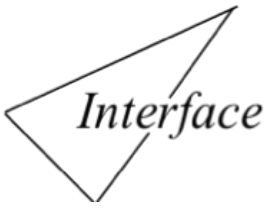
ABSTRACTS

Army Conference on Applied Statistics

19-21 October 2011, Annapolis, Maryland

Table of Contents



General Session I

Keynote Address: "Army Network Science and Statistics"
Bruce J. West, US Army Research Office

Abstract: In the past few years the Army Research Laboratory and the Army Research Office have initiated programs in Network Science that cross the boundaries separating the physical, social, communication and life sciences. A particularly significant outcome of this research into the dynamics of complex phenomena has been the crucial nature of non-stationary non-ergodic statistics. I shall outline where the traditional assumptions of stationary and ergodic statistics breakdown and why we need to extend future investigations beyond them if we hope to understand the Army's Network-of-Networks.

Social Media and the Egyptian Revolution
Maksim Tsvetovat, George Mason University

Abstract Unavailable

Contributed Session I

Network Modeling of Insurgency and Counter-Insurgency Strategies and Operations
Chris Arney, Department of Mathematics, United States Military Academy, West Point, NY 10996
Kathryn Coronges, Department of Behavioral Sciences & Leadership, United States Military Academy, West Point, NY 10996

Abstract: We model insurgency and counter-insurgency (COIN) operations with a large-scale system of differential equations and a dynamically changing coalition network. We use these structures to analyze the components of leadership, promotion, recruitment, financial resources, operational techniques, network communication, coalition cooperation, logistics, security, intelligence, infrastructure development, humanitarian aid, and psychological warfare, with the goal of informing decision makers of the options available in COIN tactics, operations, and strategy. In order to be effective, the US military must understand its capabilities and flexibility in order to match the adaptability of insurgent networks and terror cells. Our simulation model combines elements of traditional differential equation force-on-force modeling with modern social science modeling of networks, psyops, and coalition cooperation in order to understand and improve COIN strategy.

Mathematical Classification of Computational Trust Models
Craig Lennon, United States Military Academy

Abstract: A number of models exist for quantifying or computing trust in networks based on agent's actions or input. We classify computational trust by the mathematical methods it uses, and discuss how existing trust models combine these methods. We use this classification to show similarities between trust models used in social networks and in mobile ad hoc networks.

A Mathematical Model of Network Communication
James Gatewood, United States Military Academy
Donald Drew, Rensselaer Polytechnic Institute

Abstract: The behavior of a communication network can be modeled as a flow of traffic units along links connected by nodes. We derive a node/link network model and connect it to a fluid-like model of traffic flow. The discrete node/link model emphasizes packet queuing and the flow of packets from spatial point to spatial point. The model assumes that packets reside in buffers at each node, and are classified by their destination and the length of time they have resided in the buffer. An algorithm was created for packets to exit the buffer at each node according to their age and travel to the next node along a predetermined path to their destination. This algorithm calculates the rate at which packets distribute themselves to the next link in the route to their destination, assumes a source of packets originating at the node, and subtracts packets whose destination is that

particular node. The continuum model derived from this discrete flow model leads to a flow continuity equation. The continuity equation describes the density of packets as a function of time and space, so that we are able to predict changes in global flow patterns and optimal paths in the network. Solutions to the flow equations in one dimension show that if the sources are too strong or the flow is restricted, the packet density grows at the nearest upstream node. When the source strength is reduced, or when flow is restored, the buffered packets flow at capacity until the density has been reduced.

Special Session I (Army Applications)

Trends in the Incidence of Traumatic Brain Injury and Related Conditions Among Active Duty U.S. Military Personnel between 1997 and 2007

Kenneth L. Cameron[†], Stephen W. Marshall[‡], Andrew E. Lincoln[*], COL Rodney X. Sturdivant[††]

[†] John A. Feagin Jr. Sports Medicine Fellowship, Keller Army Hospital, West Point, NY; [††]Department of Mathematical Sciences, United States Military Academy, West Point, NY; [‡]Department of Epidemiology, The University of North Carolina at Chapel Hill, Chapel Hill, NC; [*]MedStar Health Research Institute, Baltimore, MD

Abstract: Traumatic brain injury (TBI) has been described as the "signature wound" associated with military operations in Iraq (OIF) and Afghanistan (OEF); however, few epidemiologic studies have examined incidence rates for TBI and related conditions in this population. The objective of this study was to examine trends in the incidence of TBI and related conditions including; 1) mild traumatic brain injury (mTBI), 2) persistent postconcussion syndrome (PCS), and 3) posttraumatic stress disorder (PTSD), among active duty US service members between 1997 and 2007. Specifically, we were interested in comparing incidence rates prior to the initiation of combat operations in Iraq and Afghanistan with subsequent years. A retrospective cohort study was conducted utilizing data extracted from the Defense Medical Surveillance System to identify all incident cases of TBI and related conditions within the study population. The primary outcomes of interest were the incidence rates of TBI, mTBI, PCS, and PTSD per 1000 person-years. Multivariable Poisson regression was used to analyze the data. Overall, there were 10,151 incident cases of TBI and 15,189,022 person-years of follow-up, for an overall incidence rate of 0.67 (95%CI: 0.66, 0.68) per 1,000 person-years. The incidence rates for mTBI, PCS, and PTSD during the study period were 6.55 (95%CI: 6.51, 6.59), 1.09 (95%CI: 1.07, 1.10), and 3.84 (95% CI: 3.81, 3.87), respectively. Trends in the incidence rate of TBI and mTBI demonstrated a steady increase during the entire study period; however, exponential increases were observed in the rate of PCS and PTSD subsequent to the initiation of OIF and OEF. The observed increase in the incidence of TBI and related conditions in this population has significant policy implications in terms of allocating appropriate healthcare resources and maintaining force readiness.

Analyzing the Effectiveness of Commander's Emergency Relief Program (CERP) Spending in Iraq and Afghanistan

MAJ Christopher Eastburg, MAJ Benjamin Thirey, and MAJ Nicholas Clark
Department of Mathematical Sciences, United States Military Academy

Abstract: The overall purpose of the Commander's Emergency Relief Program (CERP) is to enable local commanders in Iraq and Afghanistan to respond to urgent humanitarian relief and reconstruction requirements within their areas of responsibility by carrying out programs that will immediately assist the indigenous population. We identified several quantitative methodologies utilizing regression and time series that help describe the impact of different types of reconstruction spending when measured against levels of violence. Spending pattern data for different categories of reconstruction topics and violence metrics from Iraq were aggregated by month and by district. We examined the effects of various spending patterns with regards to pre-existing conditions to explore causality and the lag effects of reconstruction spending. This analysis describes mathematical relationships that may or may not be intuitive, with the goal of assisting decision makers in theater wide fund allocation and project prioritization of finite resources.

