

41st Symposium on the Interface of Computing Science and Statistics 2010 (Interface 2010)

Computational Statistics and Human Behavior

Abstracts

**Seattle, Washington, USA
16-19 June 2010**

ISBN: 978-1-62276-718-2

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© (2010) 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

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
Web: www.proceedings.com

Program for Interface 2010

Wednesday, June 16, 2010



7:30 am – 5:00 pm	Registration	
8:00 am – 12:00 noon	Short Course: Agent Based Modeling and Simulation	
1:30 pm – 5:30 pm	Short Course: Statistical Natural Languages and Text Mining	
6:30 pm – 8:00 pm	Board of Directors Meeting (Closed)	Baker
8:00 pm – 10:00 pm	Evening Mixer	Cascade Ballroom

Thursday, June 17, 2010

7:30 am – 5:00 pm	Registration	
8:00 am – 9:45 am		
	Key – 1 Keynote Address: Adrian Raftery	Cascade Ballroom
	<i>Probabilistic Projections of HIV Prevalence Using Bayesian Melding with Incremental Mixture Importance Sampling (IMIS)</i>	
9:45 am – 10:15 am	Morning Break	
10:15 am – 12:00 noon	Technical Sessions	
	Inv –1 Uncertainty and Simplifications in ABMs	St. Helens
	Organizer and Session Chair: Georgiy Bobashev	
	<i>Using Simulation Models for Decision Making under Deep Uncertainty</i> , Gary Klein, MITRE	
	<i>Robustness of Risk Maps and Surveillance Networks to the Knowledge Gaps about a New Invasive Threat</i> , Denys Yemshanov, Canadian Forestry	
	<i>Assessing risks with Agent-based Models (ABMs). Longitudinal projections and added uncertainty</i> , Georgiy Bobashev, RTI	
	Inv – 2 Statistical Analysis and Data Mining: Frontiers of Problem Data and Solution Methodology	Olympic
	Organizer and Session Chair: Arnold Goodman	
	Panel: Usama Fayyad, Jon Kettenring, Michael Leblanc, Roy Welsch	
	Inv – 3 Recent Computational Topics in Astrostatistics	Vashon
	Organizer and Session Chair: Rebecca Nugent	
	<i>Scaling the sky</i> , Andy Connolly, U. Washington	
	<i>A statistically rigorous approach to astronomical source detection</i> , David Friedenber, CMU	
	<i>Using local likelihoods to estimate gravitational lensing of the CMB</i> , Ethan Anderes, UC Davis	

Inv – 4 Perspectives on Climate Change **Adams**

Organizer: Yasmin H. Said, Session Chair: Edward J. Wegman



Testing the hypothesis of anthropogenic global warming: A continuing controversy, S. Fred Singer, Science & Environmental Policy Project 
Extracting information from large-scale computer model output, Mark Berliner, Ohio State 
Discussant: Edward J. Wegman, George Mason

12:00 noon – 1:30 pm **Lunch Break**

1:30 pm – 3:15 pm **Technical Sessions**




Inv – 5 Computational Social Science **St. Helens**

Organizer and Session Chair: Adrian Raftery

A case study in mixed membership modeling with a focus on model selection, Elena Eroshova, U Washington 
Approaches to inference from link-tracing network samples, Krista Gile, Oxford
Matrix models for relational and social network data, Peter Hoff, U Washington 




Inv – 6 Novel Methods: Symbolic Data, Layered Graphics Anatomy, and Census Layered Graphics **Olympic**

Organizer and Session Chair: Arnold Goodman

Symbolic Data, Lynne Billard, U Georgia 
Grammars of Graphics, Hadley Wickham, Rice 
OnTheMap Data and Tool, Jeremy Wu, Census Bureau 




Inv – 7 Robust Methods in Regression **Vashon**

Organizer and Session Chair: David Scott

Robust cross-validation using an L1 criterion, Dennis Cox, Rice 
Robust parametric classification and variable selection with minimum distance estimation, Eric Chi, Rice and Baylor College of Medicine 
Parametric methods for smooth quantile regression, Jonathan Lane, Rice 

