COMMUNICATING UNCERTAINTY INFORMATION ACROSS CONCEPTUAL BOUNDARIES

Paul Hyden

Naval Research Laboratory
Code 5585
4555 Overlook Ave, S.W.
Washington, DC 20375, USA

Elias Ioup

Naval Research Laboratory
Code 7442
1005 Balch Blvd, Room A-19
Stennis Space Center, MS 39529 USA

Stephen Russell

Naval Research Laboratory
Code 5584
4555 Overlook Ave, S.W.
Washington, DC 20375, USA

ABSTRACT

Information about data collection and modeling risks are frequently locked with information providers rather than shared with downstream information consumers. Information consumers often ingest products automatically. Without protocols to inject uncertainty, the ensemble modeling products common in the modeling discipline cannot accurately account for the input uncertainty inherent to those products. Future work to establish use cases and incorporate practitioner-driven rules and protocols for transmitting tiered uncertainty information between information product producers and consumers will advance the needs of environmental, social, and economic actors in the ensemble modeling production chain. This in turn will allow for improved error transmission throughout the decision making enterprise.

1   CRITICAL NEED IN ENVIRONMENTAL AND SUSTAINABILITY POLICY

Environmental and sustainability policy is driven by the need to effectively convey uncertainty information. As noted by Adams (2006), sustainability is ultimately a problem that must transcend environmental, social, and economic boundaries. In addressing the work by Edwards (2010), Allen (2010) noted that "We should aim to convert unknown-unknowns into known-unknowns, [and] not pretend we can eliminate them." To this end, a key driver in communicating environmental data products is the effective communication of uncertainty as a first-order information product. When there is a failure of the environmental domain to properly convey uncertainty to the social and economic domains, there is significant risk for loss of trust (Anon. 2010). At the same time, increasing domain specificity amplifies the amount of uncertainty experienced by elements of these groups (Gross 2010). Furthermore, the importance of understanding climate change has been identified as a key driver for national security policy with the United States Navy (Committee on National Security Implications of Climate Change for U.S. Naval Forces; National Research Council 2011). The significance of the impact of uncertainty conveyance is reflected in the literature where numerous climate change researchers have identified the communication of uncertainty as a challenge, including Bitz (2008).

The practical challenge of conveying uncertainty is heightened because the economic, social, and environmental domains do not always share a common understanding of uncertainty. Economic measures for uncertainty often focus on risk. For example, a seminal effort by Knight (1921) first identified the 'the ones we don't know', the type of risk that would be immortalized as the 'unknown-unknowns' of Secretary of Defense Donald Rumsfeld (U.S. Department of Defense 2002). Trivially, any effort to communicate uncertainty cannot communicate these risks, but can only motivate our downstream target audience to include the risks communicated by upstream uncertainty. Nobel prize winning approaches to the applications of the decisions between uncertain alternatives includes Tversky and Kahneman (1992). Hirshleifer (1979) conducts an excellent survey on some of the great thinking about information uncertainty in the economics domain, including discussion about modern portfolio theory from Markowitz (1959). The visceral nature of uncertainty as applied to information asymmetry that reducing your automobile trade-in are discussed in Akerlof (1970).

Social models for uncertainty certainly appear to lag the efforts in the environment and economic domains. Examples thought to be well solved in the environmental domain such as near term weather forecast still show cracks in conveying uncertainty. A recent effort by Morss, Demuth, and Lazo (2008) demonstrated the different ways that the general public interprets the implicit and explicit uncertainty in weather forecasts. The book by Patterson (2003) discussed the challenges faced by the public in digesting uncertainty in public policy.

2   GEOSPATIAL DOMAIN EFFORTS

The Joint Committee for Guides in Metrology (JCGM) publishes standards and shares these standards with partners such as ISO, IUPAC, ILAC, BIPM, IEC, IFCC, OIML, IUPAP. Two major approaches include the Guide to the Expression of uncertainty in measurement (GUM) (BIPM et al., 2010) and the associated document on propagation of distributions using a Monte Carlo method (BIPM et al., 2008). Numerous specific papers can be found to highlight the use of these standards in practice (Lira and Grientschnig 2010; Bich, Cox, and Harris 2006) or even teaching the communication of uncertainty. However, it is difficult to say whether these documents offer a conclusive solution to the challenges of communicating uncertainty.

