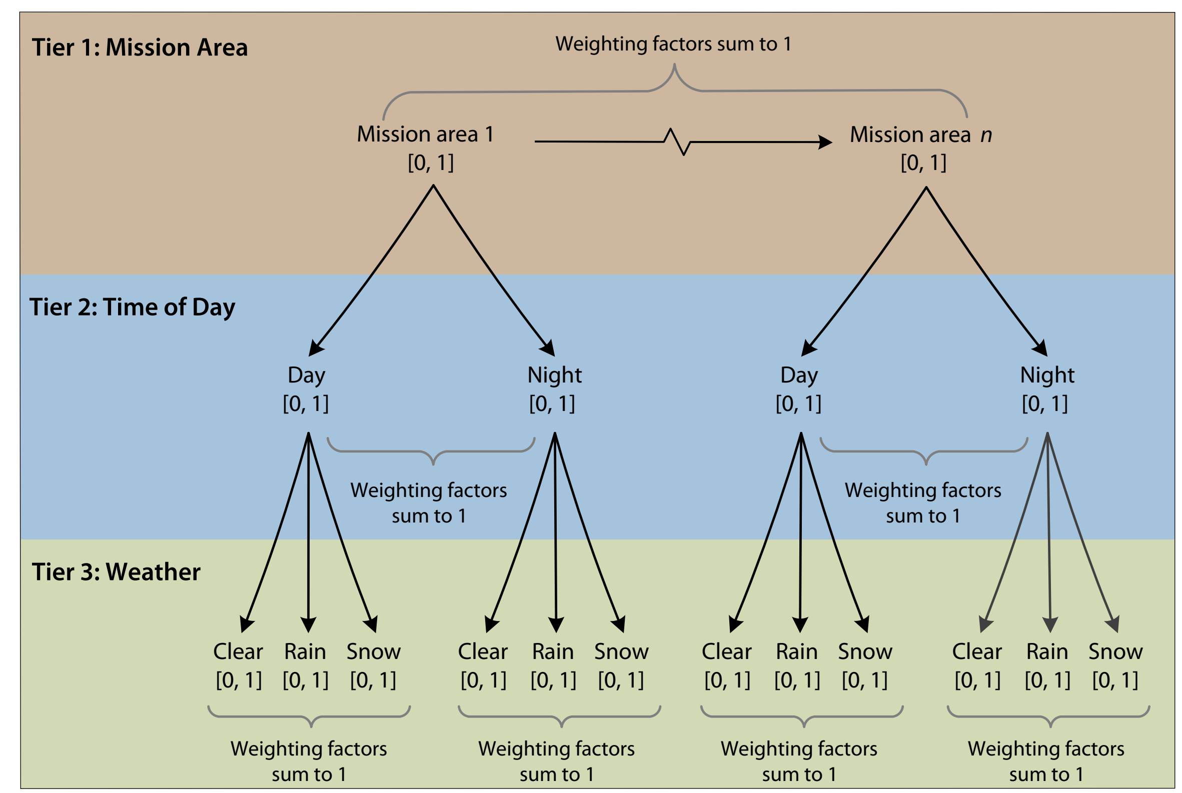


Figure 1. Probability Tree Structure Used to Determine Weighting Factors

Evaluation Matrix

As an example, consider a problem in which the user is examining the defense of goods being held at a warehouse and the protection of those goods when being transferred by convoy to a distribution center.

The user determined that for the initial set of evaluations, the warehouse mission area would be exercised about 60 percent of the time, and the convoy mission area, about 40 percent of the time. Attempts to attack at night represented the most significant threat to the facilities, leading to a weighting factor of 0.7 for the warehouse mission area at night. Convoys, on the other hand, are rarely planned for night, but extenuating circumstances might make it necessary to extend the duration of a convoy into hours of darkness. Thus, some weighting was placed on night operations for convoys (0.2).

Weather conditions were then assigned for day and night. The user expected to exercise the simulation for the warehouse half the time for clear weather conditions during daylight hours, resulting in an environmental weighting factor of 0.5 for day, and less frequent use of the simulation during inclement weather (weightings of 0.3 and 0.2, for rain and snow, respectively). At night, however, the more dangerous conditions at the warehouse facilities were operations during rain when the effectiveness of some of the sensors protecting the site would be reduced. In this case, rain was therefore assigned a weighting of 0.5, while clear weather and snow were assigned lesser values (0.3 and 0.2, respectively). In this manner, the user was able to provide weighting values for the probability tree in Figure 1 by considering the circumstances under which the simulation would be used to provide assessments.

Convoys are typically planned for daylight hours with a limited number having to extend into nighttime operation; therefore, the convoy's weighting factor for day was set at 0.8 with a corresponding value of 0.2 for night. Similarly, convoys are planned for clear weather, although daytime operations can more readily tolerate rain than can nighttime operations. Thus, the weighting factors assigned for daytime operations were 0.6 for clear weather, 0.3 for rain, and 0.1 for snow. Nighttime operations are less tolerant of inclement weather; therefore, the weighting factors set were 0.8 for clear weather and 0.1 each for rain and snow.

The evaluation matrix in Figure 2 shows these weighting values in the top four rows, which are color coded to correspond to the tiers of the probability tree in Figure 1. The cumulative weighting factor is determined by multiplying the three weighting factors (mission area, time of day, and weather) for each branch of the tree. In Figure 2, these cumulative weighting factors are found in the fifth row, the first white box under the colored rows, which represent the individual weighting factors.

The next step was conducted by the V&V agent's technical experts and involved the assignment of a criticality rating (2, 1, or 0) to each combination of mission area and condition. To illustrate this process,

- Cumulative weighting factor
- Criticality rating
- Individual priority score (product of cumulative weighting factor and criticality rating)
- Overall priority for a given representation considering all scenarios (sum of individual priority scores across all scenarios)

		Warehouse = 0.6						Convoy = 0.4						
		Day = 0.3			Night = 0.7			Day = 0.8			Night = 0.2			
		Clear	Rain	Snow	Clear	Rain	Snow	Clear	Rain	Snow	Clear	Rain	Snow	
		0.5	0.3	0.2	0.3	0.5	0.2	0.6	0.3	0.1	0.8	0.1	0.1	
Category	Requirement	0.09	0.05	0.04	0.13	0.21	0.08	0.19	0.10	0.03	0.06	0.01	0.01	
Environment	Berm	2	2	2	2	2	2	1	1	1	1	1	1	
	Vegetation: grasses	2	2	1	2	2	1	2	2	1	2	2	1	
	Precipitation	0	2	2	0	2	2	0	2	2	0	2	2	
Sensors	Light intensification devices	0	0	0	2	2	2	0	0	0	2	2	2	
	Passive IR devices	1	1	1	2	2	2	1	1	1	2	2	2	
														SUM
Environment	Berm	0.18	0.108	0.072	0.252	0.42	0.168	0.192	0.096	0.032	0.064	0.008	0.008	1.6
	Vegetation: grasses	0.18	0.108	0.036	0.252	0.42	0.084	0.384	0.192	0.032	0.128	0.016	0.008	1.84
	Precipitation	0	0.108	0.072	0	0.42	0.168	0	0.192	0.064	0	0.016	0.016	1.056
Sensors	Light intensification devices	0	0	0	0.252	0.42	0.168	0	0	0	0.128	0.016	0.016	1
	Passive IR devices	0.09	0.054	0.036	0.252	0.42	0.168	0.192	0.096	0.032	0.128	0.016	0.016	1.5

Figure 2. Evaluation Matrix

we first separated prioritized values into categories of representation types: weapons, platforms, sensors, human behaviors, and environmental factors. Every representation was evaluated for each scenario; however, for purposes of illustration, we selected requirements from two different categories of representation. We chose three environmental representations: berms, vegetation in the form of grass, and precipitation. Note that precipitation here refers to the model requirement to represent precipitation and its effects, while rain and snow are the weather conditions under which the model is expected to be used. We also chose two representations from the sensor category, light intensification devices and passive infrared (IR) devices.