Applications of the Infinite Bootstrap: The Spare Tire Problem
MAJ Christopher Marks, MAJ Kevin Cumminsky, LTC Billy Kaczynski, and COL Andrew Glen.
Department of Mathematical Sciences, United States Military Academy

Abstract: We investigate the application of the infinite bootstrap to the spare tire problem. In the spare tire problem, we are interested in the distribution of the time to failure of a vehicle with four tires and a spare. Given sets of failure data for normal and spare tires, we employ the infinite bootstrap technique to determines the distribution that would result if a traditional bootstrap were to run an infinite number of times, thus eliminating resampling error. We note the interesting properties of the resulting distribution and compare the computational expenses of the infinite and traditional bootstraps.

Contributed Session II

Measuring Cooperation in Networks
Elisha Peterson, Johns Hopkins University Applied Physics Lab

Abstract: The framework of subset team games permits precise definitions of altruistic and selfish cooperation of both individuals and subsets of players within the game. There are multiple methods for generating subset team games on graphs, given standard metrics such as node centrality. We discuss how these methods can be applied to measure various forms of cooperation on graphs, and what insights are gained from this approach.

Modeling Cooperation in Networks, Organizations, and Systems
Brian MacDonald and Chris Arney, Department of Mathematics, United States Military Academy, West Point, NY 10996
Elisha Peterson, Johns Hopkins University Applied Physics Lab, Laurel, MD 20723

Abstract: Cooperative systems, entities working together to achieve a common goal, are important elements of our information society. This presentation uses the emerging theory of subset team games, which provides a model for the principles, relationships, and metrics of cooperative phenomena. The subset team game framework of Arney and Peterson from cooperative game theory focuses on entities working together for a common good, reflecting the team-oriented cooperation that characterizes many systems and organizations. In this presentation, we explain our use of system and organizational utility to understand teamwork. The modeling framework includes cooperation space, which provides a visual means of assessing and comparing the cooperative nature of multiple algorithms or entities geared toward the same tasks. As examples, we illustrate our framework for the movements and decisions of pursuit and evasion teams, the deployment of sensor/communication networks, and competitive sports teams.

A Network View of Pessimism in Military Units
Cadet Evan Szablowski and Dr. Kate Coronges, Department of Behavioral Sciences, United States Military Academy, West Point, NY 10996;

Abstract: The attitudes of soldiers toward their unit affect their everyday performance and how they interact with other soldiers. These attitudes are often shaped by relationships within the unit, and negative attitudes can permeate throughout the social structure. Through network science, this social structure can be quantified and visualized with the hopes of better understanding how relationships are structured within the unit and how people potentially can influence each other. Our network analysis reveals interesting aspects of the unit, such as different people of importance and sources of power in communication. When the network structure is overlaid with data on the soldier's attitudes, the spread of pessimism can be better understood through analysis network metrics. With this understanding, cynicism can be dealt with to improve the quality of life and work ethic of soldiers. This network project is continuing to look at how social media and social networks will fit into future unit operations and soldiers' professional lives.

A Spectral Technique to Explore the Asymmetric Flow of Information in Directed Networks
Karl Rohe, Statistics Department, University of Wisconsin-Madison and Bin Yu, Statistics Department, University of California, Berkeley

Abstract: Relationships in social networks are often asymmetric and can represent the flow of information from one person to another. Email messages are one example. This talk will present (1) a way to explore one aspect of the asymmetric flow of information in a directed network, (2) some asymptotic theory, and (3) some preliminary analysis of the Enron email corpus. Of key interest is the discovery of bottleneck nodes that transmit information between groups.

Contributed Session III

Inferences for Matched Case-control and Stratified Semiparametric Two-sample Density Ratio Models
Tianqing Liu[1], Zhaohai Li[2] and Yuanzhang Li[1]
[1]Walter Reed Army Institute of Research, Silver Spring, MD [2]Department of Statistics, George Washington University

Abstract: We consider inference in general stratified binary response regression models under matched case-control sampling. By linking matched case-control sampling with a stratified two-sample biased sampling problem, we show that inference for the odds-ratio parameters in a stratified logistic model under matched case-control sampling can be based on a stratified semiparametric two-sample density ratio model. Large and small sample size behavior of the maximum semiparametric conditional likelihood estimator (MSCLE) of parameters in stratified semiparametric two-sample density ratio models is studied and a semiparametric conditional likelihood ratio chi-square test for testing general restrictions on the parameters is obtained. Furthermore, we prove that the estimating function obtained from the semiparametric conditional likelihood is optimal in a class of unbiased estimating functions. Therefore, semiparametric conditional likelihood can produce valid inference for the odds-ratio parameters in a stratified logistic model under matched case-control sampling. Also, we show that MSCLE can be obtained by fitting a prospective conditional logistic regression to matched case-control data.

Disclaimer: The views expressed are those of the authors and should not be construed to represent the positions of the Department of the Army or Department of Defense.

None of the authors have any associations, financial or otherwise, that may present a conflict of interest. The effort is funded by the ILIR, the Department of the Army.

Analysis of Massive Microarray Data for Understanding Effects of Water Environments to Beetle Tissues
Guilherme V. Rocha*, Indiana University; co-authors: Karen Kafadar*, Armin Moczek, Teija Kijimoto**, Emilie Snell-Rood***, Justen Andrews****
*** Dept of Statistics, Indiana University**
**** Center for Genomic Biology, Indiana University**
**** Dept of Biology, University of Minnesota**

Abstract: Multi-cellular organisms develop different tissues through cellular differentiation regulated by gene regulatory networks. Onthophagus beetles have emerged as a promising model organism in evolutionary / ecological developmental biology, due to the expression of certain phenotypic traits in response (e.g. presence or absence of horns) to different environmental factors (e.g., availability of water and nutrition during development). Identifying the genes involved in the differentiation of tissues according to gender and nutrition factors provides understanding of the molecular mechanisms involved in tissue development which in turn can be used to model the effects of toxic chemicals in water environments on the development of organisms. A very large microarray experiment was designed to assess the expression of genes in four tissue types of male and female beetles exposed to high and low levels of nutrition. We describe the analysis of the data from this study, which involved problems of multiplicity at several levels and which inspired the search for gene "signatures"

that differentiated beetles in terms of their development. Such gene signatures can be measured efficiently and inexpensively, and enable the assessment, with improved sensitivity, of changes in water quality.

Design of Genomic Studies Using the Square Combining Table
Karen Kafadar* co-author: Guilherme V. Rocha* Dept of Statistics, Indiana University

Abstract: Many experiments are conducted based on "relative" responses, where treatment A can be evaluated only in reference to treatment B; e.g., classical weighing designs, opthamology experiments ("which is clearer, lens 1 or lens 2?"), etc. Genomic experiments often are conducted using technology based on "2-color" arrays, where the responses are measured relative to one another in a paired design. Due to the number of conditions being investigated combined with the high cost of the arrays, not all combinations can be measured. In addition, each array results in tens of thousands of responses. In this talk, I will describe the experiment in connection with the Square Combining Table, a robust method of analyzing two-way tables that was critical to informing a microarray study to identify genes responsible for responses in different tissues of Onthophagus beetles. The resulting design is specifically tailored for data that arise in the form of paired comparisons and particularly well suited when relatively high amounts of data are missing.