An exciting effort is underway under the Seventh Framework Programme that addresses the need to communicate uncertainty by establishing a markup language standard called UncertML (Williams et al., 2009) for the Open Geospatial Consortium (OGC), the leading standards organization for geospatial services. UncertML is intended to describe unconditional distributions and conditional distributions where the conditioning is no longer of interest, e.g. the prediction from a Bayesian random function. UncertML is part of a broader UncertWeb project which is tasked with making "the uncertainty enabled model web a reality" (UncertWeb 2011). This includes efforts to produce APIs to support information and service models and demonstrations of UncertWeb concepts. Despite the features and generalizations provided by UncertML, it tends to emphasize random variables rather than a range of probabilistic models.

Other various efforts in communicating best practices in uncertainty as applied to climate modeling and decision-making have also been documented, including Morgan (2009). Efforts to manage uncertainty across the IPCC are documented in Swart et al. (2008) and the broader philosophical challenges to a shared uncertainty lexicon are discussed in Manning (2003). DG Robinson (1998) performs an extensive study of uncertainty analysis techniques from the period 1956-1985 applicable to complex systems.

Regan, Colyvan, and Burgman (2002) construct an ecologically focused taxonomy of uncertainty in conservation biology. An effort by Hunter and Goodchild (1993) identified the key issues of definition, communication, and management of error in spatial databases. Goodchild (1998) takes a critical look at the lack of uncertainty information in GIS systems. Land cover change modeling in Geographical Information Systems (GIS) based on historical map data is presented in presented in Leyk, Boesch, and Weibel (2005). Further efforts in the geospatial domain have included Thomson et al. (2005), Worboys (1998), AM MacEachren et al. (2005), Couclelis (2003), and Torres et al. (2004).

3 CHALLENGE ACROSS NUMEROUS DOMAINS

Transmitting uncertainty information across organizational boundaries exists as a challenge in numerous contexts. In the domains of intelligence and policy analysis, IARPA (Intelligence Advanced Research Projects Activity) has targeted the program ACE (Aggregative Contingent Estimation) to solve questions relating to the elicitation, aggregation, and communication of expert opinion used to forecast global events. Here the approach is to apply prediction markets to collectively assess event uncertainty, as promoted in Arrow et al. (2008) but with innovative weighting schemes based on forecaster traits as found in Dani et al. (2006), Ranjan and Gneiting (2010), and Cooke, El Saadany, and Huang (2008). Studies such as Tetlock (2005) establish the effectiveness of prediction markets as well as unweighted opinion pools.

Failures in communicating model uncertainty have been linked to the 2008 United States financial crisis (Salmon 2009). Salmon argues that the inability of end users to comprehend the model (Li 2000) ultimately left the financial firms vastly underestimating systemic risk. Writers such as Taleb (2008) go so far to assert that it is ultimately impossible to statistically communication uncertainty in heavy-tailed situations, an argument he advances more carefully in Taleb (2007). Visual representations of uncertainty in intelligence analysis are presented in Thomson et al. (2005).

While presented here to illustrate common problems, intelligence and financial analysis are not the only domains where the transmission of uncertainty is relevant and critical. Further, environmental, geospatial, and other domains where temporal ensemble models are common tend to be particularly susceptible to the effects of uncertainty conveyance/communication, or a lack thereof. Much of the literature in these domains focus on communicating uncertainty in a descriptive, qualitative manner rather than via analytic/quantitative means (Gigerenzer et al., 2005; Patt and Schrag 2003).

