

NISS Parameters

A quarterly newsletter from the National Institute of Statistical Sciences



NISS at Twenty: A Look Back at Our Past

Throughout this year, as NISS celebrates its twentieth anniversary and the affiliates program celebrates its tenth anniversary, *NISS Parameters* will look back and feature some of the highlights that occurred to help create and foster this unique institution.

In the 1980's a group of statisticians began to talk about the need for statistics to pursue new opportunities for cross-disciplinary collaboration. Following

discussions at a meeting of the Institute of Mathematical Statistics (IMS) at Lake Tahoe in 1984, a proposal was sent to the National Science Foundation

(NSF) for a panel study to be conducted about cross-disciplinary research. The resulting report, co-authored by panel chairs Ingram Olkin and Jerome Sacks, was published in September 1988. Among its recommendations was



Albert Bowker was on the selection committee that chose RTP as the location for NISS.



formation of an institute dedicated to cross-disciplinary research involving the statistical sciences.

Spurred by the report, NSF program directors Nancy Flournoy (statistics) and Murray Aborn (social sciences) funded a feasibility study conducted by the American Statistical Association (ASA). This study led in turn to a prospectus, a call for proposals and a selection committee composed of Albert Bowker, Janet Norwood and Olkin.

Of several strong proposals, the one to locate the institute in Research Triangle Park (RTP), North Carolina stood out. It came from a consortium that includes three universities with some of the longest histories in statistics—Duke, North Carolina State and the University of North Carolina at Chapel Hill, and reflected the

Triangle's being second only to Washington in terms of the number of statisticians. There were faculty positions for the first director and associate director, as well as faculty and student release time to



Daniel Horvitz became the interim Executive Director for NISS.

participate in NISS projects. The Research Triangle Foundation (RTF) offered a six-year startup grant, and the Triangle Universities Center for Advanced Studies, Inc. (TUCASI) provided a 99-year lease on a 5-acre building site for one dollar per year. The Research Triangle Institute (now RTI International) provided space and computer support at no cost to
(Continued on page 7)

A Note from the Director

We are starting this anniversary year in the most exciting way imaginable, with the two new NISS employees whom you can read about below in this issue of *Parameters*. We now stand at seventeen employees, an all-time



record that I fully expect to eclipse by the fall of this year!

This increase in scale, of course, reflects with continuing

growth and diversification of our research activities. Today, in broad terms our research intersects all of the major issues of our times—health, education, the environment, our physical and information infrastructure, and the security of the US. NISS is where the action is!

Of course, the more we grow, the more opportunities we see for further growth. But there are challenges to be overcome, including the current “nearly maxed out” state of our scientific leadership and the difficulty in developing major initiatives given our current revenue streams.

In response, and stimulated by the 2007 strategic positioning study, we are giving serious consideration to philanthropic-style fundraising. The goal would be to generate the resources to enable NISS to take the lead in assembling the world’s leading statisticians, in collaboration with other scientists and policy makers, to bring the most advanced statistical thinking to bear on today’s great challenges. We are developing a case statement, and are about to undertake a second round of interviews that give us others’ reaction to it. Where this will all lead, and when, is not clear, but I, the Board of Trustees,

and especially the Executive Committee—which is playing a major role in the effort—are convinced that we are taking steps in the right direction.

As a concrete first step, we have established the NISS Fund, to which contributions in any amount are welcome (and fully tax-deductible!). For more information, and to donate, please go to <http://www.niss.org/about/niss-fund>. Thank you in advance!

Alan Karr
Director

NISS Welcomes Yang and Sang

NISS is pleased to announce that it has hired David Yang as research associate and Hailin Sang as a postdoctoral fellow. Both David and Hailin will begin working at NISS in January.

David will be the second NISS member (in addition to research analyst Weiwei Cui) of the statistical standards project team at ESSI/NCES in Washington DC. He will be responsible for providing methodological and statistical support for technical review activities. He will take a lead role in projects that monitor

the quality of NCES data and track the use of NCES data products by practitioners, researchers and policy makers. David will also conduct data analyses and write research and technical reports.

David holds a masters degree in statistics from the University of Kentucky, and among his other achievements is record-breaking speed in obtaining SAS certification.

