



# Massive Datasets Transition Workshop

## May 20-22, 2013

### SCHEDULE

**Monday, May 20**  
**Radisson Room H**

- |              |  |
|--------------|--|
| 8:15-8:45    | Registration and Continental Breakfast   |
| 8:45-9:00    | Welcome + Opening Remarks<br><br>{Environment & Climate}   |
| 9:00-9:30    | <b>Dorit Hammerling</b> (SAMSI)<br><i>"A Bayesian Hierarchical Model for Climate Change Detection and Attribution"</i>   |
| 9:30-10:00   | <b>Matthias Katzfuss</b> (Heidelberg, Germany)<br><i>"Statistical Inference for Massive Distributed Spatial Data Using Low-Rank Models"</i>  |
| 10:00-10:30  | <b>Peter Thorne</b> (CICS-NC)<br><i>"The International Surface Temperature Initiative: Opportunities to engage in creating land surface air temperature datasets for the 21st Century"</i> |
| 10:30-11:00  | Break<br><br>{Online Streaming & Sketching}  |
| 11:00-11:30  | <b>Michael Mahoney</b> (Stanford)<br><i>"Revisiting the Nystrom Method for Improved Large-Scale Machine Learning"</i>  |
| 11:30-12:00  | <b>David Lawlor</b> (SAMSI)<br><i>"Regression in High Dimensions via Geometric Multi-Resolution Analysis"</i>  |
| 12:00 -12:30 | <b>Garvesh Raskutti</b> (SAMSI)<br><i>"Leverage-Score Sampling for Large-Scale Linear Regression"</i>  |
| 12:30-2:00   | Lunch<br><br>{High Energy Physics}   |
| 2:00-2:15    | <b>Steffen Bass</b> (Duke)<br><i>"Do You Speak Klingon? - The Challenges of Creating a Successful Collaboration Between Physicists and Statisticians"</i>                                  |

- 2:15-2:45 **Peter Marcy** (Wyoming)  
*"Towards Calibration in High-Energy Physical Models Using ATLAS Histogram Data"*
- 2:45-3:15 **Chris Coleman-Smith** (Duke)  
*"A Physicists Take on the Design and Analysis of Computer Experiments"*
- 3:15-3:30 **Karen Kafadar** (Indiana)  
*"Future Directions for Statistics in High-Energy Physics"*
- 3:30-5:30 Meeting of Individual WGs (discuss transition and prepare final report)
- 5:30 Poster Reception  
*SAMSI will provide poster presentation boards and tape. The board dimensions are 4 ft. wide by 3 ft. high. They are tri-fold with each side being 1 ft. wide and the center 2 ft. wide. Please make sure your poster fits the board. The boards can accommodate up to 16 pages of paper measuring 8.5 inches by 11 inches.*

**Tuesday, 21 May**  
**Radisson Room H**

- 8:30-9:00 Registration and Continental Breakfast  
  
{Inference}
- 9:00-9:30 **Xia Wang** (University of Cincinnati)  
*"Bayesian Large-Scale Multiple Testing for Time Series Data"*
- 9:30-10:00 **Tao Yu** (National University of Singapore)  
*"Local Polynomial Estimation of the Semi-Nonparametric Models: Joint Asymptotic Studies"*
- 10:00 **Yuefeng Liu** (UNC)  
*"Large-Margin Classifier Selection via Decision Boundary Stability"*
- 10:30-11:00 Break  
  
{Datamining & Clustering}
- 11:00-11:30 **Ralph Abbey** (NCSU)  
*"Stochastic Data Clustering"*
- 11:30-12:00 **Shaina Race** (NCSU)  
*"Iterative Consensus Clustering"*
- 12:00 - 2:00 Lunch

{Astrostatistics}

- 2:00-2:30 **Tamas Budavari** (Johns Hopkins)  
*"Big-Data Inference on GPUs"*
- 2:30-3:00 **Robert Wolpert** (Duke)  
*"LARK Models for Light Curves"*
- 3:00-3:30 **Mary Beth Broadbent** (Duke)  
*"Levy Adaptive Regression Kernels with Applications for Gamma-Ray Burst Lightcurves"*
- 3:30-4:00 Break
- {Discovery & Classification in Synoptic Surveys}
- 4:00-4:30 **Fabrizia Guglielmetti** (Max Planck Institut, Germany)  
*"Challenges and Possible Solutions in Image Analysis"*
- 4:30-5:00 **G. Jogesh Babu** (Penn State)  
*"Exploratory Analysis of Light Curves"*
- 5:00 Meeting of individual WGs (discuss transition and prepare final report)

**Wednesday, 22 May**  
***Radisson Room H***

- 8:30-9:00 Registration and Continental Breakfast
- {Multiscale Modeling}
- 9:00-9:30 **David Dunson** (Duke)  
*"Multiresolution Dictionary Learning for Conditional Distributions"*
- 9:30-10:00 **Mauro Maggioni** (Duke)  
*"Multiscale Estimation of Probability Measures in High Dimensions"*
- 10:00-10:30 **Marco Ferreira** (Missouri)  
*"Dynamic Multiscale Spatiotemporal Models for Poisson Data"*
- 10:30-11:00 Break

{Imaging}

- 11:00-11:30      **Ashish Mahabal** (Caltech) + **Lingsong Zhang** (Purdue)  
                          *“Analyzing Light Curves of Astronomical Sources”*
- 11:30-12:00      **Weihong Guo** (Case Western) + **Dan Yang** (SAMSI)  
                          *“Compressive Inference”*
- 12:00-12:30      **Dani Ushizima** (Lawrence Berkeley National Lab)  
                          *“Image Analysis and Scientific Computing”*
- 12:30-2:00        Lunch
- 2:00-2:30         **Ilse Ipsen** (SAMSI and NCSU) **Richard Smith** (SAMSI):  
                          Where do we go from here?
- 2:30-5:00         Meeting of Individual WGs (discuss transition and prepare final report)
- 5:00                Good Bye Party