4 CONCLUSION: SIMULATION ANALYSIS EXPERIENCE AND PRACTIONER REALITY

Simulation practitioners have long worked on the boundary between rigorous statistical and probabilistic means of representing uncertainty and practitioners' needs to make reasonable simplifications. It is a core competency of a simulation modeler's toolbox to both elicit uncertainty through input modeling techniques and communicate uncertainty through output analysis techniques. A review of classic simulation texts such as Law and Kelton (2000) and Schruben (1995) show well established techniques in this area. For example, one approach in input modeling is to lead a subject matter expert through a discussion to characterize the broad features of an input process distribution using simple distributions such as the triangular or beta distributions.

Practitioners face their own challenges when trying to incorporate uncertainty analysis in their work. Even if they have the requisite self-awareness to address the uncertainty in their analysis products, the tools available may be difficult to apply. Ultimately, proper uncertainty protocols must address the challenges of the practitioner balanced against the technical capabilities of the mathematical approach. As a demonstration of ability for statistical distributions to describe uncertainty, consider the diversity in representations of a univariate random variable and the relationships between them as shown Figure 1 from Leemis and McQueston (2008). For a statistician, the figure is beautiful given its illustration of the possibilities and the array of relationships, but a practitioner is likely to be challenged, frozen, or befuddled by the paradox of choice (Schwartz 2005). On the one hand, a theoretical univariate is unlikely to clearly match the users needs, while the sheer volume of options will create a cloud of possibilities that obfuscate appropriate problem-specific alternatives. It is understandable why practitioners tend to prefer the safety of the normal distribution.

When complete, this work will support current and future efforts to add certainty measures to environmental products. By providing users a more accurate picture of the certainty of data values, it offers the following benefits: enhanced trustworthiness of decision aid data, enabling of a greater degree of data sharing, and better decisions to survey, sample or measure the environment.
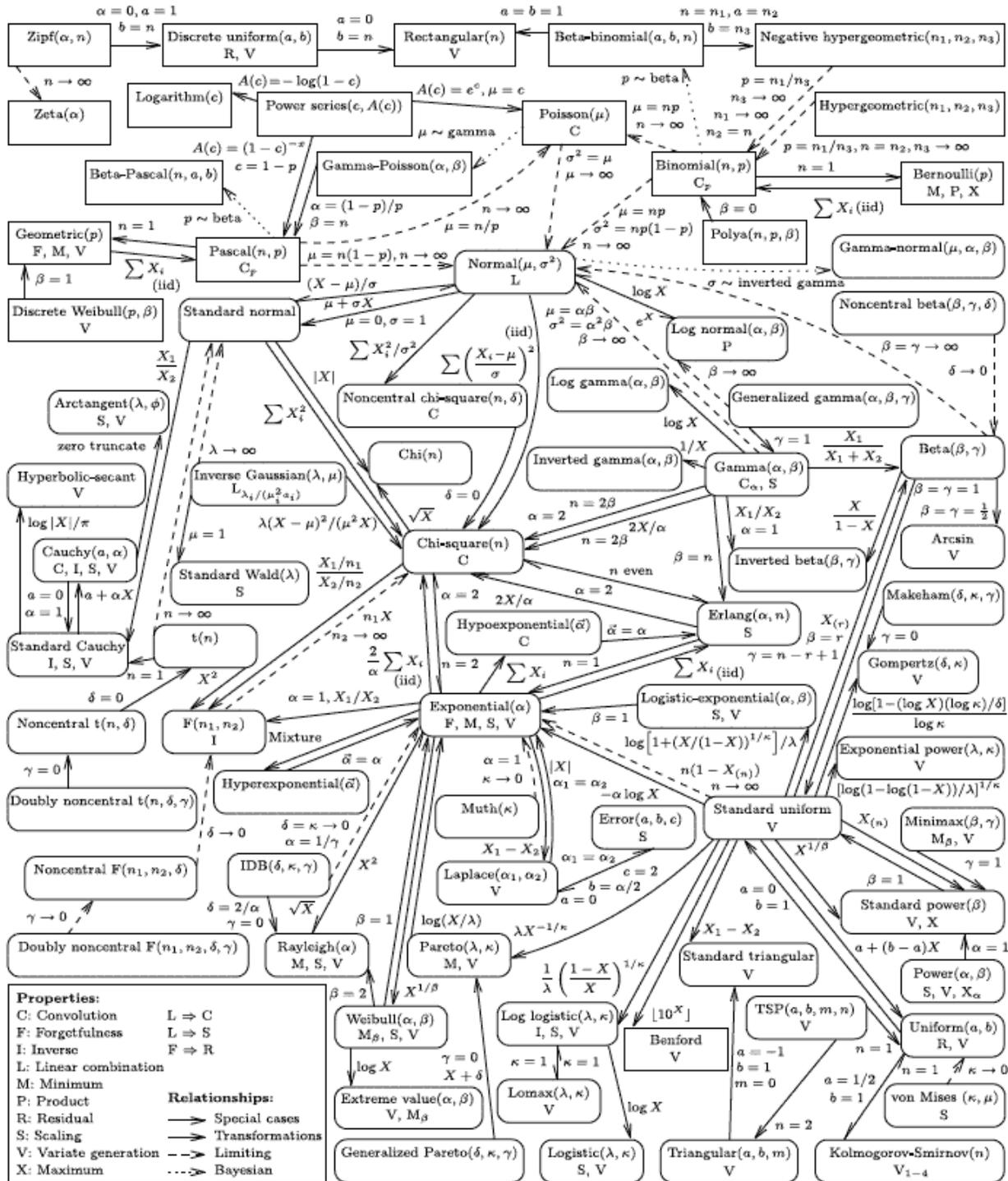
Figure 1: Univariate distribution relationships (Leemis and McQueston 2008)