Hailin is joining the NISS-NASS project as soon as visa arrangements are completed. His 2008 Ph.D. is in statistics from

the University of Connecticut, so he is rejoining his former student colleagues Xia Wang and Frank Zou, who are also postdocs at NISS. Hailin will work on the estimation of small farms from Census mail list research project.

Hailin was a visiting assistant professor at the University of Cincinnati since the fall of 2008. Like his project colleague postdocs Michael Robbins and Jay Wang, Hailin will be based at NASS in the Washington DC area during the academic year, and will spend next summer at NISS.

NISS

Celebrating 20th year of NISS

2010

10th year of affiliates program

Photos From: NISS-OCC Explorations Workshop on Financial Risk Analysis



Alan Karr, Director of NISS, welcomes the participants to the NISS and OCC Exploration Workshop II on Financial Risk Modeling.



The audience listens attentively to the speaker at the NISS and OCC Explorations Workshop which was held in Washington DC in October.



Michael Sullivan, Director, Market Risk Analysis Division of OCC chats with Richard Davis.



Rick Bookstaber, formerly Manager of Quantitative Equity Fund at FrontPoint Partners, Risk Manager at Moore Capital Management, and Managing Director for firm-wide risk management at Salomon Brothers, speaks to the audience.

Postdoc Profile:

Xia Wang

NISS is fortunate to have several new postdocs this year. One of the postdocs who is working on the Clinical Proteomic Technology Assessment Collaboration is Xia Wang.

Xia was born in Hubei, China. For several years her family lived in the northeastern province of China called Heilongjian, but moved back to Hubei when Xia was in high school.

She studied at the Huazhong University of Science and Technology and majored in economics, receiving her bachelor's degree in 1996. From 1996-98, she worked for a municipal investment

"If the experiment is designed well, the scientists can glean valuable information. Otherwise, the data is useless," - Xia Wang

company that looked at long-term infrastructure needs. She was doing financial analysis work for them. In 1998, she went to graduate school at the Zhongnan University of Finance and Economics. In 2001, she came to the United States and studied at the University of Connecticut. She was still in the economics department, pursuing her Ph.D. She started to take a few statistics courses for her research in economics and enjoyed the classes very much.

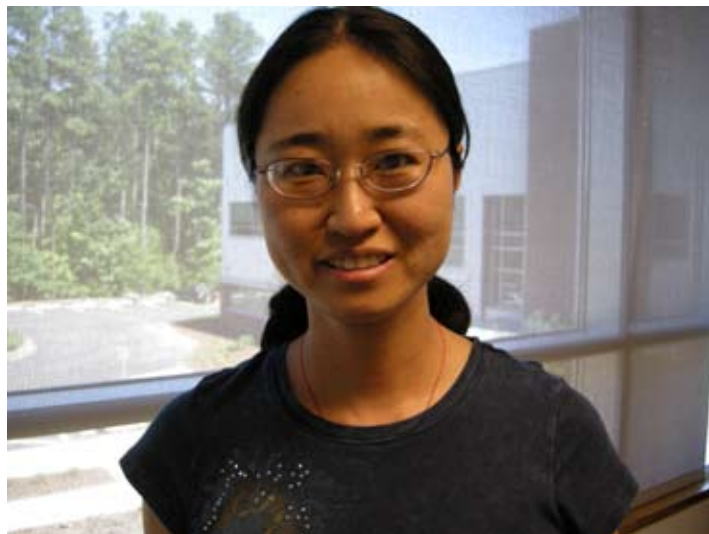
"I feel there are many methodologies that you can learn

in statistics that apply to many different fields. I was very excited about learning so many new things," said Wang.

She was collaborating with Professor Richard Langlois at the Economics Department.

He helped her connect with Professor Will Mitchell at Duke University on a project she was working on that looked at biotech companies and how the companies form alliances with each other. However, it was very difficult to get data from the companies to work on this project. The folks at Duke had a data set they shared with Xia for her research project thanks to an introduction from Dr. Langlois. He was one of the people who supported her to apply to the statistics department.

She was admitted to the Statistics Ph.D. program in Spring 2006, one semester before she successfully defended her Ph.D. in Economics. In the Statistics Department, she worked with Dr. Dipak K. Dey on a family of flexible link functions based on the generalized extreme value distributions. They have applied their model to binary, ordinal and spatial correlated data. As a part of her graduate training, she was also



Xia Wang is working on the CPTAC project at NISS

involved in consulting work on behalf of the statistics department under the mentor of Dr. Ming-Hui Chen. She worked on projects for the communications science, marine science and plant science departments to name a few. In fact, a co-authored paper is accepted in the *Applied Turfgrass Science*, which is a project dealing with a survey of golfers that she and Dr. Chen have worked on as part of the consulting work. She graduated with her Ph.D. in Statistics in 2009.

