

# Newsletter

College of Physical and Mathematical Sciences

November 2008

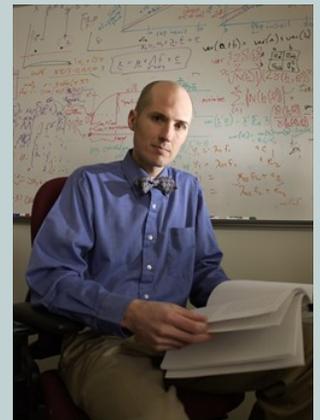


## In this issue:

1	BYU Statisticians Simulate "If the election were held today"
2	Important Dates
	Rodney Forcade receives Award and gives lecture
	Dr. William Christensen is Guest Lecturer
3	Computer Science Students take 3rd Place in Washington, DC Competition
	Grant Received in Computer Science
4	Utah Colleges Exit Poll
5	College Publications

## BYU Statisticians Simulate "If the election were held today"

Using a new method for assimilating polling information and simulating elections, BYU statisticians monitored election polls during the fall of 2008 in order to predict the outcome of the election. On Election Day, the method forecasted that Democratic presidential candidate Barack Obama would beat Republican John McCain 353 to 185 in the Electoral College, and Senate Democrats would most likely come away with 58 seats. Although some of the election results are still unclear, it appears that the approach correctly predicted 50 of 51 "states" in the Presidential election, only slightly under predicting the probable final electoral count of 365 for Obama. In the Senate election, the statisticians' approach appears to have correctly predicted all the winners with the exception of Alaska, where incumbent Ted Stevens has a slight lead with absentee and early votes still uncounted. Election night results in Alaska and Minnesota are still tentative.



Emphasis on "if the election were held today," says Brigham Young University statistician William Christensen, who notes that about a month before the elections, the simulations indicated a tie in the Electoral College was a credible scenario.

Christensen and BYU student Alan Vaughn run 50,000 simulations daily based on state-by-state polling and post the results online. The academic journal *The American Statistician* recently published the BYU team's method of simulating election outcomes.

"This approach is on solid scientific footing and provides a really interesting picture of how the probabilities of winning in each state add up," said Peter Westfall, a statistician at Texas Tech University and editor of the journal, a publication of the American Statistical Association. Westfall was not involved with developing the simulation method.

As of Election Day, Missouri and North Carolina appeared to be the most closely contested states.

To do the simulations, the researchers flip 51 virtual weighted coins – one for each state plus the District of Columbia – 50,000 times. Each figurative coin reflects the odds of that state going for McCain or Obama according to that state's combined poll results. The more recent the poll, the more influential it is in the simulation.

"This helps cut through the news chatter about national polls and directly states the candidates' odds of winning," Christensen said. "Along the way, the students in my class have fun getting their hands dirty using statistics on a topic they're already interested in."

One observation Christensen and his students made is the strong connection between Obama and Senate Democratic candidates in the South. The Senate contests in Georgia, North Carolina and Kentucky see the highest correlation between how Obama and the Democratic candidate fare in state-level polling, suggesting that they were riding the same wave. Republicans currently hold all three of those seats.

Predictions and analyses are found at web page: <http://statistics.byu.edu/faculty/wfc/2008Election/>

—This article adapted from the BYU press release by Joe Hadfield.

## Important Dates & Events in the College

November 2008

National Security  
Nephi Noble  
December 4th, 2008  
1170 TMCB  
4:00 pm

President Samuelson's Visit  
December 11th, 2008  
377 CB  
2:00pm-3:15pm

College Christmas Social  
December 12th, 2008  
ESC Pendulum Court  
11:30am-1:00pm

## Rodney Forcade Receives BYU Mathematics Award, Gives Lecture

Brigham Young University professor Rodney Forcade received the Fourth Annual Distinguished Teaching Award from the Department of Mathematics on November 13th.

Forcade then presented a lecture, "Fun with Continued Fractions and Cryptography."

"We'll have a little fun with some elementary mathematics, which just happen to apply to modern cryptography," Forcade said concerning the lecture.

In consideration of students who struggle with math, Forcade developed Mathematics 102, a general education course that teaches real-life mathematics and quantitative reasoning, such as personal finance and consumer statistics.

"Rodney succeeded in creating a course that can catch the students' interest and instill in them an appreciation for the usefulness of mathematics in their own lives," said Darrin Doud, associate professor of mathematics.

Forcade's students are very appreciative of the personal, individualized attention he gives, said David Wright, professor of mathematics. A student in one of Forcade's calculus courses said Forcade takes each class on an "in-class field trip" to his office so they know not only where to visit him, but also that they are welcome to do so.

"Professor Forcade would be the last person who would aspire for recognition of his teaching," said Wright. "He goes about his work in a very quiet way, yet he exemplifies outstanding teaching."