Con – 1 Contributed Session 1: Modeling **Adams**

Session Chair: Yasmin H. Said




Recursive modeling using xstatR, E. James Harner and Jun Tan, West Virginia U 
Bayesian analysis for exponential random graph models using the double Metropolis-Hastings sampler, Ick Hoon Jin and Faming Liang, Texas A&M 
Rainbow plots, bagplots and boxplots for functional data, Han Lin Shang, Monash 

3:15 pm – 3:45 pm **Afternoon Break**

3:45 pm – 5:30 pm **Technical Sessions**


Inv – 8 Computational Statistics Methods for Social Issues **St. Helens**


Organizer and Session Chair: Yasmin H. Said


Designing statistical tools with agent-based models in mind, Ben Klemens, Census 
Predicting edges and vertices in a network, Walid Sharabati, Purdue 
Healthcare utilization among border Hispanic seniors with diabetes: Frequencies and related factors, Xiaohui “Sophie” Wang, U Texas, Pan American 

Inv – 9 Computationally Intensive Dynamic Models **Olympic**

Organizer: C. Shane Reese, Session Chair: Derek Bingham


Simplified trans-dimensional model jumping with MCMC for complicated models, Dave Campbell, Simon Fraser 

Emulating the nonlinear matter power spectrum for the universe, Earl Lawrence, Los Alamos National Laboratory 


Estimating time-varying parameters in ODEs, Jiguo Cao, Simon Fraser 


Inv – 10 New Developments in Statistical Data Integration **Vashon**

Organizer and Session Chair: Michael Schimek

Statistical integration of ranked lists by means of data streams, M. G. Schimek, Danube University Krems, Austria 

Space oriented rank aggregation, S. Lin, Ohio State 


Application of an integrative analysis using association networks of biological data, K. Kugler, UMIT, Austria 


Consensus finding, exponential models, and infinite rankings, Marina Meila, U Washington 

Inv – 11 Bayesian Data Mining and Clustering **Adams**

Organizer and Session Chair: Roy Welsch

Hiding from decision tree detectors, David Scott, Rice 

Multivariate Bayesian logistic regression for analysis of clinical trial safety issues, William DuMouchel, Phase Forward Lincoln Safety Group 

A graph-based method for estimating the cluster tree of a density, Werner Steutzle, U Washington 

7:00 pm – 10:30 pm **Conference Banquet**


Friday, June 18, 2010


7:30 am – 5:00 pm **Registration**


8:00 am – 9:45 am **Technical Sessions**

Inv – 12 Novel Methods: Internet Data and Targeted Marketing, Seriation and Pattern Discovery, Plus Collaboration and Social Networks

Organizer and Session Chair: Arnold Goodman **St. Helens**

Internet data and targeted marketing, Usama Fayyad, Open Insights 


Collaborative seriation among disciplines and people, Innar Liiv, Tallinn University of Technology, Estonia 


Suggesting enrichment of social network analysis beyond graphs, Arnold Goodman, Collaborative Data Solutions 

Inv – 13 Separating the Wheat from the Chaff - Feature Selection in High-Dimensional Regression **Olympic**

Organizer and Session Chair: Tim Hesterberg





Feature selection with constrained L1 regularization, Leming Qu, Boise State 

Bootstrap inference for network construction, Pei Wang, Fred Hutchinson Cancer Research Center 

Boosting for nonparametric high-dimensional models, Lifeng Wang, Michigan State 

Inv – 14 All at See: Snapshots of Modern Visualization Research




Organizer: Antony Unwin, Session Chair: Jay Emerson **Vashon**

An object-oriented approach in R for the visualization of functional actigraphy data, Jürgen Symanzik, Utah State 
Visualization and statistical modeling, Adi Wilhelm, Jacobs University Bremen 
Every plot must tell a story - even in R, Heike Hofmann, Iowa State 
iPlots eXtreme - next generation of interactive graphics for analysis of large data, Simon Urbanek, AT&T Labs 