Berms are built as barriers against incursion by unwanted visitors. They are not likely to be found along a convoy route, but they might be present as artifacts of prior events. For example, fortifications built along coastal roads on both the East and West Coasts of the United States for fixed gun batteries have defensive features in common with berms and could be represented as such. Thus, berms have a criticality rating of 2 for the warehouse mission area and 1 for the convoy mission area, across all conditions. Grasses are found around both warehouse sites and along the roadside; however, during snow, they are likely to be weighted down and, hence, less important as potential cover for threats. Criticality for grasses is rated as 2 for clear and rainy weather and 1 for snow for both the warehouse and convoy mission areas. Precipitation is irrelevant for clear days or night and thus rates a 0 under clear weather conditions. Light intensification devices are used during low light conditions and are thus rated as irrelevant (0) during the day and critically important at night (2). IR devices may have some use during the day, but are critically important at night, as reflected by their scores of 1 and 2 for those circumstances.

Once the cumulative weighting factors and the critical factors are determined, the prioritization can be computed, first for each individual scenario and then summed for all scenarios to find the overall priority. For purposes of illustration, the representations used in the above computations are replicated below the grey row in Figure 2. The values represented in each of the cells in the lower part of the matrix result from multiplying the cumulative weighting factor by the criticality rating. The overall priority, shown in the column on the right labeled *SUM*, is the sum of all the values for that row—effectively the sum of the priority for that representation for each scenario summed for all scenarios.

The simulation has hundreds of individually listed representations, making it desirable to group them into categories of similar entities. (The term *entity* is used in simulation to mean the thing, behavior, or condition being represented.) The representations are listed in Figure 2 under the term *requirement* because these are the required entities. While all the representations could be listed in a single set in priority order using the values computed in the SUM column, the results made more sense when the ordering was done within the category.

The computational framework presented above lacks the ability to account automatically for interdependence among representations. For example, to use a weapon successfully, the actor in the simulation might have to be able to assume different positions and seek cover. If the use of the weapon had a high priority, the accompanying behaviors on the part of the actor would have to also have that high priority, even if seeking cover was not a high priority when considering the actor's behaviors in isolation. However, having all the representations ordered within their respective categories facilitates cross-category comparisons for detection of such interdependencies. The ability to determine interdependencies is important when resources for validation are limited. The cut-off points for investment have to include all the related representations needed for coherent operation.

Conclusion

The use of this prioritization framework is readily understandable from the perspective of the user and technologist, and it allows the user to establish needs in clear terms. The framework gives the user the ability to provide rational, repeatable, and documented evidence for decisions concerning where to focus V&V resources, thereby providing increased confidence that the simulations selected adequately portray the conditions appropriate to the intended use.

The prioritization framework presented here is easy to implement and is based on user needs and intended use of the simulation. While the method was developed to support a V&V effort directed toward potential acceptance of a simulation developed for other users, it can be adjusted for use in managing investment in any new simulation tool. Prioritization is not a definitive assessment, but a triage that can help the user determine the final selection of requirements to be validated for a given model. It also provides a rational, defensible, and repeatable process for choosing what to validate and what to leave out of the V&V process.

Reference

Department of Defense. 2007. *Modeling and Simulation Verification, Validation and Accreditation Recommended Practices Guide.* On-line Reference Guide, Washington, DC: Office of the Secretary of Defense.

S. K. (Sue) Numrich (second from right), a Research Staff Member in the Joint Advanced Warfighting Division (JAWD) of IDA's Systems and Analyses Center, holds a doctorate in physics from the American University.

Bob Zirkle (second from left), a Research Staff Member in the Strategy, Forces and Resources Division (SFRD) of IDA's Systems and Analyses Center, holds a doctorate in political science from the Massachusetts Institute of Technology (MIT).

Jim Ayers (far left), a Research Staff Member in JAWD, holds three master's degrees, the first in strategic studies from the Air War College, the second in air mobility from the Air Force Institute of Technology, and the third in business from Webster University.

Forrest Smith (far right), a Research Associate in SFRD, holds a master's degree in operations research from North Carolina State University.

Anna Vasilyeva (not pictured), a former Research Associate in the Information Technology and Systems Division of IDA's Systems and Analyses Center, holds two master's degrees, one in technology policy from the University of Cambridge Judge Business School and another in aeronautics and astronautics from MIT.

The full paper "Prioritization Framework: A Step Toward Cost-Effective VV&A," was published in the online proceedings of the 2016 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), held in Orlando, Florida, November 28–December 2, 2016.

Reshaping Space Policies to Meet Global Trends

Bhavya Lal

Fifty years ago, the United States and the Soviet Union conducted the only significant national space programs, and only a small number of commercial entities were involved in space activities. Now, while the United States remains the main player, the space sector includes many more countries and integrates technologies and innovations from other sectors. Private funding for space-based ventures has increased dramatically, contributing to rapid growth of the private space sector. As a result, the space sector is changing from being largely driven by government and several large commercial enterprises to being more segmented and globally integrated and driven by commercial activity (Figure 1). What do these trends mean for the U.S. government agencies and departments that spend in excess of $43 billion annually on space-based activities?

Participants in the private sector are focusing on cost innovation, following a philosophy of developing products that are good enough rather than perfect and prioritizing low cost over performance or reliability.

Space Enterprise Is Not an Island

In the early years of the space age, technologies were developed in and for the space sector and "spun out" into other sectors. Increasingly, though, the reverse is occurring, and technologies are spinning "into" the space sector from others, principally from advances in materials science, robotics, and information technology (IT) sectors, and often in the form of commercial off-the-shelf products. Falling costs and dramatic improvements in areas such as processing power, data storage, camera technology, solar array efficiency, and micro-propulsion have fed into a variety of space-related areas, including remote sensing and Earth observations, telecommunications, space science and technology, and exploration.

As a result, newer and lower-cost applications of space are emerging, making investing in space more beneficial and lucrative. Smaller, lighter, and more capable satellites make Earth observation and remote sensing within the reach of countries, corporations, and individuals alike. Use of high-throughput satellites can provide high-speed data communication that is many times faster than with traditional satellites. Using newer technologies and new business models, companies like SpaceX have developed reusable boosters, and are disrupting the launch market that had been controlled by heavyweights such as United Launch Alliance and Arianespace. Firms such as AGI, ExoAnalytic Solutions, and LED Labs are besting legacy government systems to provide space situational awareness (SSA) services that improve the ability to view, understand, and predict the physical location of objects in space, with the goal of avoiding collisions. The trend toward smaller satellites has yielded the use of CubeSats developed by private firms not only for commercial purposes like weather prediction, but also for national security–related activities such as rendezvous and proximity operations, and for scientific research into heliophysics, planetary science, and astrophysics.