Detection of Subgroups for Differential Treatment Effects
Wei-Yin Loh, Department of Statistics, University of Wisconsin, Madison, WI 53706

Abstract: In the analysis of data from a clinical trial, the effect of a treatment may be found to be no different from that of a placebo across all subjects. But there may be subsets of subjects where the treatment is much better or much worse than the placebo. The existence of such subgroups is the key to the current emphasis on "personalized medicine" and "tailored therapy" in the pharmaceutical industry. In this talk, a solution based on an extension to the GUIDE regression tree algorithm is introduced. Results from computer simulations and re-analyses of published data sets show that the new method is as or more powerful than existing methods in the literature.

Contributed Session IV (Stochastic Optimization)

Combinatorial Framework for the Modeling and Solution of Probabilistic Optimization Problems
Miguel A. Lejeune, George Washington University

Abstract: We propose a new modeling and solution method for probabilistically constrained optimization problems. The methodology is based on the integration of the stochastic programming and combinatorial pattern recognition fields. It permits the fast solution of stochastic optimization problems in which the random variables are represented by an extremely large number of scenarios. The method involves the binarization of the probability distribution, and the generation of a consistent partially defined Boolean function (pdBf) representing the combination (F; p) of the binarized probability distribution F and the enforced probability level p. We show that the pdBf representing (F; p) can be compactly extended as a disjunctive normal form (DNF). The DNF is a collection of combinatorial p-patterns, each of which defining sufficient conditions for a probabilistic constraint to hold. We propose two linear programming formulations for the generation of p-patterns, which can be subsequently used to derive a linear inner approximation of the original stochastic problem. A formulation allowing for the concurrent generation of a p-pattern and the solution of the deterministic equivalent of the stochastic problem is also proposed. Computational results showing that large-scale stochastic problems, in which up to 50,000 scenarios are used to describe the stochastic variables, can be consistently solved to optimality within a few seconds.

Stochastic Network Design for Disaster Preparedness

Xing Hong (PhD student at George Washington University) Joint work with Miguel A. Lejeune and Nilay Noyan (Assistant Professor in Engineering at Sabanci University, Turkey)

Abstract: We propose a new stochastic modeling approach for the design of a network enabling an efficient response in case of a disaster. The demand for commodities and the arc capacities are defined as random variables and the inherent uncertainty is finely represented with a large number of scenarios. The problem is formulated as a probabilistically constrained programming problem. It determines the size and the location of the facilities that should be opened, the inventory levels of commodities, and ensures that the demand for commodities at all the disaster-stricken areas across the network can be satisfied with high reliability. The solution method rests on three pillars. First, we use the Gale-Hoffman inequalities to define the probabilistic feasible flow in the network. Second, we identify a preprocessing algorithm (Prekopa, Boros, 1991; Wallace, Wets, 1995) to identify the subset of non-redundant Gale-Hoffman inequalities. Third, we use a method based on the derivation of combinatorial patterns (Lejeune, 2010) to solve the stochastic programming problem. The method allows the reformulation of the stochastic problem as a mixed-integer programming problem in which the number of binary variables does not depend on the number of scenarios used for modeling the uncertainty. Computational results based on real networks describing, for example, the risk of earthquakes in the Seattle area will be presented.

Evaluating the Validity of Probability Assessments in Strategic and Realistic Environments

Prashant Doshi, Adam Goodie, Daniel Hall, Matthew Meisel and Roi Ceren; Dept. of Computer Science, Dept. of Psychology, Dept. of Statistics, University of Georgia, Athens, GA 30622, pdoshi@cs.uga.edu

Abstract: Human judgment of uncertainty suffers from various cognitive biases. In the military context, operators of autonomous vehicles such as unmanned aerial vehicles (UAV) may assess probabilities at levels that are not objectively justified. Some of the psychological issues that affect probability judgment include inaccuracy in verbally reporting probabilities and a lack of honesty in expressing the true probability assessments. In this paper, we report on a study that we conducted in order to evaluate the validity of probability assessments using a realistically simulated UAV theater.

We investigated whether direct verbal reports of probabilities are consistent with degrees of uncertainty inferred from choice data. For example, if a subjective probability of 0.5 is attributed to an outcome A, is the participant indifferent between betting on A and betting on a fair coin flip? Previous investigations into biases in judgment utilized simple settings in which the relevant probabilities were few, were given and the calculations were often simple. This study took place in the context of the Georgia test bed for autonomous control of vehicles (GaTAC), which is a hyper-realistic computer simulation framework for evaluating autonomous control of aerial robotic vehicles such as UAVs.

Subjects were recruited from the Psychology Research Pool and the Army and Air Force ROTC. Participants experienced a theater involving two UAVs. The subject is tasked with guiding her UAV I from the initial sector to a target sector without being spotted by an opponent UAV J, while being queried at intervals about her probability judgment on successfully moving in different directions and ultimately reaching the target without being spotted. Judgments of participants are validated by assessing preference relative to another random event with clear, objective probabilities.

We statistically analyzed the data to identify the bias on average in the research pool and ROTC populations, if any, and whether it differed between the two populations. Our analysis proceeded by noting that a participants believed odds are, Obelief =X'*Ostated while the odds of the random event are: ORE=X*Ostated. Here, X' is not observed and must be inferred from the X obtained from a participant's choices. This is accomplished through the use of a mixed-effect probit regression model that incorporates subject-specific random effects to accommodate between subject heterogeneity. An analysis of the data of 42 research pool participants in the experimental group reveals a 95% confidence interval of (0.0048, 17.11) for X' with an estimated population mean of 0.2866. On the other hand, data from the 9 ROTC participants in the experimental group reveals a

confidence interval of (0.9415, 1.58) with an estimated population mean of 1.22. Because both intervals contain 1.0, no systematic inflationary or deflationary bias in the uncertainty expressions is currently inferred. However, the significantly smaller interval for ROTC participants despite the small sample indicates a much better behaved population in the context of this experiment, as compared to the general research pool.

Stochastic Resource-Constrained Project Scheduling and Its Military Applications
Haitao Li, Assistant Professor of Logistics and Operations Management, College of Business Administration, University of Missouri – St. Louis. Co-author: Keith Womer, Professor of Management Science, College of Business Administration, University of Missouri – St. Louis.

Abstract: In this research, we develop computationally tractable algorithms to obtain near-optimal closed-loop policy for the stochastic resource-constrained project scheduling problem (SRCPSP). Approximate dynamic programming algorithms integrating optimization, simulation and statistical techniques are designed to tackle the "curses-of-dimensionality" of the resulting high-dimensional Markov decision process (MDP) model for SRCPSP. The research outcomes have a wide range of military applications such as mission planning, path planning, and logistics network configuration.