Xia joined NISS in June 2009. She is working on the Clinical Proteomic Technology Assessment Collaboration (CPTAC) program, along with postdocs Jessie Xia and Xingdong Feng, associate director Nell Sedransk and assistant director of bioinformatics, Stan Young. Xia is involved with the discovery working group of the CPTAC program. They are looking for the biomarkers that will help
(continued on page 6)

Affiliate Profile:

Yahoo! Labs

The world is getting more data rich every year. Portals like Yahoo! have millions of visitors each day. Yahoo! Labs manages and studies many of the largest and richest data repositories in the world. Its scientists focus on data-driven analysis, high-quality search, algorithms and economic models. It is a perfect place for a statistician to work and to flourish. Deepak Agarwal, the Yahoo! Labs representative for the affiliates program, notes, “The web critically depends on

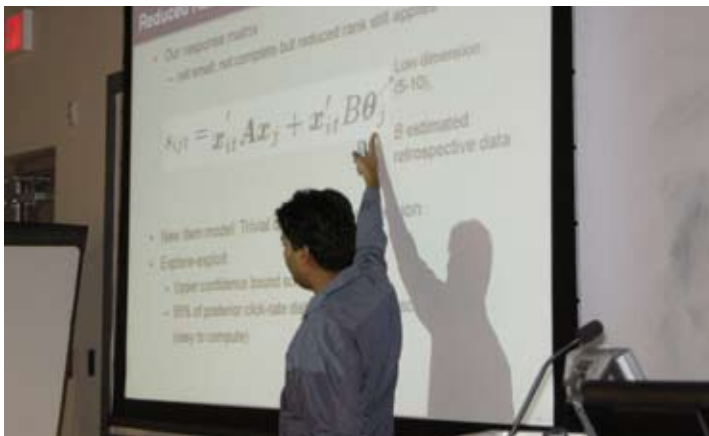
and search results are delivered to a particular user when visiting a website. A primary source of revenue for Yahoo! is through advertising. Online advertising is already a multi-billion dollar industry and continues to grow at a rapid rate. Not surprisingly, it is practiced in various forms to cater to diverse advertiser objectives. For instance, advertisers often opt to show banner ads displayed to users visiting a website on the Yahoo! portal to build “brand awareness” that promotes future

sales. Controlled experimental designs conducted at Yahoo! have clearly shown the positive impact of such advertisements on sales. Other advertisers with rare repeat sales (e.g. auto insurance,

specials, or sales on bikinis. Some ads are not as obvious as this example though. “Advertisers are happy only if people who see their ads are ones who like it,” notes Agarwal, “We have huge amounts of data coming through our system every day with several million visits. The goal is to process this data, do statistical modeling, and then figure out the best advertisements to show this user.” In order to guarantee good return on investment to advertisers, it is also important to serve content that maximizes user engagement, reach and satisfaction. This provides an unique opportunity to develop statistical methods to recommend the best content to users visiting various websites on Yahoo!. Statistics is a key technology that have had a profound impact on many products at Yahoo!; the importance of statistics is well recognized at all levels of the company and Yahoo! is dedicated to building a strong statistical presence in the company. Getting involved in the affiliates program was a natural fit to help them with fulfilling this mission.

Recently, NISS held a two-day workshop on computational advertising. The workshop introduced the general topic to many statisticians who have not been working in this type of statistics in the past. “It was a small group by design. There was a lot of

(Continued on page 7)



Deepak Agarwal presents his work at the Affiliates Computational Advertising workshop in November.

statistics, machine learning and computer science. One of our big problems to solve is what ads to show to users.” Agarwal explains that there is a tremendous amount of data collected from Yahoo! servers each day. Several terabytes are collected and the department has to mine the data to figure out people’s habits and behaviors so that the right kinds of ads, content

refrigerators) may care more about immediate than future sales. Such advertisers may opt for performance-based advertising where the goal is to maximize clicks on ads and promote immediate sales. So, for example, you might be searching for articles on beach resorts and on the side bars advertisements will start to appear for cruises, vacation

Xia Wang Profile Continued

(From page 4)

doctors treat cancer patients at an early stage, or perhaps to identify the susceptibility before a person gets the disease at all. NISS is helping with the input of the design of the experiment and subsequently analyzing the data with sound statistical methodologies.