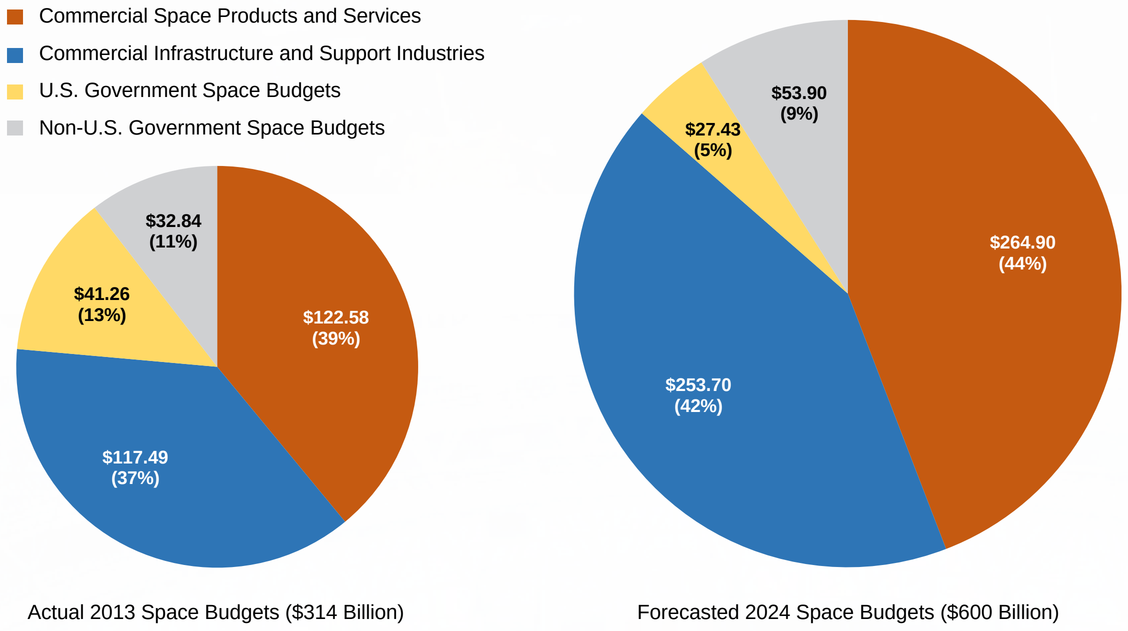


Figure 1. International Space Budgets

Government Funding and Policies

Government agencies in the United States and around the world are under pressure to reexamine policies restricting the commercial development and sale of space goods and services, as illustrated by the debate in the U.S. Congress on using commercial rockets, extracting space-based resources, or selling space-based imagery. There is also pressure on agencies to begin to view and regulate space as a mainstream economic endeavor, not solely as a sector relevant to national security and science. This shift in emphasis is especially evident in the United States and Europe, where commercial solutions are increasingly being used to meet government needs, technology export controls are being liberalized, and regulations are being relaxed to allow the private sector to provide services such as high-resolution imagery and SSA that were previously restricted to the government.

Signals of Change

Global Investment

That space is changing is evident in many measurable ways. Although there have been government cutbacks in the United States, globally there has been an increase in funding of space activity. Global investment in space activities increased at an annual rate of 6 percent between 2009 and 2013. In the broadest picture, almost 170 countries have some level of financial interest in satellites, up from 20 in the 1970s. Further, 60 of the more than 80 countries engaged in space-based activities in 2015 have invested $10 million or more in space-related applications and technologies, twice as many as in 2004. The increase has been especially noteworthy in countries such as Saudi Arabia (60-percent increase since 2009) and Brazil (40-percent increase). Overall, even by the most conservative estimates, global activity in space is expected to almost double in the next 10 years.

Other countries, such as India, are also demonstrating growing expertise in space exploration and technology development, while countries such as Israel, Singapore, South Korea, and the United Kingdom have begun to specialize in niche areas such as avionics, alternative approaches to launch, and data analytics, among others. The United Arab Emirates has plans to build a Mars probe and the first space research center in the Middle East, all the more impressive since the nation began its space activities only in the 1990s. Leveraging commercial products and services, including those from the United States, these and other countries are poised to become major space players and may well rival more established countries in a few years, particularly given the enduring perception that a presence in space brings prestige, geopolitical advantages, and economic opportunities.

Different Stakeholders

The presence of private companies in space is not a new phenomenon, but in the 1950s and 1960s, private companies operated under a model where investments went into a monolithic, capital-intensive industry driven by government. Now, the investments are less capital intensive, with different investors, especially those from the IT sector, being spurred by emerging markets. Some of these private investors are not motivated solely by profit, and they can provide long-term capital, previously the domain of governments only. Companies such as SpaceX, Blue Origin, and Bigelow undoubtedly intend to make money, but these companies' founders seem to be driven by a zeal—and a time horizon—that transcends that of a typical venture capital investor.

Another stakeholder in the space sector—one that did not exist when space was solely a government-driven sector—is the private consumer, who is both demanding and willing to pay for space-based services such as ubiquitous broadband access and near real-time situational awareness. These consumers are now contributors to the growing private sector. Add to this the emergence of concepts such as crowdfunding and citizen-led space activities, and the number of stakeholders in the space sector is dramatically higher than it was even a decade ago.

New Approaches

These stakeholders are also following different approaches to developing their space enterprises. Governments in less-industrialized countries are increasingly using technology transfer and partnerships to build capabilities in specific areas of interest rather than investing in developing indigenous systems. It is no longer necessary to build a satellite or even operate it to get data from one. At the same time, there is also a shift from buying technology and products to buying services.

Participants in the private sector are focusing on cost innovation, following a philosophy of (1) developing products that are good enough rather than perfect and (2) prioritizing low cost over performance or reliability. This approach is reflected in the increased use of streamlined processes, cheaper components, open-source hardware and software, agile manufacturing, and production models (as distinct from the production of one-off products).

These trends are most evident in the small satellite sector, where risk and reliability are seen differently than in the traditional aerospace sector. For example, the Earth observation company Planet Labs can have a fifth of its CubeSats fail in orbit without losing a meaningful amount of its imaging capacity. Such an architecture becomes feasible only when satellites cost 2 to 3 *orders of magnitude* less than traditional satellites. Many small satellite firms (Spire, for example) see themselves not as aerospace firms but as information technology or media companies, so they are takeover targets for technology giants such as Google or Facebook, not for traditional aerospace firms such as Lockheed Martin or Boeing.

Implications for the United States

Developments such as those described here reveal that many of the subsectors of space, including Earth observation, space science and technology, exploration, and even SSA, are beginning to diverge into two segments. The first segment is a government-driven one that develops massive systems such as the James Webb Space Telescope or Space Launch System rockets. These systems have exquisite capabilities that require hundreds of millions to billions of investment dollars to develop and operate. The second is a less-capable but also less-expensive consumer-oriented segment. Largely centered in Earth observation data and services today (and telecommunication and other services in the future), the segment is globalizing rapidly and will inevitably spread to other subsectors of space. It is therefore not difficult to believe that the future holds both domestic and international implications for the United States.