Contributed Session V

Imputation of Missing Round Data with Acoustic Target Systems to Estimate Dispersion Characteristics
Craig D. Andres, Jenna L. Preston, Paul S. Kruse US Army Aberdeen Test Center

Abstract: Acoustic target systems are very efficient at gathering round by round impact locations when testing weapon systems. There are cases where round dispersion is high enough or accuracy is biased enough where round data is missed, thus any dispersion results that strictly use recorded results are very likely to be underestimated. The authors propose a procedure for imputing the missing data to reduce or eliminate the bias of the dispersion estimates. Validation for the procedure will be illustrated using simulations of bivariate normally distributed data. Practical application of the procedure will be illustrated using actual test data.

Statistical Methods of Target Impact Dispersion Metric for Small-Arms MIL-HDBK Development
Douglas Ray and Chad Bareither, Armament Research, Development and Engineering Center

Abstract: Through the years there has been some controversy over how to best measure the dispersion pattern of rounds on targets. Target impact dispersion can usually be modeled adequately using the Bivariate Normal pdf, and there are a number of common metrics in use, from Extreme Spread, Radial Standard Deviation, Mean Radius, Mean Horizontal/Vertical Deviation, and Circular Probably Error. Much work has been done by statisticians such as F. Grubbs and S. Wilks regarding this topic and presented to ACAS (or the DOE conference) in the past 50 years, but with advances in targeting systems and the great need from the user for precision weapons this topic remains to be settled. Recently we have seen many requests from the user's communities and material developers for system performance to be evaluated using specific measures, and simulations have shown these metrics to have very different properties and varying levels of alpha and beta risk when the operating characteristic curves are compared. US Army - ARDEC statisticians are leading a joint-services effort to develop a comprehensive MIL-HDBK which will guide the armaments engineering community, material developer, and precision-shooter in the selection and implementation of different dispersion metrics for a variety of applications at different phases of the acquisition lifecycle.

Aluminum 7020-T651 Characteristics and Ballistic Protection, with Regression Statistics of 5083 and 7039 Aluminum, and RHA Steel
John F. Chinella, U.S. Army Research Laboratory, Weapons and Materials Research Directorate

Abstract: This study characterizes V50 protection performance of experimental 7020-T651 aluminum, (Al) a medium strength Al-4.5Zn-1.2 Mg alloy manufactured in Europe. A material properties description includes optical microstructure, certified chemical analysis, and mechanical properties. The experimental V50 results are

compared by regression predictions and statistics of 5083-H131 and 7039-T6 Al, and rolled Homogeneous Armor Steel (RHA), the Comparison Materials (CMs). The CM statistics include the mean response of Areal Density (AD)-dependent V50-performance, called the Mean Estimate (ME), and the Confidence Intervals (CIs) for the ME and Single Point Future Prediction (SPFP). T-statistic values and t-distribution cumulative probabilities standardize and summarize the 7020 V50 protection levels over the CMs' ranges of SPFP. Insights are gained on the effects of the materials deformation and failure modes by the standardized comparative performance of the protection materials. Results of this study, and literature of Al-Zn-Mg alloys, reveal 7020 has high payload capability for improved protection, and excellent to superior performance at a lower cost in more categories of manufacture and service performance than alternative fielded or candidate Al armors.

The Application of Statistical Equivalence Testing by a Military Operational Test Agency
Brittney Cates, Mathematical Statistician, Marine Corps Operational Testing and Evaluation Activity

Abstract: Null hypothesis statistical testing of differences is commonly used by Military Operational Test Agencies (OTAs) to determine if one military system is better than an alternative system. However, in some situations OTAs are asked to determine if one military system is practically indistinguishable from another system. In these situations, a statistical test of equivalence may yield more valid conclusions for evaluators and decision-makers than standard statistical tests of differences.

To address evaluation questions that ask for proof of equivalence, the Marine Corps Operational Test and Evaluation Activity has implemented the use of an equivalence test, specifically the F-test for equivalency of k normal distributions. This test is similar in principle to the Analysis of Variance (ANOVA) with the exception that it is suited for equivalency vice difference problems. Equivalency means here *equality except for practically irrelevant deviations*. This paper will describe the use of the F-test for equivalency of k normal distributions, which involves the use of the Noncentral χ^2 distribution.

This represents a new application for equivalency testing even though equivalency testing is not new. Generic drug manufacturers use this procedure in clinical trials for generic drugs when comparing to a known, well-established drug of proven efficacy and tolerability to obtain Food and Drug Administration approval.

General Session II

Vertex Nomination
Carey E Priebe, Johns Hopkins University

Abstract: The vertex nomination problem involves a graph in which vertices have associated categorical attributes. However, we observe these attributes for only a (small) subset of the vertices. Some of the vertex attributes identify vertices of particular interest. We wish to nominate from the collection of vertices with unobserved attribute (the candidate set), hoping that the nominees are truly of interest.

Contributed Session VI

Pushing Geospatial Analytics to the Edge
MAJ Christopher Eastburg, Department of Mathematical Sciences, United States Military Academy

Abstract: A mobile device app exists for just about everything now. The genius behind app development is to reduce a complex action to essential information. Computers continue to improve processing times, the volume of data continues to increase, software is growing in capability and complexity, the Army is more networked, and military organizations are increasingly joint, multinational, and involve other DoD and federal agencies. With the National Geospatial Intelligence Agency absorbing about 500 terabytes of data each month, getting the right geospatial information to the point of decision can be quite complex. Not only are geospatial analytics increasingly important to the military; geospatial tools allow better integration and sharing between first

responders, intelligence agencies, FBI, DHS, FEMA, the Department of Agriculture, NGOs, coalition nations, and the list goes on. Two recent projects at the Army Geospatial Center, the Common Ground Joint Capabilities Technology Demonstration and the Battlespace Reasoning and Awareness – Battle Command, designed tools that process, capture, and push critical information to the platform level. These geospatial analytic tools are extracted from vector terrain feature data like the US Army Theater Geospatial Database (TGD), from Digital Elevation Models (DEM) such as Digital Terrain Elevation Data (DTED), and Digital Surface Models (DSM) from Light Detection and Ranging (LIDAR). The terrain features are grouped, optimized, and analyzed to provide both commander and staff with responsive terrain information expressed in war fighter terms tailored to the mission and tasks. The general idea is to develop automated algorithms and request processes. These algorithms are capable of processing large amounts of terrain data in a rapid, consistent, and standardized manner. With the rapid production and dissemination of such tailored knowledge products, commanders and supporting staffs are able to apply judgment much more quickly throughout all phases of the decision making process and develop a thorough understanding of their operational situation.

An Analysis of Crime Using High Dimensional and Collinear Predictors
MAJ Nicholas Clark, Department of Mathematical Sciences, United States Military Academy

Abstract: In this presentation, we analyze 125 socio-economic factors to predict the crime rate in 2215 cities in the United States. The potential predictors are not only of high dimension, but are highly collinear, causing issues with standard data mining techniques. We discuss several variable selection techniques and explore their impact on our prediction error rate. This problem is applicable to US Army Officers who are working to determine where to place collection assets on an underdeveloped battlefield. We aim to explore applications of these techniques in order to assist Intelligence Officers with this task**.**

Mixed Hidden Markov Models for Joint Outcomes
LTC John C. Jackson, Department of Mathematical Sciences, United States Military Academy

Abstract: Hidden Markov models provide a means to describe processes that experience sudden rather than gradual changes over time. I present a mixed hidden Markov model to investigate associations in binary and longitudinal count data where the hidden states describe the mean of these two outcomes. The mixed model is unique in that it allows heterogeneity in the conditional model as well as in the hidden process. I introduce a unique estimation procedure where the forward-backward algorithm is embedded to compute the likelihood conditional on the random effect. Extensions to multiple outcomes are discussed as well as comparisons to other estimation methods.

