

# **38th Symposium on the Interface of Statistics, Computing Science, and Applications 2006 (Interface 2006)**

## **Massive Data Sets and Streams**

**Abstracts**

**Pasadena, California, USA  
24-27 May 2006**

**ISBN: 978-1-62276-715-1**

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2006) by the Interface Foundation of North America  
All rights reserved.

Printed by Curran Associates, Inc. (2013)

For permission requests, please contact the Interface Foundation of North America  
at the address below.

Interface Foundation of North America  
PO Box 7460  
Fairfax Station, VA 22039

Phone: (703) 993-1212  
Fax: (703) 993-1700

[interface@galaxy.gmu.edu](mailto:interface@galaxy.gmu.edu)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2634  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# TABLE OF CONTENTS

<b>Statistical Issues in Model Validation</b> .....	1
<i>David Banks</i>	
<b>Computer Experiment Designs to Achieve Multiple Objectives</b> .....	1
<i>Leslie M. Moore</i>	
<b>A Hierarchical Model for the Reliability of an Anti-aircraft Missile System</b> .....	1
<i>Shane Reese</i>	
<b>Quantitative Tools for the Counter-IED Campaign</b> .....	2
<i>Thomas J. Sullivan</i>	
<b>Systems Reliability and Experiment Planning</b> .....	2
<i>Alyson Wilson</i>	
<b>Alcohol Problems: System Complexity as a Challenge to Data Systems</b> .....	3
<i>Paul Gruenewald</i>	
<b>Developing an Analytical Approach to Massive Data Sets on Alcohol Use</b> .....	3
<i>William F. Wieczorek</i>	
<b>Mixture Modeling for Genome-wide Localization of Transcription Factors</b> .....	4
<i>Sunduz Keles</i>	
<b>An Infinite-state Generalized Hidden Markov Phylogeny for Multi-species Regulatory Module Detection</b> .....	4
<i>Jon McAuliffe</i>	
<b>Dense SNP Genotyping on Chromosome 22 in 200 Persons from Each of 12 Populations</b> .....	5
<i>Susan Service</i>	
<b>Adaptive Analysis of Neural Plasticity in the Hippocampal System</b> .....	5
<i>Loren Frank</i>	
<b>Dynamic Imaging of Brain Function with MEG and EEG</b> .....	6
<i>Matti Hamalainen</i>	
<b>Current Developments in a Cortically Controlled Brain Machine Interface</b> .....	7
<i>Nicho Hatsopoulos</i>	
<b>Uncertainty in DNA-Protein-Regulation Cycle and a New Model of Its Processes: Dedicated to the Vision and Memory of Paul Silverman</b> .....	7
<i>Arnold F. Goodman, Claudia M. Bellato, Lily Khidr</i>	
<b>Virtual Astronomy, Information Technology, and the Revolution in Scientific Methodology</b> .....	8
<i>George Djorgovski</i>	
<b>Quasi-adaptive Confidence Bands and the Dark Energy Equation of State</b> .....	9
<i>Christopher Genovese</i>	
<b>Modern Statistical Methods for GLAST Event Analysis</b> .....	N/A
<i>Robin Morris</i>	
<b>Data Assimilation for Weather and Climate Models with Ensemble Filters</b> .....	10
<i>Jeffrey L. Anderson</i>	
<b>Assessments of Climate Change using Regional Climate Models</b> .....	10
<i>Steve Sain</i>	
<b>Management of Massive Seismic Data Sets: Activities within the IRIS Data Management System</b> .....	11
<i>Tim Ahern</i>	
<b>Process, Pattern, Prediction: Using Space Data to Understand and Predict Complexity in Driven Dynamical Earth Systems</b> .....	12
<i>John Rundle, Andrea Donnellan</i>	
<b>Some Ideas About the Use of Global Sensitivity Analyses in Geoscience Applications</b> .....	12
<i>Carola Tiede</i>	
<b>Full Matrix Visualization of Network Traffic</b> .....	14
<i>Steve Marron</i>	
<b>Semi-experiment Investigations of Network Traffic</b> .....	14
<i>David Rolls</i>	
<b>Stream-based Data Analysis and Visualization of High-speed Wide-area Wireless Networks</b> .....	14
<i>David A. James, A. Buvanewari, Lawrence M. Drabeck, John M. Graybeal, W. Mike Macdonald</i>	
<b>Measurement and Analysis of Large-Scale Stochastic Service Systems</b> .....	15
<i>Pat Tendick</i>	
<b>Classification of High-dimensional Data by Semiparametric Generalized Regression Models</b> .....	15
<i>Marlene Muller</i>	

<b>A Novel Application of Functional Data Analysis to High-resolution Data from Environmental Epidemiology</b> .....	16
<i>Joan Staniswalis</i>	
<b>Efficient and Fast Spline-backfitted Kernel Smoothing of Additive Regression Models</b> .....	17
<i>Lijian Yang</i>	
<b>Using Ensemble Statistical Procedures for Imputing Homeless Counts for Los Angeles County</b> .....	17
<i>Richard A. Berk</i>	
<b>Integration of Deterministic and Statistical Algorithms for Retrieval and Analysis of Geophysical Parameters</b> .....	18
<i>Zoran Obradovic</i>	
<b>What is Special About Mining Spatial Datasets?</b> .....	18
<i>Shashi Shekhar</i>	
<b>A Streaming/Sequential Algorithm for Learning Sparse Classifiers</b> .....	19
<i>Suhrid Balakrishnan</i>	
<b>Analysis of Streaming Text</b> .....	19
<i>David Marchette</i>	
<b>A General Framework for Mining Massive Data Streams</b> .....	20
<i>Pedro Domingos</i>	
<b>Time-evolving Adaptive Regression</b> .....	20
<i>William F. Szewczyk</i>	
<b>Dense Information Grinding</b> .....	N/A
<i>Olivier Verscheure</i>	
<b>Multi-Mode Co-clustering of Different Text Data Attributes</b> .....	21
<i>John Rigsby</i>	
<b>Text Mining Using Statistical Topic Models</b> .....	21
<i>Padhraic Smyth</i>	
<b>Literature-based Discovery for the Identification of New Methods of Water Purification</b> .....	21
<i>Jeffrey L. Solka</i>	
<b>Monitoring the Converged Network</b> .....	22
<i>Akshay Adhikari, Lorraine Denby, Jim Landwehr, Jean Meloche</i>	
<b>The Evolution of Science: Time Series Models of Scientific Journals and Other Large Text Databases</b> .....	22
<i>John Lafferty</i>	
<b>Earth Science Data: A Look Ahead</b> .....	23
<i>Anthony Freeman</i>	
<b>Problems in Meta-Analyses: Studies are many, but cases are few</b> .....	24
<i>Tom Torda</i>	
<b>Statistical Techniques Applied to the Surface Topography Analysis of Firearms-related Forensic Evidence</b> .....	25
<i>Benjamin Bachrach</i>	
<b>Statistical Methods for Assessing the Feasibility of a National Ballistics Imaging Database</b> .....	25
<i>James J. Filliben</i>	
<b>Author Index</b>	