For example, the emergence of new applications (e.g., commercial radio frequency sensing or signals intelligence) presents unprecedented challenges not only to U.S. government organizations such as the National Reconnaissance Office that control such national security–centered activities, but also to regulatory agencies such as the Department of Commerce that have no systems in place for such new applications Similarly, with SpaceX, OneWeb, and other private companies planning to launch satellite constellations comprising thousands of satellites, the same system of spectrum licensing by the Federal Communications Commission will not work.

Globally, the challenges will be even more complex given how space is increasingly described as congested, contested, and competitive. The guidelines surrounding space debris are currently nonbinding and difficult to enforce. Compounding the challenge are the high cost of debris mitigation and the prospect of the United States having to share debris mitigation technologies with less-wealthy nations and nonspace actors that are launching spacecraft.

Today's space community must also address previously unknown challenges such as the loss of electromagnetic spectrum; the lack of global standards and regulations for activities related to serving satellites or other objects on-orbit; the development of deep space mining or in situ resource utilization; the rise of cyber terrorism; and the legacy of pollution from launches. These and many other challenges now confronting global space powers require an appropriate response.

With more countries and private sector firms operating in space and seeking to take on additional roles by participating in international space organizations, the domestic and global governance landscapes will continue to become more complex. The United States and other traditional space-faring countries will have diminished control of global decisions related to space activities, and they will be under greater pressure to accommodate the needs of the private sector and countries with emerging space capabilities.

Ready for Wildcards

As efforts to develop and implement policy changes to address these and other challenges proceed, it will be useful to ensure that any changes are alert to unknown and unforeseen situations—*wildcards*—that might overturn these trends. Wildcards could be related to technology developments. A dramatic breakthrough, such as perfecting the ability to reliably and cheaply reuse multiple stages of rocket engines or developing specialized carbon nanofibers that make technologies such as space elevators feasible, could dramatically reduce the cost of access to space.

Wildcards can also emerge from geopolitical developments. Drastic changes or responses to the Outer Space Treaty or other international rules governing space, or aggressive weaponization of space, could affect how liberal the U.S. government will be with respect to international collaborations. Other wildcards that could upend the current trajectory include a debilitating space weather disaster or cyber-event that cripples space-based services for an extended period, a space-debris cascading event that degrades use of space, or the discovery of a large asteroid or comet headed toward Earth. Any policies need to be robust to these wildcards.

Even if no wildcards enter the picture in the near term, or if policies are implemented that are responsive to multiple alternative futures, hasty change is not advised. For example, the government does not necessarily need to use capabilities that become available from outside the government just because they are available.

Policy makers need to decide which capabilities are so important that they should not be outsourced, procured, or purchased from outside the government. Such decisions are likely to be complex, and will probably have political implications, so policy makers must also plan for the consequences of these decisions.

Change in the Wind

It seems clear that the space sector will continue to undergo transformation as it increasingly, if gradually, diverges from the military/government users. More governments worldwide can be expected to act on their space aspirations by participating in space activities in different ways, and a globalized private sector (even if mostly centered in the United States) will begin to provide more space-based products and services. As the number of actors increases, the space sector will likely see increased competition and overcrowding, both literally and metaphorically. This, in turn, will serve as a driver for more products, services, and governance structures that can support the needs of the ever-expanding sector.

It is also clear that the U.S. government will need to adapt to these changes by reshaping its space departments and agencies and by leveraging developments beyond their conventional boundaries. Toward this end, the government will need to harness its vision, openness, agility, and risk tolerance; incorporate a well-matched mix of centralized planning and decentralized execution; and expend the resources required to implement these changes.

Bhavya Lal, a Research Staff Member in IDA's Science and Technology Policy Institute, has a doctorate in science and technology policy from the George Washington University School of Public Policy and Public Administration.

The original article, "Reshaping Space Policies to Meet Global Trends," was published in *Issues in Science and Technology: A New Global Environment for Space Policy,* Summer 2016, http://issues.org/32-4/reshaping-space-policies-to-meet-global-trends/.

Statistical Approach to the Operational Testing of Space Fence

Daniel L. Pechkis, Nelson S. Pacheco, and Tye W. Botting

Our approach quantifies the probabilities of meeting requirements; determines how performance varies as a function of an object's altitude, inclination, or size; estimates a 25-day test duration; and determines that modeling and simulation methods may be needed to represent 125 additional satellites.

The next-generation space surveillance system, known as Space Fence, uses radar to track space debris and operational satellites in low and medium Earth orbit that may threaten U.S. space assets. The enhanced capability provided by Space Fence is expected to increase the number of routinely tracked orbiting objects from approximately 17,000 to more than 100,000 (U.S. Air Force 2014; NASA 2012). Given that existing sensors cannot quickly validate the new radar's complete set of observations, the question becomes how to test successfully this new space surveillance system.

Difficulty of Testing Space Fence

Space Fence is a new ground-based, space-directed radar system the United States is acquiring to detect, track, and catalog orbiting space objects, including the growing population of space debris. The system will consist of two S-band (2–4 GHz) phased-array radar sites from which it will perform autonomous cued and uncued surveillance and cued searches for objects in low and medium Earth orbits. Space Fence will provide tracking and radar characterization data on orbiting objects to the U.S. Air Force Joint Space Operations Center to support maintenance of the Satellite Catalog (SATCAT) and support other space situational awareness needs (U.S. Air Force 2014).

Space surveillance radar performance has traditionally been tested by comparing radar observations against truth data on position, velocity, and time for a small number of well-understood objects with known positions measured to within 1 meter using laser ranging or onboard beacons (Mochan and Stophel 1968; Noll and Pearlman 2011; Joint Range Instrumentation Accuracy Improvement Group 1995; Martin et al. 2011). However, truth data for the majority of the approximately 17,000 objects tracked by existing radar systems and optical telescopes may not be sufficiently accurate to validate the data on the much larger number of satellites expected to be tracked by the highly accurate Space Fence. Further, truth data on the relatively small number of objects may not extrapolate to an operationally representative population of space objects of different types, inclinations, altitudes, sizes, shapes, and rotational motions. Because Space Fence will have a larger field of view and higher sensitivity than existing radar systems, it is expected to be able to routinely track more than 100,000 orbiting objects (U.S. Air Force 2014; NASA 2012).

Testing a space surveillance system whose complete set of observations cannot be validated in a timely manner by existing radar systems and optical telescopes presents some challenges.

- How can testers know Space Fence is capturing all the objects it is intended to observe?
- Are the radar measurements on all the objects observed by Space Fence of sufficient accuracy and precision to both meet its requirements and support orbital prediction and SATCAT maintenance?
- Can adequate testing of an operationally representative sample population, covering all intended object sizes, altitudes, and inclinations, be performed in a timely manner?

Proposed Testing Method

To address the issue of Space Fence performance across the full operational space, we propose extending initial calibration tests into broader rigorous statistical test designs, using on-orbit test targets that span the orbital limits of Space Fence's operational requirements. Through this approach, we characterize Space Fence performance by using a relatively small subset of the publicly available SATCAT (~1,500 out of ~17,000 objects),[1] grouped by altitude, inclination, and size (Pechkis, Pacheco, and Botting 2014, 2016).