For more information, contact David Wright at (801) 422-4027 or [wright@math.byu.edu](mailto:wright@math.byu.edu).

—BYU News



## Dr. William Christensen is Guest Lecturer

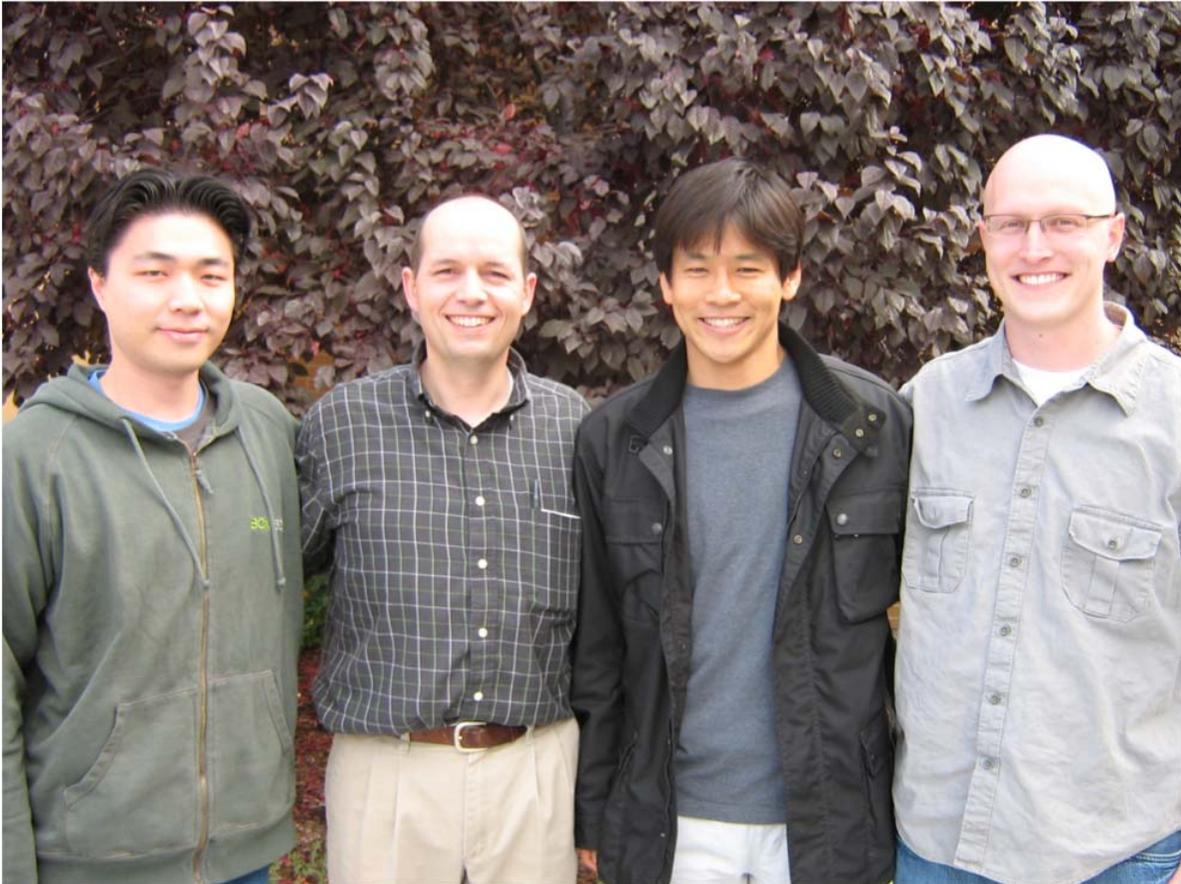
Dr. William F. Christensen, Associate Professor of the Department of Statistics, was the featured speaker for the Fourth John C. & Susan S. G. Wierman Lecture in Air Quality Data Analysis. The annual lecture was held on October 30, 2008 at The Johns Hopkins University in Baltimore, Maryland. This lecture series, hosted by the Johns Hopkins Department of Applied Mathematics and Statistics, features talks on developments in air quality data analysis that are relevant for policy development. It seeks to bring together faculty and researchers in engineering and sciences with state and local air quality officials to enhance understanding and stimulate collaboration on important air quality issues. The lectures are intended to showcase new developments, to encourage the quantitative analysis of scientific issues related to air quality, and to elucidate the policy implications of recent research.