Density Dependent Utilities with Transaction Costs
Julius N. Esunge, University of Mary Washington, Fredericksburg, VA, USA
Eriyoti Chikodza, Great Zimbabwe University, Masvingo, Zimbabwe, Southern Africa

Abstract: We investigate the combined singular and impulse control problem in the context of jump diffusions. Problems of this nature often arise when both fixed and proportional transaction costs are considered, for instance, in finance. We formulate and prove a verification theorem for the generalized combined singular and impulse control. This theorem establishes sufficient conditions for the existence of both the value function and optimal combined controls. An illustrative example of this result is presented.

Special Session II (Game Theory)

Network Routing in an Adversarial Environment
Nozer Singpurwalla, Department of Statistics, George Washington University

Abstract: In this talk I will describe an approach for the optimal routing of a commodity (personnel, materials or information) via a network whose links can be laden with obstacles. An archetypal scenario, is military

logistics under an IED laden environment. My approach is decision theoretic but with the novelty that the adversarial conduct of an adversary is modeled via the likelihood function. This is in contrast to using game theoretic models. The idea of using the likelihood function to describe adversarial behavior and a decision maker's strategies for defeating such behaviors is new.

Game Theory for Security: Lessons Learned from Deployed Applications
Milind Tambe, Department of Computer Science, University of Southern California

Abstract: Security at major locations of economic or political importance or transportation or other infrastructure is a key concern around the world, particularly given the threat of terrorism. Limited security resources prevent full security coverage at all times; instead, these limited resources must be deployed intelligently taking into account differences in priorities of targets requiring security coverage, the responses of the adversaries to the security posture and potential uncertainty over the types of adversaries faced. Game theory is well-suited to adversarial reasoning for security resource allocation and scheduling problems. Casting the problem as a Bayesian Stackelberg game, we have developed new algorithms for efficiently solving such games to provide randomized patrolling or inspection strategies: we can thus avoid predictability and address scale-up in these security scheduling problems, addressing key weaknesses of human scheduling. Our algorithms are now deployed in multiple applications. ARMOR, our first game theoretic application, has been deployed at the Los Angeles International Airport (LAX) since August 2007 to randomizes checkpoints on the roadways entering the airport and canine patrol routes within the airport terminals. IRIS, our second application, is a game-theoretic scheduler for randomized deployment of the Federal Air Marshals (FAMS) requiring significant scale-up in underlying algorithms; IRIS has been in use since October 2009. Similarly, GUARDS is under evaluation for national deployment by the Transportation Security Administration (TSA), and a new set of algorithms are being tested in Boston for a system called PROTECT for randomizing US coast guard patrolling; PROTECT is intended to be deployed at more locations in the future. These applications are leading to real-world use-inspired research in scaling up to large-scale problems, handling significant adversarial uncertainty, dealing with bounded rationality of human adversaries, and other fundamental challenges. This talk will outline our algorithms, key research results and lessons learned from these applications.

A Stochastic Differential Game for the Infinity Laplacian
Amarjit Budhiraja, Department of Statistics and Operations Research, University of North Carolina-Chapel Hill

Abstract: A two-player zero-sum stochastic differential game, defined in terms of an m-dimensional state process that is driven by a one-dimensional Brownian motion, played until the state exits the domain, is studied. The players controls enter in a diffusion coefficient and in an unbounded drift coefficient of the state process. We show that the game has value, and characterize the value function as the unique viscosity solution of an inhomogeneous infinity Laplace equation. This is a joint work with R. Atar.

Defending Critical Infrastructure Systems
David Alderson, Naval Postgraduate School

Abstract: Our nation faces a daunting problem: how do we fortify and defend our vast national critical infrastructure from damage inflicted by terrorist attack, natural disaster, or catastrophic accident? U.S. policy on infrastructure protection has recently emphasized probabilistic risk assessment (PRA) as a central theme in a large number of models used to allocate billions of dollars for protecting infrastructure. PRA may be acceptable for assessing risk from acts of Nature or accidents at large scale, but we have deep misgivings about using it to assess deliberate terrorist threats, and about using probabilities to represent unknowable terrorist intent. We propose an alternative where we model what we do know: how we operate our own infrastructure. We model strategic, game theoretic interactions between an attacker and a defender of an infrastructure and demonstrate how to solve such models at reasonable scale for systems of practical interest. We illustrate the technique with examples from the 150+"red team" case studies that have been conducted by military 150+officer-students at the Naval Postgraduate School.

Special Session III (Reliability)

Prediction of Remaining Life of a Fleet of Assets Based on Left Truncated and Right Censored Lifetime Data
William Q. Meeker, Iowa State University

Abstract: This talk describes a statistical procedure, based on age-adjusted life distributions, for computing prediction intervals for remaining life for individuals in a fleet of assets. We then extend these ideas to provide predictions and prediction intervals for the cumulative number of removals, over a range of time, for the overall fleet of assets. The methods will be illustrated developing predictions for the remaining life of high-voltage power transformers. Predictions for the population were required for planning maintenance and capital expenditures. Predictions for individual transformers were required for purposes of allocating special expensive monitoring equipment for those units with the highest risk of failure. In this application, as in many others, the lifetime data are complicated because transformer lifetimes can extend over many decades and transformer designs and manufacturing practices have evolved over the years. Another complicating factor is that the company's data records begin in 1980, providing information on installation and failure dates of transformers. Although the data set contains many units that were installed before 1980, there is no information about units that were installed and failed before 1980. Thus, the data are left truncated and right censored. We use a parametric model to describe the lifetime distribution of individual transformers, leading to our prediction model.

This is joint work with Yili Hong, Virginia Tech, and James D. McCalley, Iowa State University.

The Failure Rate of Mixtures
Henry W. Block, University of Pittsburgh

Abstract: For the past several years, there has been a great deal of interest in determining the shape of the failure rate of mixtures. Mixtures are important in reliability since it is widely believed that most populations are comprised of heterogeneous subpopulations. For extensive discussions concerning mixtures of survival functions see the recent books of Lai and Xie (2006) and Marshall and Olkin (2007). In this presentation we review the contributions of the author along with various co-authors.

Topics covered in this literature include the result that the asymptotic failure rate of a mixture of failure distributions converges to the failure rate of the strongest component. A similar result holds for the asymptotic shape of the failure rate. A complete catalog of all shapes of mixtures of distributions with linear failure rates has been determined. Similarly mixtures of exponential and gamma distributions are limited to having increasing, decreasing, bathtub, upside down bathtub, and modified versions of the last two types, depending on the parameters. Also, results on the cases where the mixture can have bathtub or upside down bathtub have been determined.