ACKNOWLEDGMENTS

REFERENCES

Adams, W. M. 2006. "The Future of Sustainability: Re-thinking Environment and Development in the Twenty-first Century: Report of the IUCN Renowned Thinkers Meeting, 29–31 January 2006." *IUCN. Online at:* http://www.iucn.org/members/future_sustainability/docs/iucn_future_of_sustainability.pdf.

Akerlof, G. A. 1970. "The market for 'lemons': Quality uncertainty and the market mechanism." *The quarterly journal of economics* 84 (3): 488–500.

Allen, M. 2010. "Embracing an uncertain future." *Nature* 466 (7302) (July 1): 31. doi:10.1038/466031a.

Anon. 2010. "A Question of Trust." *Nature* 466 (7302) (July 1): 7. doi:10.1038/466007a.

Arrow, K. J, R. Forsythe, M. Gorham, R. Hahn, R. Hanson, J. O Ledyard, S. Levmore, R. Litan, P. Milgrom, and F. D Nelson. 2008. "The promise of prediction markets." *Science* 320 (5878): 877.

Bich, W., M. G. Cox, and P. M. Harris. 2006. "Evolution of the '*Guide to the Expression of Uncertainty in Measurement* '." *Metrologia* 43 (4): S161-S166.

BIPM, IEC, IFCC. ILAC, ISO, IUPAC, IUPAP and OIML. 2008. *Evaluation of Measurement Data—Supplement 1 to the 'Guide to the Expression of Uncertainty in Measurement'—Propagation of distributions using a Monte Carlo method*. Joint Committee for Guides in Metrology, JCGM 101. http://www.bipm.org/utils/common/documents/jcgm/JCGM_101_2008_E.pdf.

———. 2010. *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement GUM 1995 with minor corrections*. Joint Committee for Guides in Metrology, JCGM 100. http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf.

Bitz, C. M. 2008. "Some aspects of uncertainty in predicting sea ice thinning." *Geophysical Fluid Dynamics* 5: 6.

Committee on National Security Implications of Climate Change for U.S. Naval Forces; National Research Council. 2011. *National Security Implications of Climate Change for U.S. Naval Forces*. Washington, D.C.: The National Academies Press.