Building on recent experimental design work for assessment of naval surface radar performance (Cortes and Bergstrom 2012), we used the target altitude, size, and inclination as predictor variables (or factors) in statistical tests for measures of radar performance requirements (e.g., range accuracy) as dependent (or response) variables. This approach quantified the probabilities of Space Fence meeting its performance requirements, determined whether and how satisfaction of individual requirements depends on an object's orbit and size, and estimated the sample sizes needed for statistical confidence in this evaluation. Comparing the resulting sample sizes with the number of currently known targets, we determined the areas where augmentation with modeling and simulation (M&S) may be needed because of an insufficient number of targets. Finally, we estimated the necessary test duration by assuming a radar coverage solely for the first radar site (located in Kwajalein in the Marshall Islands) and a conservative number of radar tracks per object per day.

Evaluating Space Fence in Terms of Operational Requirements

We chose four Space Fence operational requirements—metric accuracy, probability of track, object association, and data latency—as the response variables to illustrate the different statistical test design methodologies needed to support Space Fence operational test and evaluation.

[1] Analyses are based on data from the entire publicly available SATCAT on space-track.org as of June 2013.

Metric Accuracy

The metric accuracy of Space Fence is stated in terms of measurement errors (error variance) for each of five radar observation components: time, elevation, azimuth, range, and range rate. Metric accuracy is key to establishing orbital precision and for supporting the coverage and flexibility of radar surveillance.

We determined the sample sizes necessary for computing time, elevation, azimuth, range, and range rate accuracy of uncued objects entering the observation field of view by evaluating measured errors in these metrics with a hypothesis test for their variance, assuming a normal distribution. The hypothesis test results were agnostic to specific requirement thresholds; instead, they depended on the effect size (the amount a parameter exceeds its threshold), the desired statistical power (the probability of correctly determining that the requirement is met), and the significance level, referred to as α error (the probability of incorrectly determining that the requirement is not met).

Accuracy requirements were initially tested against a subset of SATCAT objects with highly accurate information available and then against the entire SATCAT inventory. For the initial testing, we selected two subsets of satellites known to contain accurate position, velocity, and time data—the International Laser Ranging Service (ILRS) satellites (Noll and Pearlman 2011) and the High Accuracy Satellite Drag Model (HASDM) satellites (Storz et al. 2005).

Six hundred object tracks were necessary to achieve a statistical power level of 95 percent for an effect size of 10 percent and an α error of 5 percent. We chose a 10-percent effect size because it would be sufficient to detect meaningful improvement or shortfall between Space Fence and legacy systems. Assuming that half of the satellites in the two subsets (60 satellites) are available and each had a conservative number of two *acceptable tracks* per day over a Kwajalein-based radar,[2] we calculated that 600 tracks could be obtained in as few as 5 test days:

$$5\ \text{days} = 600\ \text{tracks} \times \left(\frac{1}{60} \times \frac{1}{satellites}\right) \times \left(\frac{1}{2} \times \frac{days}{(tracks/satellites)}\right)$$

We used the analysis of variance (ANOVA) method to determine the probability of detecting whether or not any factor, or a combination of factors, affects the metric accuracy measurements of Space Fence. The factors we considered are altitude, inclination, and size, and we chose levels for each factor consistent with Space Fence requirements (Table 1). To implement this ANOVA approach, we first searched through the SATCAT for satellites likely to be observable from a Kwajalein-based Space Fence and estimated the average number of tracks per day.

[2] An acceptable track is a radar track of an object passing through the radar's field of view at a sufficient elevation and for a sufficient distance to allow the radar to gather enough data to generate observations.

Table 1. Number of Available SATCAT Objects by Inclination, Altitude, and Size

Inclination (Degrees)	Altitude (Kilometers)	Number by Size (Centimeters)*: SATCAT Objects < 10	SATCAT Objects ≥ 10	Real Tracks/ Minimum Test Days† < 10	Real Tracks/ Minimum Test Days† ≥ 10	M&S Tracks Needed‡ < 10	M&S Tracks Needed‡ ≥ 10
9 ≤ I ≤ 45	250–600	1	32	25/25	25/1	0	0
	600–2,000	4	101	25/4	25/1	0	0
	2,000–6,000	0	6	—	25/3	25	0
	6,000–22,000	0	2	—	25/7	25	0
45 < I ≤ 80 (centered on the highly populated band in the mid-60s)	250–600	16	85	25/2	25/1	0	0
	600–2,000	534	2,498	25/1	25/1	0	0
	2,000–6,000	0	10	—	25/2	25	0
	6,000–22,000	1	246	25/13	25/1	0	0
80 < I ≤ 171 (representing near-polar and retrograde orbits)	250–600	28	276	25/1	25/1	0	0
	600–2,000	1,372	5,728	25/1	25/1	0	0
	2,000–6,000	0	89	—	25/1	25	0
	6,000–22,000	0	2	—	25/7	25	0
Total	—	**1,956**	**9,075**	**175/25**	**300/7**	**125**	**0**

* Objects <10 cm are included to capture sensitivity improvements from Space Fence; objects ≥10 cm sizes are tracked by current radars.

† The notation 25/*n* indicates that 25 tracks can be obtained in a minimum of *n* days.

‡ The number of M&S tracks that would be needed to augment the real track to meet the 25-track limit.

For this calculation, we assumed a conservatively low number of one acceptable track per day for altitudes less than 600 kilometers, and two acceptable tracks per day for all targets above 600 kilometers. The ANOVA design evenly divides the 600 tracks needed to test the radar calibration across all factor-level combinations to ensure that all combinations are tested.

As shown in Table 1, there are a total of 4 × 3 × 2 = 24 combinations of object altitude, inclination, and size levels, so each combination requires 600 ÷ 24 = 25 data points. Table 1 also contains the number of objects expected to be available from the SATCAT over an approximate one-month test period for each factor-level combination, compared with the 25 tracks needed. (As of June 2013, the publicly available SATCAT contained 16,845 objects, of which 15,842 were in Earth orbit and had complete data.) A one-month test period allows for schedule flexibility and is consistent with historical cost-effective operational test periods.

Tracks from objects in the SATCAT are available in all inclination, altitude, and size regimes, except for objects smaller than 10 centimeters at altitudes between 2,000 and 22,000 kilometers, for which M&S would be needed.

Probability of Track and Object Association

Space Fence has probability requirements for tracking objects that pass through its field of view (*probability of track*[3]) and for associating those tracks with objects in the catalog (*object association*[4]). Unlike metric accuracy requirements, which are expressed as continuous responses, probability of track and object association requirements are stated in terms of binary responses (tracked or not tracked, associated or not associated). As such, we propose statistical hypothesis tests on binomial distributions to assess the system against documented system requirements. We then apply a logistic regression/Monte Carlo method to determine if system performance varies with an object's altitude and inclination.

Unlike in the metric accuracy analysis, the sample sizes necessary to demonstrate Space Fence can meet its probability of track and object association requirements depend on the specific requirement threshold values. For illustrative purposes, we chose threshold requirements of 50-percent probability of track and 97-percent object association. Sample sizes of 268 and 81 tracks, respectively, can demonstrate the radar's probability to meet these threshold requirements for 10-percent effect size, 5-percent α error, and 95-percent power. Using logistic regression/Monte Carlo methods, we determined that the effects of altitude and inclination on probability of track and object association can be tested with 1,530 and 540 tracks, respectively, at 10-percent effect size, 5-percent α error, and at least 90-percent statistical power in 8 days. Table 2 shows the required number of data points for 3 × 3 = 9 combinations of altitude and inclination levels for both probability of track and object association. Each combination requires 1,530 ÷ 9 = 170 and 540 ÷ 9 = 60 data points, respectively.