The title of Dr. Christensen's presentation was "Identifying Pollution Source Locations for Air Quality Monitoring." The lecture discussed the pollution source apportionment (PSA) problem, which involves quantifying the impact of major sources of pollution on air quality. The identification of pollution source directions is an important part of PSA. Estimated source directions are used both as inputs to a Bayesian source apportionment analysis, and as part of a post-analysis check to associate identified pollution factors with potential pollution sources. The lecture discussed two approaches for source location identification which can be used in different settings. The first requires wind direction data measured at the air quality receptor and makes use of statistical and/or deterministic (AERMOD) models for chemical transport of particulate matter from source to receptor. The second makes use of HYSPLIT back-trajectory estimates and a kriging estimator which filters heterogeneous measurement errors.

# Computer Science Students take 3rd place in Washington, DC International Competition

Jun won Lee, a Masters student in the Computer Science Department, recently returned from Washington, DC, where he presented a paper at the Annual Symposium of the American Medical Informatics Association (AMIA). The AMIA Annual Symposium is a peer-reviewed conference and the largest and most comprehensive biomedical and health informatics symposium held anywhere in the world. Jun and two of his fellow students in Dr. Giraud-Carrier's Data Mining Lab, Yao Huang Lin (an undergraduate student) and Matt Smith (a PhD student) submitted a paper in the student competition, the focus of which was the study of best practices related to knowledge discovery in health care data. Out of thousands of papers submitted from around the world, the Data Mining Lab's paper was one of just a handful to be invited to present at the Symposium. Jun's presentation at the symposium was well received; he ended up taking third place in the overall competition.

In their research for the paper, Jun, Yao, and Matt used data mining tools to derive useful information from the data and statistical information collected by the National Center for Health Statistics. The students performed two experiments using decision tree and association rule modeling to find relationships in the data that may prove valuable to medical researchers. They found several implicit relations among the data, showing, for example, a correlation between diabetes and wheezing in the chest, and high blood pressure and hearing disabilities. Jun, Yao, and Matt are excited and gratified by the recognition given to them by the AMIA; however, even more significant to them is the fact that the discoveries they made over the course of their research provide interesting insights that may open additional doors in the study of major diseases and risk factors for these diseases.



From Left to right: Yao Huang Lin, Dr. Christophe Giraud-Carrier, Jun won Lee, and Matthew Smith

## Grant received in Computer Science Department

Dr. Mike Goodrich of the BYU Computer Science Department, along with Mark Colton (Mechanical Engineering), Bonnie Brinton (Graduate Studies), and Martin Fujiki (Communication Disorders), were recently awarded a Honda Initiation Grant to work on assistive robotics.

# Utah Colleges Exit Poll

In order to know the winner of an election with 100% certainty, all the votes must be tallied even if that requires hours, days or months. However, an unbiased probability sample of only a relatively small number of voters may get an estimate with sufficient precision to identify a winner.

For example, a sample of 3800 of the almost 880,000 Utah voters who went to the polls this year produced a maximum margin of error for statewide contests of 2.6%. That means that a winner can safely be inferred in races where the difference between the two candidates is greater than 5.2%. When a race is closer than that we label the race 'too close to call' and wait for the returns like everyone else. For congressional district races involving U.S. Representatives, the maximum margin of error was 4.6% or less based on sample sizes of approximately 1100 respondents and would require differences in excel of 9.2% in order to "call a race."

One of the most effective ways to learn is by doing, in a real-world setting. The Utah Colleges Exit Poll (UCEP) provides such a learning experience for students of statistics and political science from BYU and the other participating schools: USU, Westminster College, Weber State, UVU, SUU and Dixie College. The first Utah Colleges Exit Poll was conducted in 1982 and has been repeated every 2 years since then. It has been described as the most accurate exit poll in the state, and a prominent visitor on campus with national stature in the discipline of statistics referred to it as the "Cadillac" of exit polls. In recent years, the exit poll has involved nearly 1000 student volunteers from participating institutions who have carried it out. Brigham Young University is the sponsoring institution with overall organization provided by students in the Department of Statistics and Department of Political Science under the direction of professors from both of those disciplines.

This year interviewers visited over 107 different polling places across the state. They administered questionnaires to randomly selected individuals after they had finished voting. The data from these questionnaires were called into the "data center" at BYU and the data was compiled and analyzed every few hours through the day. By the end of the day, nearly 6400 questionnaires were included in the database from which projections of the outcomes of the races of interest were made. This included a careful tally of non-responses as well as responses. (About 60% of individuals contacted by our interviewers were willing to fill out a questionnaire.)

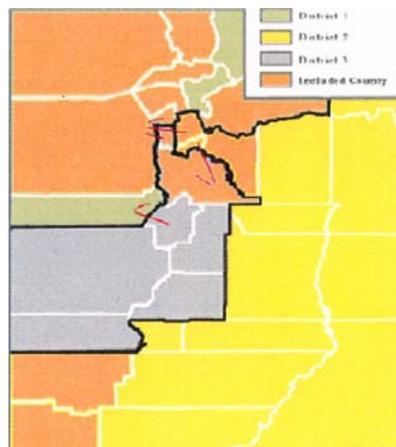
In addition to the head-to head political races that were included on the questionnaire, an array of demographic information was collected as well as attitudes and opinions on a variety of political and economic topics. Three primary questionnaires were used with as many as 50-60 questions being asked of each respondent.