Estimating Component Characteristics Based on Lifetime Data from Multiple Systems
Francisco J. Samaniego, University of California, Davis

Abstract: The failure time data one collects from a sample of fielded systems provides indirect information about the performance of the system's components. Since it is often difficult to create or simulate field conditions in laboratory settings, the process of drawing inferences about component characteristics from data on system performance is of practical importance. Under the assumption that the system under study has components with independent lifetimes and a common distribution F, Bhattacharya and Samaniego (Naval Research Logistics, 2010) identified the asymptotic behavior of the nonparametric MLE of the underlying component reliability function $\overline{F}(t)$ and demonstrated the estimator's good performance, even for moderate sample sizes. In this paper, this problem is nontrivially extended to that of estimating $\overline{F}$, under the same assumptions, but based on independent samples from m coherent systems, with $m \geq 2$. Two estimators are studied. The first is a natural estimator based in the treatment of each estimation problem separately – a convex combination of the m individual estimators of $\overline{F}$. The estimator is intuitively appealing and is both consistent

and asymptotically normal. As expected, it improves substantially of any of the individual estimators of $\overline{F}$. The second estimator is obtained by solving multiple point-wise maximum likelihood problems, i.e., one for each fixed time point t, and cobbling together these separate estimates to obtain an overall estimator of the function $\overline{F}$. We show that the resulting estimator is a legitimate survival curve and that it is consistent and asymptotically normal. The large sample variances of the two estimators are compared, and one is found to be uniformly superior. The performance of the winning estimator in investigated via simulation for moderate sized samples from three particular systems. This work in joint with Peter Hall and Yin Jin.

Contributed Session VII (Inference Modeling)

A Decision Theoretic Approach to Stochastic Linear Programming
Joshua Landon and Nozer Singpurwalla, Department of Statistics, The George Washington University, Washington, DC. 20052

Abstract: This talk is based on the premise that stochastic optimization problems are de facto problems of decision making under uncertainty, making them akin to problems in statistical decision theory. The satisfactory resolution of such problems calls for two ingredients: an overarching probabilistic structure, and a meaningful utility function. Here, we consider the special case of a simple linear program with uncertainties in the constraints and in the objective function to demonstrate the workings of our point of view. This talk is expository and represents work in progress. As is the tradition in ACAS meetings, participant comments are welcome.

Stable Matchings with Additional Objectives
Craig Lennon, United States Military Academy

Abstract: In this paper, we consider a version of the stable matching problem, but with multiple objectives. Given n applicants and n employers, each with complete preference lists, we want to find matchings which are stable, respect the aggregate preferences of the applicants and employers, and minimize / maximize any additional objectives. We develop a genetic algorithm which searches among the lattice of stable matchings, design and execute an experiment to improve it, and then compare its performance to a single variable optimization algorithm which optimizes each objective function independently.

A Mixed-integer Programming Model for Enforcing Priority List Policies in Markov Decision Process Models
Laura A. McLay, Department of Statistical Sciences & Operations Research, Virginia Commonwealth University

Abstract: Optimal dispatching policies for server-to-customer systems can be identified using Markov decision process models and algorithms, which indicate the optimal server to dispatch to each customer type in each state. Since the optimal policy is dependent on state variables, it may not always correspond to a simple set of rules. Restricted policies that use a priority list policy for each type of customer may be desirable for use in practice. In this paper, we demonstrate the optimal priority list policy can be identified by using constrained Markov decision processes. We present an efficient model for identifying optimal priority list policies using mixed integer programming models. The benefit of this approach is that it does not extend the Markov decision process state space and relies on standard algorithms.

Recent Advances in Reliability-Based Design Optimization for Broader Army Applications
K.K. Choi Ikjin Lee, Liang Zhao, and Yoojeong Noh, Department of Mechanical and Industrial Engineering, College of Engineering, The University of Iowa, Iowa City, IA and David Gorsich and David Lamb,US Army TARDEC, Warren, MI

Abstract: In deterministic design optimization, the designs are typically pushed to the design constraint bounds to obtain an optimum design, leaving little or no room to consider input uncertainties such as manufacturing dimensional variability, material property variability, operational load variability, and simulation model uncertainty. Thus, the optimum designs obtained without consideration of these input uncertainties are unreliable. During last two decades, computation methods have been explored to incorporate uncertainty analysis at an affordable computational cost and, more recently, to carry out design optimization with the additional requirement of reliability, which is referred to as reliability-based design optimization (RBDO).

Over a number of years, the University of Iowa (UI) research team has been working with the U.S. Army TARDEC to develop a simulation-based RBDO method to obtain optimum designs to meet the durability and reliability requirements while minimizing Army ground vehicle weight. The UI research team has integrated the UI developed durability analysis code *DRAW* and sensitivity analysis code *DSO* with commercial codes such as CAD-Pro/E, FEA-NEiNastran, multibody dynamics code DADS, and design optimization code DOT, along with the UI developed *Most Probable Point (MPP)-Based Dimension Reduction Method (DRM)*. This integrated software system was successfully demonstrated to provide reliable optimum designs with significantly reduced weight and improved fatigue life of U.S. Army High Mobility Trailer drawbar, Stryker A-arm, and HMMWV A-arm components.

With the success of the Iowa developed MPP-Based DRM and software tools, the UI research team was looking for an extension of the RBDO method that can be applied to broader problems beyond the durability. Under the joint funding of ARO and TARDEC, the UI team started extending the RBDO methods by developing a sampling-based system level RBDO method to support broader application problems. For the sampling-based RBDO, the stochastic sensitivity analysis is developed using the score function and copula to compute sensitivities of probabilistic constraints with respect to independent and correlated random design variables (TARDEC funding); and the Dynamic Kriging (DKG) method is developed for accurate surrogate modeling (ARO funding). The new RBDO tool is planned to be applied for reliability analysis and RBDO of multi-functionality (durability and safety) of Army ground vehicle structures, wind energy system, ship motions in rough seas; and integration of manufacturing process and RBDO of armor.

Contributed Session VIII

Maintainability Data Decision Methodology (MDDM)
John Nierwinski, Jr., Army Material Systems Analysis Activity

Abstract: Organizations within the U.S. Army [e.g. Communications-Electronics Command (CECOM)] and other government organizations have the need to evaluate Maintenance Manpower requirements for systems (i.e. Power Generators, etc.) where fully developed maintenance data is NOT available. Maintenance Manpower requirements are computed by multiplying an estimated maintenance ratio (man-hours per usage) by a one year wartime usage, which results in a total number of recommended maintenance man-hours. Army organizations need to know how much maintenance ratio data needs to be collected until the sample can be used to generate maintenance manpower requirements.