Data Latency

Our final response variable is data latency—the time from when the sensor has finished collecting the data to when the U.S. Air Force Joint Space Operations Center has received the data. For Space Fence, we assumed data latency will be no more than 2 minutes 99 percent of the time and used tolerance intervals to determine the number of data points necessary to evaluate the latency requirement. Data latency is not typically sensitive to the characteristics of the orbiting objects, so we did not account for factor analyses.

A sample of 856 data transmissions can achieve a 90-percent power level for a 10-percent effect size and a 5-percent α error. A 10-percent effect size for latency corresponds to a 12-second delay in the 2-minute latency threshold. Although this may seem like a short delay, it could prove significant for certain conjunction alerts and consequent collision avoidance maneuvers.[5] For the International Space Station, for example, with a collision-avoidance-maneuver velocity of 0.5–1 millisecond (Hutchinson 2013), a 12-second delay would mean being 6 to 12 meters closer to a potential conjunction.

[3] Probability of track is the probability of keeping track of the position and velocity of a given object that penetrates the radar's field of view.

[4] Object association is the probability of associating detected objects with known SATCAT objects (to determine if a detected object is already known or newly discovered).

Table 2. Number of Available SATCAT Objects to Test Probability of Track and Object Association, Ordered by Inclination and Altitude

Inclination (Degrees)	Altitude (Kilometers)	Quantity	Real Tracks/ Minimum Test Days* Probability of Track	Real Tracks/ Minimum Test Days* Object Association
9 ≤ I ≤ 45	250–550	22	> 170/8	> 60/3
	550–800	60	> 170/2	> 60/1
	800–3,000	37	> 170/3	> 60/2
45 < I ≤ 80	250–550	67	> 170/3	> 60/1
	550–800	1,094	> 170/1	> 60/1
	800–3,000	1,536	> 170/1	> 60/1
80 < I ≤ 171	250–550	156	> 170/2	> 60/1
	550–800	1,356	> 170/1	> 60/1
	800–3,000	4,039	> 170/1	> 60/1
Total	—	**8,367**	**> 1,530/8**	**> 540/3**

* Indicates the minimum number of days needed to obtain 170 tracks.

Summary

Space Fence will be a ground-based radar designed to perform surveillance on Earth-orbiting objects. Its capabilities will increase the number of objects tracked in the current SATCAT from approximately 17,000 to over 100,000. Testing a system whose complete set of observations cannot be validated in a timely manner by existing systems presents challenges for gathering detection and accuracy truth data while ensuring a reasonable test duration. We proposed a rigorous statistical test design with candidate on-orbit test targets that span orbital limits defined by Space Fence operational requirements. We characterized Space Fence performance across the entire operational envelope by using relatively small subsets (containing no more than 1,530 satellites) of the public SATCAT grouped by altitude, inclination, and size. We identified the type and number of on-orbit test targets needed for evaluating metric accuracy, probability of track, object association, and data latency. Our approach quantifies the probabilities of meeting requirements; determines how performance varies as a function of an object's altitude, inclination, or

[5] A conjunction alert occurs when the predicted time and location at which two or more objects in space will cross orbital paths, creating the potential for a collision. Satellite operators use these alerts to assess the need for collision avoidance maneuvers.

size; estimates a 25-day test duration; and determines that modeling and simulation methods may be needed to represent 125 additional satellites. These results provide testers and users with a statistical basis of evaluation for Space Fence operational deployment decisions.

References

Cortes, L. A., and D. Bergstrom. 2012. "Using Design of Experiments for the Assessment of Surface Radar Detection Performance and Compliance with Critical Technical Parameters." 80th MORS Symposium. Colorado Springs: Military Operations Research Society.

Hutchinson, L. 2013. "How NASA Steers the International Space Station around Space Junk." *Ars Technica* (July 4). https://arstechnica.com/science/2013/07/how-nasa-steers-the-international-space-station-around-space-junk/.

Joint Range Instrumentation Accuracy Improvement Group. 1995. *Using Satellites for Radar Performance Monitoring and Calibration*. Document, U.S. Army White Sands Missile Range: Secretariat Range Commanders Council.

Martin, L. K., N. Fisher, W. Jones, J. Furumo, J. A. Heong Jr., M. Umeda, and W. Shiroma. 2011. "Ho'opono pono: A Radar Calibration CubeSat." *Proceedings of the AIAA/USU Conference on Small Satellites*. Technical Session VI, Small but Mighty, SSC11-VI-7. http://digitalcommons.usu.edu/smallsat/2011/all2011/42/.

Mochan, J., and R. A. Stophel. 1968. "Dynamic Calibration of Space Object Tracking." *The Space Congress Proceedings*. Paper 1. http://commons.erau.edu/space-congress-proceedings/proceedings-1968-5th/session-1/1/.

NASA. 2011. "Orbital Debris Frequently Asked Questions." Last updated September 2, 2011. http://orbitaldebris.jsc.nasa.gov/faqs.html.

Noll, C., and M. Pearlman. 2011. *International Laser Ranging Services 2009–2010 Report*. Recurring Report, Greenbelt: NASA Goddard Space Flight Center.

Pechkis, D. L., N. S. Pacheco, and T. W. Botting. 2014. "Identifying On-Orbit Test Targets for Space Fence Operational Testing." In 15th Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS 2014). https://amostech.com/TechnicalPapers/2014/Poster/PECHKIS.pdf.

———. 2016. "Statistical Approach to the Operational Testing of Space Fence." *IEEE A&E Systems Magazine* 31, no. 11 (November): 30–39. https://doi.org/10.1109/MAES.2016.150176.

Storz, M. F., B. R. Bowman, J. I. Branson, S. J. Casali, and K. W. Tobiska. 2005. "High Accuracy Satellite Drag Model (HASDM)." *Advances in Space Research* 36, no. 12: 2497–2505. https://doi.org/10.1016/j.asr.2004.02.020.

U.S. Air Force. 2014. "Space Fence Contract Awarded." Published June 5, 2014. http://www.af.mil/News/Article-Display/Article/485271/space-fence-contract-awarded/.

Dan Pechkis (right), a Research Staff Member in the Operational Evaluation Division (OED) of IDA's Systems and Analyses Center, holds a doctorate in physics from the College of William and Mary.

Tye Botting (left), also a Research Staff Member in OED, holds a doctorate in physical and nuclear chemistry from Texas A&M University.

Nelson Pacheco, an OED consultant, holds a doctorate in mathematical statistics from Colorado State University.

Acknowledgments: The authors are thankful for the insights provided by Kelly Avery and Kassandra Fronczyk of OED and Michael Tuley of the Science and Technology Division of IDA's Systems and Analyses Center.

The original article "Statistical Approach to Operational Testing of Space Fence" by D. L. Pechkis, N. S. Pacheco, and T. W. Botting was published in I*EEE Aerospace & Electronics Systems Magazine* 31, no. 1: 30–39. https://doi.org/10.1109/MAES.2016.150176.