Cooke, R. M., S. ElSaadany, and X. Huang. 2008. "On the performance of social network and likelihood-based expert weighting schemes." *Reliability Engineering & System Safety* 93 (5) (May): 745-756.

Couclelis, H. 2003. "The Certainty of Uncertainty: GIS and the Limits of Geographic Knowledge." *Transactions in GIS* 7 (2): 165-175.

Dani, V., O. Madani, D. Pennock, S. Sanghai, and B. Galebach. 2006. An empirical comparison of algorithms for aggregating expert predictions. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI). Citeseer*.

Edwards, P. N. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. MIT Press.

Gigerenzer, G., R. Hertwig, E. van den Broek, B. Fasolo, K. V. Katsikopoulos, 2005. "A 30% Chance of Rain Tomorrow: How Does the Public Understand Probabilistic Weather Forecasts?" *Risk Analysis* 25 (3), 623-629.

Goodchild, M. F. 1998. "Uncertainty: The Achilles Heel of Gis." *Geo Info Systems* 8 (11): 50–52.

Gross, M. 2010. *Ignorance and Surprise: Science, Society, and Ecological Design*. MIT Press.

Hirshleifer, J. 1979. "The Analytics of Uncertainty and Information-An Expository Survey." *Journal of Economic Literature* 17 (4): 1375–1421.

Hunter, G. J, and M. F. Goodchild. 1993. Managing uncertainty in spatial databases: Putting theory into practice. In *Papers from the Annual Conference-Urban and Regional Information Systems Association*, 15–15.

Law, A.M. and W.D. Kelton. 2000. *Simulation modeling and analysis*. McGraw Hill.

Knight, F. 1921. *Risk, uncertainty, and profit*. New ed. Kissimmee Fla.: Signalman Publishing.

Leemis, L. M, and J.T. McQueston. 2008. "Univariate Distribution Relationships." *The American Statistician* 62 (1): 45-53.

Leyk, S., R. Boesch, and R. Weibel. 2005. "A Conceptual Framework for Uncertainty Investigation in Map-based Land Cover Change Modelling." *Transactions in GIS* 9 (3): 291-322.

Li, D. X. 2000. "On Default Correlation." *The Journal of Fixed Income* 9 (4): 43-54.

Lira, I., and D. Grientschnig. 2010. "Bayesian assessment of uncertainty in metrology: a tutorial." *Metrologia* 47 (3): R1-R14.

MacEachren, A.M., A. Robinson, S. Hopper, S. Gardner, R. Murray, M. Gahegan, and E. Hetzler. 2005. "Visualizing geospatial information uncertainty: What we know and what we need to know." *Cartography and Geographic Information Science* 32 (3): 139–160.

Manning, M. R. 2003. "The difficulty of communicating uncertainty." *Climatic Change* 61 (1): 9–16.

Markowitz, H. 1959. *Portfolio selection: efficient diversification of investments*. Wiley.

Morgan, M. G. 2009. *Best practice approaches for characterizing, communicating and incorporating scientific uncertainty in climate decision making*. DIANE Publishing.

Morss, R. E, J. L Demuth, and J. K Lazo. 2008. "Communicating uncertainty in weather forecasts: A survey of the US public." *Weather and Forecasting* 23 (5): 974–991.

Patt, A. and D. Schrag, 2003. "Using specific language to describe risk and probability." *Climatic Change*, 61, 17-30.

Patterson, T. E. 2003. *The Vanishing Voter: Public Involvement in an Age of Uncertainty*. Random House Digital, Inc.

Ranjan, R., and T. Gneiting. 2010. "Combining probability forecasts." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72 (1): 71-91.

Regan, H. M, M. Colyvan, and M. A Burgman. 2002. "A taxonomy and treatment of uncertainty for ecology and conservation biology." *Ecological applications* 12 (2): 618–628.

Robinson, D.G. 1998. "A survey of probabilistic methods used in reliability, risk and uncertainty analysis: Analytical techniques i." *Sandia National Lab, Report SAND98-1189*.