AMSAA developed a maintainability data decisioning methodology (MDDM) which determines if enough sample maintenance data exists in order to infer the true fleet maintenance ratio (MR). This will allow the Army to make manpower requirement determinations. MDDM uses a systems aging model, parametric & nonparametric empirical Bayes models, two stochastic inferencing and four coverage validation models using nonparametric & parametric bootstrapping, percentile method with bias correction & acceleration using jackknifing, Monte Carlo simulation, and other stochastic modeling techniques & processes.

AMSAA has applied MDDM to CECOM Power Generators. Current results show all of the generator groups require more data to infer a fleet MR. Accuracy and precision of the inference is what determines if enough data exists. A coverage validation model is used to measure accuracy and precision is measured by the size of the inference. MDDM is a decisioning process and can be applied to other Army weapon systems where the need exists.

The Navy Officer Lateral Transfer Process and Retention: A Matched Analysis
Yevgeniya (Jane) Pinelis and Ann Parcell, CNA (Center for Naval Analyses)

Abstract: Managing the officer corps well depends on accurately estimating officer losses from the Navy. Many factors affect loss rates, including the ability of officers to transfer laterally from one community to another. In order to successfully lateral transfer, officers must apply and be approved by a lateral transfer board. We examine the loss rates of officers who applied for lateral transfer but were disapproved and compare them with the loss rates of those who applied and were approved.

Whether the officer is 'best and fully qualified' is part of the board decision, and data show that there are large differences on observed characteristics between officers approved and denied by the board. To suggest a link between the approval decision and leaving the Navy, we need to properly adjust for these differences.

First, we use logistic regression to model the loss of lateral transfer applicants from the active-duty Navy after the lateral transfer board as a function of the board approval decision, controlling for the effect of their military characteristics and demographics. Our findings support the hypothesis that officers who were disapproved for lateral transfer have a higher 36-month loss rate than those officers who were approved. Although we show the link between the board result and an officer's decision to leave the Navy, regression analysis is not sufficient to assert causality. We thus supplement our analysis with a method often used in observational studies to search for causality: propensity score matching.

We divide officers into strata using full matching on the estimated probability of board approval (propensity score). We ensure that each stratum has at least one accepted and at least one rejected officer, while using a formal, recently developed covariate balance diagnostic to guide the choice of the matching structure. Our resulting stratification balances all observed variables important to the outcome (e.g. officers' gender, promotion information, etc.) between the two groups in a way that we might expect to see had the officers been randomly selected for acceptance or rejection. By recovering the 'hidden' randomized experiment from this observational study, we link the board results to the officers' decisions to leave the Navy in a causal way. Our final analysis results are slightly different from those based on logistic regression, although the main conclusions about variables important to the probability of leaving remain the same.

A Statistical Analysis of Proposed Military Retirement Plans
CDT Joseph Lucas, MAJ Kevin Cummiskey, and Dr. Robert Wooster Department of Mathematical Sciences, United States Military Academy

Abstract: The Department of Defense is considering changes to the military retirement system. One proposal would replace the current pension system with a 401(k) style retirement savings plan that makes monthly contributions to a service member's Thrift Savings Program account. We attempt to establish conditions (years of service, discount rate, etc.) that make compensation under the current and proposed systems equivalent at age 65 for the career military officer. Using statistical modeling, we estimate the likelihood of these conditions. In addition, we investigate the sensitivity of the cost of the proposed plan to the distribution of officers (rank and time in service).

Clinical Session I

Evaluation of Weapon System Performance in a Limited-Data Environment: Estimating the Probability of Engagement Success (PES) for Missile Defense Systems

Carl C. Gaither, III, Kenneth Brown, Dawn Loper, Michael Luhman, Jasmina Marsh, Robert Salow
Institute for Defense Analyses, Alexandria, Virginia 22311

Abstract: Estimating the probability of engagement success (PES) for missile defense systems is a major challenge for the defense test and evaluation community. Missile defense flight tests are very expensive, complex, and logistically demanding endeavors. This limits the amount of testing that can be reasonably accomplished, and in turn the amount of data available to assess the effectiveness of missile defense systems. Despite this lack of data, reliable estimates of PES are crucial for foreign and defense policy decision-making. Key stakeholders who can benefit from PES estimates include the defense acquisition community, Congress, and most importantly, the military operators of these systems.

We have developed a methodology for estimating the probability of engagement success (PES) for missile defense systems for the portion of their engagement battlespace sampled through developmental and operational testing. This methodology is based on the concept of DuPont analysis commonly used by the corporate finance community to evaluate and/or predict a company's return on equity. Our methodology incorporates ground- and flight-test data to the maximum extent possible. It expands on traditional success/failure (binomial) models that rely solely on end-to-end flight tests (that is, flight tests that start with a target launch and end with an interceptor either hitting or missing the target) by incorporating partial test results such as target tracking-only exercises, interceptor-only flight tests, ground-based lethality tests and command and control ground tests. The advantages of this methodology include: (1) it is straightforward and offers clear insight into the missile defense functions that drive performance; (2) it identifies over- and under-performing missile defense functions; (3) it can incorporate historical testing, component-level testing, and end-to-end testing of missile defense systems; and (4) it lends itself to adaptation as missile defense systems mature. Despite these advantages, additional study is needed to fully understand the statistical properties of this methodology. We will present a description of this methodology as well as some preliminary numerical studies of its performance for realistic test outcomes to stimulate a discussion of potential improvements and/or modifications that could enhance this methodology to either provide a better theoretical statistical foundation or to produce more robust results.

Uninformative Confidence Bands in a Probit Analysis

David W. Webb, U.S. Army Research Laboratory

Abstract: In probit analysis, one attempts to estimate the relationship between a stimulus variable and the probability of a successful outcome by assuming that the relationship is a Gaussian cumulative distribution function. When a zone of mixed results (ZMR) is present in the data, one can use maximum likelihood estimation to determine the Gaussian parameters (μ and σ), which are most likely to have given rise to the data. Based on the asymptotic normal distribution of these maximum likelihood estimators confidence intervals for the success probability can be calculated. A plot of the confidence limits over a range of stimulus values yields confidence bands for the estimated curve. In most cases the confidence limits follow a monotonically increasing relationship with stimulus, which intuitively has appeal since the estimated curve shares the same property. However, in some cases the confidence bands can expand outside of the ZMR, losing the monotonicity property and asymptoting towards the uninformative confidence interval of [0, 1]. Conditions under which this phenomenon may occur and methods for detecting it are explored.