Strategic Material Shortfall Risk Mitigation Optimization Model

James S. Thomason, D. Sean Barnett, James P. Bell, Jerome Bracken, and Eleanor L. Schwartz

The model solves a mathematical programming problem to identify an optimal set of strategies for mitigating the shortfalls, within cost and risk constraints set by the user.

The Defense Logistics Agency Strategic Materials office recommends strategies to Congress for managing risk that arises from shortages of strategic and critical materials that could occur during military conflicts. Strategies typically considered include material stockpiling, substitution, and spot market purchases. Heuristic methods traditionally used to select the strategies do not allow these strategies to be optimized under budgetary or nonzero expected risk constraints. We developed linear and nonlinear programming models to identify strategies that minimize expected total risk subject to upper bounds on expected total cost (budget) and to upper bounds on expected risks for individual materials.

Background and Issues

Section 14 of the Strategic and Critical Materials Stock Piling Act requires the Department of Defense (DoD), specifically, Defense Logistics Agency (DLA) Strategic Materials, to assess periodically the potential for shortfalls of strategic and critical non-fuel materials that could occur in the context of a national emergency planning scenario. The scenario consists of one or more major regional military conflicts followed by a period of military force recovery and regeneration. DoD then recommends to Congress mitigation strategies for materials that could potentially suffer shortfalls during the scenario. This paper presents the Strategic Material Shortfall Risk Mitigation Optimization Model (OPTIM-SM), which identifies shortfall mitigation strategies that would minimize expected total risk while satisfying an expected total cost constraint and constraints on the expected risks arising from possible shortfalls in individual materials.

OPTIM-SM is part of an assessment procedure that moves beyond the traditional National Defense Stockpile (NDS) planning process of estimating material shortfalls and recommending that the shortfall amounts be acquired and stored in the NDS, to a risk-based process of evaluating stockpiling along with other cost-effective alternatives for mitigating material shortfall risk.

The assessment procedure consists of the following steps:

1. Select materials of interest.
2. Estimate material shortfalls in the planning scenario.
3. Assess shortfall risk.
4. Identify promising shortfall mitigation options.
5. Assess the options' relative costs and mitigation effectiveness.
6. Identify/recommend the most promising option set—potentially within a budget constraint.
7. Begin the cycle again, as appropriate.

OPTIM-SM addresses step 6. Steps 1 and 2 constitute the traditional NDS process. Steps 3 through 5 are evaluated as part of the assessment process, and the results of steps 1 through 5 become inputs to OPTIM-SM.

The assessment procedure was implemented in support of the *Strategic and Critical Materials 2013 Report on Stockpile Requirements* (Department of Defense 2013, referred to hereafter as the 2013 report). Step 2 found 23 materials that had shortfalls, of which 19 were analyzed via OPTIM-SM. For step 4, five different possible mitigation strategies were identified and studied:

- *Stockpiling:* acquisition and storage in the U.S. NDS.
- *Buffer Stocks:* acquisition by vendor and storage in vendor-managed buffer stock inventories.
- *Export Guarantee:* reduced government guarantees of supplies of material used to produce goods to be exported during the scenario.
- *Substitution:* use of substitute materials or goods during the scenario.
- *Extra Buy:* increased U.S. buys of foreign supplies from reliable suppliers during the scenario.

Each mitigation strategy acts as an effective source of supply (or, equivalently, reduction in demand) for one or more materials in shortfall. Each strategy has a different capacity (maximum supply provided or demand reduced) and a different expected cost for each material. The effectiveness of each strategy in reducing risk depends on how much risk is created by each material shortfall and how much each strategy reduces each shortfall. The probability of the emergency scenario occurring, the negative consequences of unmitigated shortfalls, the extent to which each strategy can reduce each shortfall, and the cost of each strategy were evaluated in the preparation of the 2013 report.

Model

The OPTIM-SM model solves a mathematical programming problem to identify an optimal set of strategies for mitigating the shortfalls, within cost and risk constraints set by the user. The model is generally regarded as a linear programming problem, but there is an option for a nonlinear formulation, depending on the form of the assumed relationship between shortfall size and shortfall consequences.

Risk Measure

In the model, risk is expected risk and is defined as follows:

Expected risk = Initial shortfall risk × Expected shortfall remaining risk factor,

where:

Initial shortfall risk is the product of *Probability of war* and *Shortfall consequences.*

Probability of war is the annual probability that the emergency scenario causing the shortfalls will occur. It was estimated by subject matter experts (SMEs) for the 2013 report.

Shortfall consequences are the consequences to the nation of each material shortfall projected to occur by the DoD planning process. These consequences were estimated for the 2013 report by SMEs using a common ratio scale that focused on economic effects.

Expected shortfall remaining risk factor is *(Expected shortfall remaining/ Initial shortfall)*^{Exponent}.

Expected shortfall remaining is *Initial shortfall* minus the supply increase or demand decrease resulting from the mitigation strategies, each of which has a different capacity and effectiveness.

Initial shortfall is determined by supply and demand modeling for each material.

Exponent, which can be equal to or greater than 1, is a factor that is capable of accounting for the effect of *Shortfall consequences* that increase nonlinearly with shortfall amount. (Applications of a material that are less important would tend to be forgone before applications that are more important.) In the linear programming formulation, *Exponent* is set equal to 1.

Cost of Mitigation Options

Expected net cost formulations are devised for each mitigation strategy evaluated by the model. These formulations vary linearly with the amount of material planned to be acquired by the strategy. Discount factors are applied to all future costs and benefits. Costs given here are incurred by the U.S. government, so the Export Guarantees and Substitution strategy options have costs of zero. Net cost is particularly important for the Stockpiling strategy because it accounts for recoupment—the sale of a stockpiled material after it is no longer needed to mitigate shortfall risk. In the 2013 report, recoupment was assumed to take place after 20 years. Expected cost is particularly important for the Buffer Stocks and Extra Buy strategies because the costs of acquiring materials using those options would not be incurred unless the scenario were to occur. Several strategies have a limited capacity to mitigate shortfalls, so even a zero-cost option cannot necessarily be counted upon to eliminate any given shortfall.

Optimization Problem Formulation

The decision variables of the mathematical programming problem are the amounts of each material planned to be provided by each mitigation strategy. For the 2013 report, this means 19 materials times 5 mitigation strategies (i.e., 95 decision variables). In accordance with the formulas stated previously, each decision variable will induce its own amount of cost. Together, the decision variables for a material will lead to an expected risk for that material. The total cost of mitigation is simply the sum (over materials and mitigation strategies) of the individual costs, and the overall remaining risk is regarded as the sum of the expected risks for the individual materials.

Upper limits can be imposed on each decision variable, corresponding to the maximum amount of material that can be obtained by a given mitigation strategy. The total cost must be less than a given budget amount. The *Expected shortfall remaining risk factor* values for each given material, which generally corresponds to the fraction of *Initial shortfall* left unmitigated, can be constrained to be less than an upper limit for that material.

Overall, the optimization problem is to choose the decision variables to minimize overall remaining risk subject to:

- Upper bounds on each decision variable,
- Budget constraint on total cost of mitigation, and
- Upper bounds on *Expected shortfall remaining risk factor* values for each material.