Inv – 15 Policy Issues on Climate Change

Adams

Organizer: Yasmin H. Said, Session Chair: Edward J. Wegman

Global warming: Nexus of politics, economics and science, Jeff Kueter, The Marshall Institute 
Global warming--fact, fiction, and fraud, Don Easterbrook, Western Washington 
Climate change policy and the climategate scandal, Yasmin H. Said, George Mason 



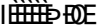
9:45 am – 10:15 am **Morning Break**

10:15 am – 12:00 noon **Technical Sessions**

Inv – 16 Computational and Statistical Issues in ABMs

St. Helens




Organizer and Session Chair: Georgiy Bobashev

Simplifying complex systems into multi-level agent based models, Rainer Hilscher, Altarum Institute 
The role of population heterogeneity and human mobility in the spread of pandemic influenza, Marco Ajelli, Fondazione Bruno Kessler, Italy 
Statistical issues posed by interaction-based models in social systems, Chris Barrett, VBI 

Inv – 17 Graphical Methods for Classification Based on Dimension

Reduction Organizer and Session Chair: Simon Sheather




Olympic

SMVCIR dimensionality test, Charles Lindsey, Texas A&M 
Robust dimensional reduction via invariant coordinate selection, David Tyler, Rutgers 
Sufficient dimension reduction based on normal and its connection with principal components, Liliana Forzani, Universidad Nacional del Litoral 

Inv – 18 Non-English Text Data Mining Via the Vector Space Model

Organizer and Session Chair: Jeff Solka



Vashon

Computing within the foreign vector space framework, Nick Tucey, NSWCD 
Wikipedia as a test bed for implicit translation, Kristin Ash, NSWCD 
An experiment in implicit translation, David Marchette, NSWCD 

Con – 2 Contributed Session 2: Inference

Adams




Session Chair: Rida Moustafa




Model-averaged L1-penalized logistic regression, Mark Seligman, Chris Fraley, Insilicos LLC 
Regression tree boosting to adjust health care cost predictions for diagnostic mix, John W. Robinson, Statistical and Health Informatics Consulting 


12:00 noon – 1:30 pm **Lunch Break**

1:30 pm – 3:15 pm **Technical Sessions**

Inv – 19 Statistical Analysis and Data Mining: Frontiers of Social Science and Network Science **St. Helens**
Organizer and Session Chair: Arnold Goodman
Panel: Pedro Domingos, Adrian Raftery, Lee Wilkinson




Inv – 20 Quantitative Horizon Scanning **Olympic**
Organizer and Session Chair: David Marchette
Detecting anomalies and estimating anomaly characteristics in time series of graphs, Glen Coppersmith, Johns Hopkins 
Multi-feature clustering and visualization of large document collections, Jeff Solka, Avory Bryant, NSWCD 
Quantitative horizon scanning for mitigating technological surprise, Avory Bryant, NSWCD 




Inv – 21 Computational Statistics and Robotics **Vashon**
Organizer and Session Chair: Barry Bodt
Issues and approaches in on-line modeling of environments from 3D data, Marshal Hebert, Carnegie Mellon 
Partial least squares applications in computer vision, Larry S. Davis, U Maryland and Aniruddha Kembhavi, U Maryland 
Test methods and metrology for evaluating human detection and tracking systems, Tsai Hong, NIST and Barry Bodt, ARL 

Work – 1 Advances in Machine Learning and Data Mining Technology Workshop **Adams**
Part 1: Advances in Interaction Detection using New Machine Learning and Data Mining Technology, Presented by Dan Steinberg, the President of Salford Systems 