General Session III
Chair:

Heavy Tailed Statistics for Modeling Data Network Sessions
Sidney Resnick, School of Operations Research and Information Engineering, Cornell University

Abstract: A session is a higher order entity in data network modeling resulting from amalgamating packets, connections, or groups of connections according to specified but not unique rules. For example, the flow of packets past a sensor can be amalgamated into sessions using a threshold rule based on gaps between packet arrivals. We review and illustrate heavy tailed probability and statistical modeling based on sessions. Statistical analysis of these sessions based on packets is complex: session duration (D) and size (S) are jointly heavy tailed but average transmission rate (R = S/D) is sometimes not heavy tailed and arrival times of sessions is typically not Poisson. By segmenting sessions in various ways, for instance using a peak rate covariate, we find conditional on a peak rate decile, within this decile segment session initiations can be modeled as Poisson. Some reasons why Gaussian traffic seems ubiquitous are discussed. This is a joint work at various times with Bikramjit Das, ETH Zurich, Janet Heffernan, Lancaster, UK, Luis Lopez-Oliveros (Cornell University and Murex), Bernardo D'Auria (University Carlos III of Madrid), Thomas Mikosch (University of Copenhagen), Holger Rootzen (Chalmers University, Goteborg, Sweden), A. Stegeman (Groningen, The Netherlands)

General Session IV

Nested Latin Hypercube Designs
Peter Qian, University of Wisconsin-Madison

We propose an approach to constructing nested Latin hypercube designs intended for multi-fidelity computer experiments, sequential experiments with computer models and large-scale computer experiments. A nested Latin hypercube design with two layers is defined to be a special Latin hypercube design that contains a smaller Latin hypercube design as a subset. Our method is easy to implement and can accommodate any number of factors. We extend this method to construct nested Latin hypercube designs with an arbitrary number of layers. Sampling properties of the constructed designs are derived. Also will be discussed are several refinements of nested Latin hypercube designs, including nested orthogonal array-based Latin hypercube designs, Sudoku-based space-filling designs and sliced Latin hypercube designs.

Contributed Session IX

Large Fractional Factorial DIODE DOE
Curt Laumann & Bill Thomas, Raytheon Missile Systems

Abstract: The performance of a wirebonding process was fully optimized through the application of recurring DOE's. Process performance was measured by both strength of the wirebond (a continuous variable) and ability of the wire to "stick" (pass / fail). The DOE designs were novel: three factor, seven level fractional factorial, with five replications. Data were obtained to identify an extremely narrow process window yielding strength Cpk>1.0, and "sticking" yield of ~100%. Previous attempts to optimize this process using one-factor-at-a-time methods were not successful."

Design of Experiments in Highly Constrained Design Spaces
Laura J. Freeman, Institute for Defense Analyses

Abstract: In operational testing there are often constraints that limit the applicability of standard statistical designs. These constraints may be imposed by safety restrictions, physics, or that a certain combination of factors lacks operational meaning. This presentation covers common examples of constrained design spaces in operational testing. I provide statistical designs that meet the constraints. Additionally, these designs provide some degree of statistical optimality. The presentation concludes by highlighting general selection criteria for nonstandard statistical designs.

Prioritizing Unaided Human Search in Military Simulations
CPT James K. Starling, Department of Mathematical Sciences, United States Military Academy

Abstract: Search and Target Acquisition (STA) in military simulations is the process of first identifying targets in a particular setting, then determining the probability of detection. This study will focus on the search aspect in STA, particularly with unaided vision. Current algorithms in Combat XXI use an antiquated windshield wiper search pattern when conducting search that is based on studies using aided vision, such as binoculars or night vision devices. Very little research has been conducted for unaided vision and especially in urban environments. Based on empirical data, this study will show that a human observer using the naked eye does not uniformly search a given urban environment. It will then propose an algorithm to be used in military simulations that will more closely reflect human behavior, given that a target is not immediately identifiable.

Contributed Session X

Distributed Modal Identification Using Restricted Auto-Regressive Models
Guilherme V. Rocha, speaking, Statistics Department, Indiana University, Shamim Pakzad, Department of Civil and Environmental Engineering, Lehigh University, Bin Yu, Statistics Department, UC Berkeley

Abstract: Advances in Wireless Sensor Networks (WSN) technology have provided promising possibilities in detecting a change in the state of a structure through monitoring its features estimated using sensor data. The natural vibration properties of the structure are a set of features commonly used for this purpose and are often estimated using a multivariate autoregressive model (AR model) for the measured structure's response to ambient vibrations. Fitting a multivariate AR model to the observed acceleration requires the computation of the lagged covariance between the measurements in all nodes. The resulting volume of data transmission causes significant latency due to the low data bandwidth of WSNs in addition to having a high transmission energy cost. In this paper, a set of restrictions to the estimation of the AR model is introduced. Such restrictions significantly reduce the volume of data flowing through the WSN thus reducing the latency in obtaining modal parameters and extending the battery lifetime of the WSN. A physical motivation is given for the restrictions based on a linear model for a multi-degree of freedom vibrating system. Stabilization diagrams are compared for the restricted and full AR models fitted using data simulated from linear structures and real data collected from a WSN deployed on the Golden Gate Bridge. These stabilization diagrams show that the estimated modes using the restricted AR models are of comparable quality to that of the full AR model while substantially reducing the volume of transmitted data.

Symbolic ARMA Model Analysis
Keith H. Webb & Lawrence M. Leemis, Department of Mathematics, College of William & Mary

Abstract: Many problems in time series analysis rely on approximate values from Monte Carlo simulations or the central limit theorem rather than exact results. The computer algebra system Maple and the APPL (A Probability Programming Language) package can help find exact results that would be impractical to compute by hand (Glen, et al. 2001). This paper describes our time series extension to APPL which contains tools for analyzing ARMA (autoregressive moving average) models, including but not limited to finding exact values for

autocorrelation, partial autocorrelation, covariance, and forecasts. The extension can be used to find many new results, particularly for ARMA models which are not covariance stationary.

A New General Purpose Method of Time Series Forecasting
Michael Frey, Bucknell University, and Kyle Caudle, South Dakota School of Mines and Technology

Abstract: Forecasting future values of a time series is one of the most difficult and conceptually challenging problems facing statisticians, and many methods, traditional regression and its variants, Box-Jenkins ARIMA forecasting, and neural networks most prominently, have been devised to address this challenge. We describe a new general purpose method of statistical forecasting based on the concept of a flow field, which describes change in the time series. This new method of forecasting uses standard tools of semiparametric regression and Gaussian process regression to estimate the flow field and then builds forecasts by iterative application of the estimated flow field. This new method of forecasting is relatively computationally efficient and readily yields standard errors with its forecasts. We give an elementary presentation of this new forecasting method, demonstrate it, and compare it with standard methods. In the course of this presentation, we explore fundamental issues associated with the challenge of forecasting, distinguishing near-future statistically reliable forecasts and far-future plausible forecasts.

General Session V

Info-Metrics and Estimation of Extreme Events
Amos Golan, Info-Metrics Institute and Department of Economics, American University

Abstract. Reasoning under conditions of incomplete information challenges researchers across disciplines. Info-metrics is the science and art of processing information. It crosses the boundaries of all sciences. It is inherently interdisciplinary and provides the universal mathematical and philosophical foundations for inference with finite and incomplete information. Info-metrics is the intersection of information theory, inference, mathematics, statistics, complexity, decision analysis and the philosophy of science. The study of info-metrics helps in resolving a major challenge for all scientists and all decision makers of how to reason under conditions of incomplete information. In this talk I will discuss briefly the state of info-metrics in general and will concentrate on the class of Information-Theoretic Estimators and relate it to the study of extreme events in particular. Two examples of extreme events analyses will be presented.