Salmon, F. 2009. "Recipe for disaster: The formula that killed Wall Street." *Wired Magazine* 17 (3): 17–03.

Schruben, L. W. 1995. *Graphical Simulation Modeling and Analysis: Using SIGMA for Windows*. Course Technology Press, Boston, MA.

Schwartz, B. 2005. *The paradox of choice: Why more is less*. Harper Perennial.

Swart, R., L. Bernstein, M. Ha-Duong, and A. Petersen. 2008. "Agreeing to disagree: uncertainty management in assessing climate change, impacts and responses by the IPCC." *Climatic Change* 92 (1-2): 1-29.

Taleb, N.N. 2008. *The Black Swan*. Random House, Inc.

Taleb, N. 2007. "Black Swans and the Domains of Statistics." *The American Statistician* 61 (3): 198-200.

Tetlock, P. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press.

Thomson, J., E. Hetzler, A MacEachren, M. Gahegan, and M. Pavel. 2005. "A typology for visualizing uncertainty." In *Proc. SPIE*, 5669:146–157.

Torres, R., G.R. Keller, V. Kreinovich, L. Longpré, and S.A. Starks. 2004. "Eliminating Duplicates under Interval and Fuzzy Uncertainty: An Asymptotically Optimal Algorithm and Its Geospatial Applications." *Reliable Computing* 10 (5): 401-422.

Tversky, A., and D. Kahneman. 1992. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty* 5 (4): 297–323.

U.S. Department of Defense. 2002. Defense.gov News Transcript: DoD News Briefing - Secretary Rumsfeld and Gen. Myers. February 12. http://www.defense.gov/Transcripts/Transcript.aspx?TranscriptID=2636.

UncertWeb. 2011. UncertWeb - About the Project. April 8. http://www.uncertweb.org/project.

Williams, M., D. Cornford, L. Bastin, and E. Pebesma. 2009. "Uncertainty Markup Language (UncertML)."

Worboys, M. 1998. "Computation with Imprecise Geospatial Data." *Computers, Environment and Urban Systems* 22 (2): 85–106.

AUTHOR BIOGRAPHIES

PAUL HYDEN is an Operations Research Analyst in the Information Management and Decision Architectures branch of the US Naval Research Laboratory. He received a BS, MS, and PhD in Operations Research and Industrial Engineering from Cornell University. He has worked on research projects in academia, industry, and the military using tools of analytical simulation modeling, statistics, and decision analysis in areas such automated equities trading, semiconductor manufacturing, financial engineering, hydrology assisted reproductive technology, and recently in large data and geospatial/environmental sciences. His email address is paul.hyden@nrl.navy.mil.

ELIAS IOUP is a Computer Scientist in the Geospatial Sciences and Technology branch of the US Naval Research Laboratory. He received a PhD in Computer Science from the University of New Orleans in 2011 and Bachelors degrees in Mathematics and Computer Science from the University of Chicago in 2003. Dr. Ioup's research is in the areas of geospatial/environmental data processing, Web services, and geospatial ontologies. His email address is elias.ioup@nrlssc.navy.mil.

STEPHEN RUSSELL is a computer scientist at the Naval Research Laboratory. Prior to joining NRL, he was faculty in in the Department of Information Systems and Technology Management, School of Business at the George Washington University. Before that time, he was a serial entrepreneur founding companies that specialized in information management, software development and telecommunications. He has over 20 years of industry and entrepreneurial experience in such industries as telecommunications, healthcare, and manufacturing. He has authored or co-authored articles in Expert Systems with Applications, Decision Support Systems, the Encyclopedia of Decision Making and Decision Support Technologies, and Frontiers in Bioscience. His primary research interests have been in the area of decision support systems and the application of intelligent systems to operations management. He received his Ph.D. in Information Systems from University of Maryland Baltimore County as well as a M.S. in Information Systems Management and a B.Sc. in Computer Science from the same. His email address is stephen.russell@nrl.navy.mil.