3:15 pm – 3:45 pm **Afternoon Break**

3:45 pm – 5:30 pm **Technical Sessions**


Inv – 22 JCGS Highlights **St. Helens**
Organizer and Session Chair: David van Dyk
Combining mixture components for clustering, Adrian Raftery, U Washington 
Understanding GPU programming for statistical computation: Studies in massively parallel massive mixtures, Andrew Cron, Duke 
Pairwise display of high dimensional information via Eulerian tours and Hamiltonian decompositions, Wayne Oldford, U Waterloo 


Inv – 23 Computer Models, Virtual Laboratories **Olympic**
Organizer and Session Chair: Stephan Sain
Predictive modeling of a radiative shock physics, Derek Bingham, Simon Fraser 
Statistical analysis of regional climate model ensembles, Steve Sain, NCAR 
Posterior exploration for computationally intensive forward models, Shane Reese, BYU 


Ref – 1 Refereed Session

Vashon

Session Chair: E. James Harner


ChiD, A χ^2 -based discretization algorithm, Ross Bettinger 

A Bayesian decision theoretic approach to multiple hypotheses problems,
Naveen K. Bansal, Marquette and Klaus Miescke, U Illinois, Chicago 

Vulnerability of US hospitals from terrorist attack: Pilot study on hospitals in California, Byeonghwa Park and Yasmin Said, George Mason 

Work – 2 Advances in Machine Learning and Data Mining Technology Workshop

Adams

Part 2: Hands-on Introduction to New Machine Learning and Data Mining Technology: Jerome Friedman's TreeNet and Leo Breiman's Random Forests,
Presented by Dan Steinberg, the President of Salford Systems 

Saturday, June 19, 2010


8:00 am – 9:45 am


Technical Sessions

Inv – 24 Sampling and Inference for Hidden Networked Populations

Organizer and Session Chair: Mark Handcock

St. Helens

Respondent-driven sampling: Risks and benefits of a novel sampling strategy, W. Whipple Neely, U Washington 


Network model-assisted inference from respondent-driven sampling data, Krista J. Gile, Oxford 


Developments in network sampling, Steven K. Thompson, Simon Fraser 

Inv – 25 Interfacing Text Mining and Image Analysis

Olympic

Organizer and Session Chair: Adi Wilhelm

Combining text and image processing in an automatic image annotation system,
Iulian Ilies, Jacobs University, Bremen 

Extraction of endogenous metadata for text and image databases, Edward Wegman, George Mason 


Keeping your own familiar file organization, Dong Cao, CMU 

Inv – 26 Computing on Streams

Vashon

Organizer and Session Chair: William Szewczyk

Stream algorithms and workflows, J. David Harris, DoD 

Analytics for streaming applications, Daby Sow, IBM/T. J. Watson Research 

Index Learning Omid Madani, SRI 

9:45 am – 10:05 am

Morning Break


10:05 am – 11:50 am


Technical Sessions


Inv – 27 Model Selection Problems in Kernel-Based Methods with Applications

Organizer and Session Chair: Hamparsum Bozdogan

St. Helens

Robust Bayesian relevance vector machines using information complexity and the genetic algorithm, Brant Quinton, U Tennessee 

A non-inferiority trial design without the need for a conventional margin, Xi Chen, PharmClint Co 

Hybridized support vector machine and recursive feature elimination with information complexity, Seung Hyun Baek, U. Tennessee 

Inv – 28 Visualizing Intrusion Detection Data

Olympic

Organizer: Rida Moustafa, Session Chair: Yasmin H. Said

On some visualization methods for computer networking data, Rida E.A.

Moustafa, dMining Technology, LLC

Visualizing streaming data, John W. Emerson, Yale

Change detection in multivariate streaming data, Kyle A. Caudle, U.S. Naval Academy

Inv – 29 Learning from Proximity Data

Vashon

Organizer and Session Chair: Michael Trosset

Local learning methods for proximity data, Maya Gupta, U

Washington

Interactive concept search, Ashish Kapoor, Microsoft Research

Learning from heterogeneous data sources by combining dissimilarities, Brent Castle, Indiana