The constraints on the *Expected shortfall remaining risk* values for each given material can be imposed to address the concern that if some shortfalls were left unmitigated or undermitigated (out of a desire to pursue the most cost-effective overall shortfall solution), those shortfalls might prevent certain industries from producing important goods. Constraining *Expected shortfall remaining risk factor* values for each separate material forces all—or at least most—shortfalls to be reduced more evenly among the different materials, which reduces the likelihood that a shortfall that might be more costly to mitigate would end up preventing the production of important goods during a crisis scenario.

Although the analysis for the 2013 report considered five specific mitigation strategies, a mitigation strategy should be recognized as any activity that can increase material supply or decrease material demand and is characterized by its cost and by the change it generates in supply or demand. Risk, upon the application of any strategy, is a function of the still-unsatisfied shortfall. Thus, other shortfall mitigation strategies, such as increased material production, material recycling, and futures contracts, can also be modeled in OPTIM-SM, as long as their attributes are characterized in the terms set out here.

Model Results

Three initial cases are considered, using the data from the 2013 report. Each case is characterized by its own upper bound for the *Expected shortfall remaining risk factor* values for each material. The upper bound values for the different cases are set to 1.00, 0.30, and 0.24, respectively. In each case, total mitigation cost (i.e., budget) is constrained at $50 million, and upper bounds on the capacities of the shortfall mitigation strategies are those used in the 2013 report.

The results show that as residual risk constraints are tightened (in the second and third cases), the shortfalls of some materials must be reduced below the levels to which they are reduced in the unconstrained, minimum total risk (for the given budget) case (the first case). That extra reduction, in turn, requires the diversion of resources that had been spent in the first case to reduce risk arising from the shortfalls of other materials. However, the further shortfall reductions in the second and third cases cost more (in dollars spent per unit of risk reduced) than the original (unconstrained) reductions. Therefore, because total cost is fixed, total risk increases.

In addition to the three optimal solution cases, three experiments were performed to show how the model responds to other changes in input data:

- First, the constraint on total cost is raised from $50 million to $80 million. The overall remaining risk becomes nearly zero.
- Second, *Probability of war*—the occurrence of the scenario—is raised significantly. As expected, the model shows that a higher *Probability of war* raises the expected costs of increased U.S. buys of foreign material supplies at the time of war. This situation can make buys of foreign material at the time of war unattractive relative to other possible shortfall mitigation strategies.
- Third, *Exponent* is set to 1.5 so that *Shortfall consequences* will increase nonlinearly with shortfall amount (representing that less important applications of a material would be forgone before more important applications). As expected, results show further reduced risk as shortfalls are mitigated below their original values.

This paper has developed a mathematical programming model to assess strategies for mitigating shortfalls of strategic and critical materials that could occur during a military conflict. The model identifies an optimal mix of such strategies: one that minimizes risk subject to constraints on total cost.

Bibliography

Bracken, J., and G. P. McCormick. 1968. *Selected Applications of Nonlinear Programming.* New York: John Wiley & Sons.

Dantzig, G. B. 1998. *Linear Programming and Extensions.* Princeton: Princeton University Press.

Denardo, E. V. 2011. *Linear Programming and Generalizations: A Problem-Based Introduction with Spreadsheets.* New York: Springer Science + Business Media.

Department of Defense. 2013. *Strategic and Critical Materials 2013 Report on Stockpile Requirements.* Washington, DC: Under Secretary of Defense for Acquisition, Technology, and Logistics (USD AT&L).

Dorfman, R., P. A. Samuelson, and R. M. Solow. 1958. *Linear Programming and Economic Analysis.* New York: Dover Publications, Inc.

Frontline Systems, Inc. 2012. "Optimization Solutions: Premium Solver Pro." https://www.solver.com/premium-solver-platform-.

Gass, S. 2003. *Linear Programming: Methods and Applications.* 3rd edition. New York: Dover Publications.

Haimes, Y. Y. 2004. *Risk Modeling, Assessment, and Management.* 2nd edition. New York: John Wiley & Sons.

McCormick, G. P. 1983. *Nonlinear Programming: Theory, Algorithms and Applications.* New York: John Wiley & Sons.

Singpurwalla, N. D. 2006. *Reliability and Risk: A Bayesian Perspective.* New York: John Wiley & Sons.

Jim Thomason (right), Assistant Director of the Strategy, Forces and Resources Division (SFRD) of IDA's Systems and Analyses Center, holds a doctorate in political science from Northwestern University.

Elly Schwartz (left), a Research Staff Member in SFRD, holds a master of science degree in management from the Massachusetts Institute of Technology (MIT).

Sean Barnett (not pictured), a former Research Staff Member in SFRD, holds a doctorate in nuclear engineering from MIT and a juris doctor degree from Georgetown University Law Center.

Jim Bell (not pictured), a former Research Staff Member in SFRD, holds a doctorate in economics from the University of California, Berkeley.

Jerome Bracken (left), an Adjunct Research Staff Member in SFRD, holds a doctor of business administration degree from Harvard University.

The original article, "Strategic Material Shortfall Risk Mitigation Optimization Model," appeared in *Military Operations Research* 20, no. 4 (November 2015), http://www.mors.org/Publications/MOR-Journal/Online-Issues.

Past Welch Award Winners

2016

"Mining Measured Information from Text"

Proceedings of the 38th International ACM CIGR Conference on Research and Develoment in Information Retrieval

A. Maiya, D. Visser, and A. Wan

2015

"Visible Signatures of Hypersonic Reentry"

Journal of Spacecraft and Rockets

J. Teichman and L. Hirsch

2014

"Mixed Models Analysis of Radar Residuals Data"

IEEE Access

C. Gaither, D. Loper, C. Jackson, and J. Pozderac

2013

"Comparison of Predicted and Measured Multipath Impulse Responses"

IEEE Transactions on Aerospace and Electronic Systems

K. Haspert and M. Tuley

2012

The Saddam Tapes: The Inner Workings of a Tyrant's Regime 1978–2001

Cambridge University Press

K. Woods and co-editors D. Palkki and M. Stout

Past Issues

2018 | Multidisciplinary Research for Securing the Homeland

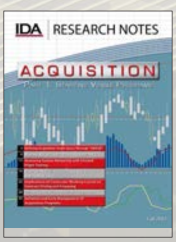
- Countering Terrorism One Technology at a Time
- Does Imposing Consequences Deter Attempted Illegal Entry into the United States
- Improving Shared Understanding of National Security and Emergency Preparedness Communications
- and more...

2016 | Acquisition, Part 2

- Cost Growth, Acquisition Policy, and Budget Climate
- Improving Predictive Value of Indicators of Poor Performance
- Root Cause Analysis of VTUAV Fire Scout’s Nunn-McCurdy Breach
- and more...

2014 | Technological Innovation for National Security

- Countering Terrorism One Technology at a Time
- Does Imposing Consequences Deter Attempted Illegal Entry into the United States
- Improving Shared Understanding of National Security and Emergency Preparedness Communications
- and more...



2013 | Acquisition, Part 1

- Defining Acquisition Trade Space Through “DERIVE”
- Supporting Acquisition Decisions in Air Mobility
- Assessing Reliability with Limited Flight Testing
- and more...

2012 | Security in Africa

- Trends in Africa Provide Reasons for Optimism
- China’s Soft Power Strategy in Africa
- Sudan on a Precipice
- and more...

IDA | RESEARCH NOTES

© Institute for Defense Analyses

4850 Mark Center Drive • Alexandria, VA 22311-1882
ida.org

@IDA_org