“If the experiment is designed well, the scientists can glean valuable information. Otherwise, the data is useless,” notes Xia.

She is also sitting on the Tumor Proteomics Discovery Group which is just forming. This group will compare Discovery proteomics platforms in the ability to generate inventories of tumor tissue proteomes and to detect proteomic differences that correspond to phenotypic differences.

“I love being here at NISS. I am getting really good direction from my mentor, Nell Sedransk. She gives me clear direction on where to go when I am trying to figure out a problem. She is really good at picking up the details and showing me the big picture. I have learned a

lot from Nell,” remarks Xia.

Xia has enjoyed attending some of the workshops that SAMSI sponsored. She attended two of the spatial program workshops and the Sequential Monte Carlo transition workshop.

Xia explains, “It gives me a chance to meet people from all over the world and be exposed to the most current and frontier research. I really don’t know of another place where you have that many different people all in one place.”



Xia Wang (R) and Jessie Xia (L) at the NISS Turkey Fry. Both are NISS postdoctoral fellows.

Computational Advertising Workshop



Charles Elkan, Jim Landwehr and David Banks chat at a break at the workshop.



The audience listens intently to the speaker.



Bob Bell, AT&T Labs Research, talks about how his team won the \$1 million Netflix prize.

Calendar of Events

Frontiers of Statistical Decision Making and Bayesian Analysis in Honor of James O. Berger

March 17 - March 20, 2010

Downtown Campus of the University of Texas - San Antonio

NISS Co-Sponsored and ARA-eligible

Conference on Resampling Methods and High Dimensional Data

March 25-26, 2010

Hilton Hotel and Conference Center, College Station, TX

NISS Co-Sponsored and ARA-eligible

Affiliates Annual Meeting

April 2010 (exact dates TBD)

Chicago, IL

ARA eligible

ITSEW Workshop

The Ongoing Evolution of Survey Methodology and the Impact on Total Survey Error

June 13-16, 2010

Stowe, VT

NISS Co-Sponsored and ARA eligible

History of NISS Continued

(from page 1)

NISS. Finally, the State of North Carolina promised funds to meet one-half of the cost of a building for the institute. The proposal also laid out a plan for self-sufficiency of the institute.

This proposal was selected. Daniel Horvitz of RTI, the centralizing force in pulling together the RTP proposal, became interim Executive Director. Bowker was elected the first chair of the Board of Trustees. The founding ceremony took place on December 3, 1990 on the RTP campus of Glaxo (now GlaxoSmithKline). Sherwood Smith, who was then vice-chair of RTF and chair, president and CEO of Carolina Power and Light (now Progress Energy), as well as chair of the

board of trustees of TUCASI emceed the event. North Carolina Governor James Martin was the keynote speaker. Although a chemist, he told jokes about statisticians that were both on point and mathematically correct!

So at the end of 1990, the National Institute of Statistical Sciences was a reality. More to come in the next issue of *NISS Parameters*.



Yahoo! Continued

(from page 5)

energy and there were lots of great discussions that resulted from the workshop,” comments Agarwal. It is quite possible that more workshops of this nature will be held in the future as computational advertising is a budding industry and there is still plenty of room for an enterprising statistician to make new strides in research in this area.

Yahoo! Labs is continuing to grow in leaps and bounds as the Internet continues to develop and change. “It’s an exciting place to work, especially for statisticians. The main business model critically depends on statisticians. Statisticians at Yahoo! Labs are actually entrepreneurs who can start their own projects and make an impact. Statisticians can potentially make new products, not just assist others in making products,” explains Agarwal. There are many exciting topics to explore in addition to computational advertising: recommender systems, social networks and more. The data are very rich and massive, but at the same time noisy and sparse, providing brand new opportunities for a lot of research.

What is so exciting to Agarwal about Yahoo! Labs is that he is able to pursue a research agenda that leads to state-of-the-art methodological research, makes a tremendous impact to the company and improves the lives of millions of users who use Yahoo! each day. The research environment is highly cross-disciplinary, with statisticians working in close collaboration with machine learners, computer scientists, economists, engineers and product managers. The dilemma to a researcher is not “what to work on” but “what not to work on.”

Updates from Former Postdocs

Murali Huran says, “I am still an assistant professor at Penn State Statistics. My collaborative projects involving geoscientists, infectious disease modelers and ecologists (and several bright statistics grad students), have been gathering momentum this past year. I graduated my first Ph.D. student in the spring of 2009, and am expecting to graduate another this summer. I am having a lot of fun with my cross-disciplinary projects while simultaneously working on research related to parallelizing, blocking and slicing Markov chain Monte Carlo algorithms. Being at SAMSI this fall semester for the space-time program has been wonderful, partly because of all the interesting ideas and projects I’ve been exposed to, and partly because it’s been nice to

be around all the familiar faces at NISS, including Katherine, James, Martha, Jamie and Alan. I was sorry to leave in December!”

Matthias Schonlau was awarded a fellowship from the Max Planck Institute for Human Development in Berlin, Germany. “The fellowship enables me to spend a sabbatical located at the German Institute for Economic Research (DIW), which hosts the longest running German household panel survey, the G-SOEP. I will return to RAND (Pittsburgh) in July 2010,” says Schonlau.

P.O. Box 14006
19 T.W. Alexander Drive
Research Triangle Park, NC 27709
919.685.9300 (phone) 919.685.9310 (fax)
www.niss.org

NISS/SAMSI Affiliates

Industry

AT&T Labs-Research, Florham Park, NJ
Avaya Labs, Basking Ridge, NJ
Bayer HealthCare Pharmaceuticals, West Haven, CT
GlaxoSmithKline, Research Triangle Park, NC and Collegeville, PA
Merck Research Laboratories, West Point, PA
MetaMetrics, Inc., Durham, NC
PNYLAB, LLC, Princeton, NJ
RTI International, Research Triangle Park, NC
SAS Institute, Cary, NC
SPSS, Chicago, IL
Telecordia Technologies, Piscataway, NJ
Yahoo! Research Laboratory, Silicon Valley, CA

Government Agencies & National Laboratories

Bureau of Labor Statistics, Washington, DC
US Census Bureau, Washington, DC
Energy Information Administration, Washington, DC
National Agricultural Statistics Service, Fairfax, VA
National Center for Education Statistics, Washington, DC
National Center for Health Statistics, Hyattsville, MD
National Security Agency, Ft. George W. Meade, MD
Office of the Comptroller of the Currency (Treasury Department), Washington, DC

University

University of California - Berkeley, Department of Statistics
Carnegie Mellon University, Department of Statistics
Columbia University, Department of Biostatistics
University of Connecticut, Department of Statistics
Duke University, Departments of Statistical Science and Mathematics
Duke University Medical Center; Department of Biostatistics and Informatics
Emory University, Department of Biostatistics
University of Florida, Department of Statistics
Florida State University, Department of Statistics
George Mason University, Department of Statistics
Georgetown University Medical Center, Department of Biostatistics, Bioinformatics, and Biomathematics
University of Georgia, Department of Statistics
University of Illinois Urbana-Champaign, Department of Statistics
Indiana University, Department of Statistics
University of Iowa, Department of Statistics
Iowa State University, Department of Statistics
Johns Hopkins University, Department of Applied Mathematics and Statistics

Medical University of South Carolina, Department of Biostatistics, Bioinformatics and Epidemiology
University of Michigan, Departments of Statistics and Biostatistics
University of Missouri-Columbia, Department of Statistics
North Carolina State University, Department of Statistics
North Carolina State University, Department of Mathematics
University of North Carolina at Chapel Hill, Department of Statistics and Operations Research
University of North Carolina at Chapel Hill, Department of Biostatistics
University of North Carolina at Chapel Hill, Department of Mathematics
Oakland University, Department of Mathematics and Statistics
Ohio State University, Department of Statistics
Pennsylvania State University, Department of Statistics
Purdue University, Department of Statistics
Rice University, Department of Statistics
Rutgers University, Department of Statistics
University of South Carolina, Department of Statistics
Southern Methodist University, Department of Statistical Science
Stanford University, Department of Statistics
Texas A&M University, Department of Statistics
Virginia Commonwealth University, Departments of Biostatistics and Statistical Sciences