At about 6:00 p.m., a final analysis was performed on the data and results were prepared for broadcasting on KBYU-TV at 8:00 p.m., after the polls had closed.

Participating students received experience in questionnaire design, data management, sample survey design, and statistical analysis. They also experienced the business side of managing a large-scale, complex sampling operation similar to those carried out by public opinion research firms and commercial sample survey organizations. They also came face-to-face with hard deadlines that have to be met, facing the pressures of having a show ready to go for both a 6 pm and an 8 pm broadcast. Students learned about state demographics, candidates and political issues, and the voting process used in the state of Utah.

Many important statistical concepts frequently encountered only in a text-book discussion immediately become real to them.

The following is a preliminary comparison of the UCEP results with actual results from the state elections office. The State numbers do not include absentee ballots and the results from provisional ballots. But all precincts had reported their election results from the November 4 balloting.



Individual	Actual Result	8 pm Projection	Margin of error
McCain	62.86	58.6	2.6
Obama	34.15	36.28	2.4
Other	2.99	5.12	1.0
Huntsman	77.73	78.3	1.6
Springmeyer	19.66	18.5	1.6
Other	2.61	3.2	0.6
Shurtleff	69.71	70.9	2.0
Hill	26.58	25.81	2.4
Other	3.71	3.29	0.6
Bishop	64.97	65.85	4.0
Bowen	30.43	31.04	4.6
Other	4.6	3.11	1.0
Matheson	63.27	66.96	3.2
Dew	34.6	30.4	2.9
Other	2.13	2.64	0.6
Chaffetz	65.84	68.12	3.6
Spencer	28.07	28.41	3.0
Other	6.09	3.47	1.4

# College Publications

## Chemistry & Biochemistry

D.E. Austin, Y. Peng, B.J. Hansen, I.W. Miller, A.L. Rockwood, A.R. Hawkins, and S.E. Toley, "Novel Ion Traps Using Planar Resistive Electrodes: Implications for Miniaturized Mass Analyzers," *J Am Soc Mass Spectrom*, **19**, 1435-1441 (2008).

T.F. Culwell, C.D. Thulin, K.J. Merrell, and S.W. Graves, "Influence of Diet on the Proteome of *Drosophila Melanogaster* as Assessed by Two-Dimensional Gel Electrophoresis and Capillary Liquid Chromatography-Mass Spectrometry: The Hamburger Effect Revisited," *Journal of Biomolecular Techniques*, **19**, 244-250 (2008).

## Computer Science

Jay McCarthy and Shriram Krishnamurthi. Minimal Backups of Cryptographic Protocols Runs. In Formal Methods in Security Engineering, 2008.

Jay McCarthy and Shriram Krishnamurthi. Cryptographic End-Point Projection and Explication. In European Symposium on Research in Computer Security, 2008.

L. Xu and D.W. Embley, Categorization of Web Documents using Extraction Ontologies, *International Journal of Metadata, Semantics and Ontologies*, Vol. 3, No. 1, 2008, 3-20.

R. S. Abbott, T. W. van der Horst, and K. E. Seamons. CPG: Closed Pseudonymous Groups. Workshop on Privacy in the Electronic Society (WPES 2008), Alexandria, VA, October 2008.

## Geological Sciences

McBride, J. H., Stephenson, W. S., Thompson, T. J., Harper, M. P., Eipert, A. A., Hoopes, J. C., Tingey, D. G., Keach, R. W., II, Okojie-Ayora, A. O., Gunderson, K. L., Meirovitz, C. D., Hicks, T. C., Spencer, C. J., Yaede, J. R., and Worley, D. M., 2008, A geophysical investigation of shallow deformation along an anomalous section of the Wasatch fault zone, Utah, USA, *Environmental & Engineering Geoscience*, v. 15, no. 3, p. 183-197.

Rupper, S. and G. Roe, 2008: Glacier changes and regional climate: a mass and energy balance approach. *Journal of Climate*, 21(20), 5384-5401.

## Mathematics

Peter Bates, Kening Lu and Chongchun Zeng, "Approximately invariant manifolds and global Dynamics of Spikes States," *Inventiones mathematicae* 174 p. 355-433 **2008**

Lennard Bakker, "Measurably Nonconjugate Higher-Rank Abelian Non-Cartan Actions," *Proceedings of Dynamic Systems and Applications*, 5 p. 53-59 **2008**

Kening Lu, Björn Schmalfuß, "Invariant foliations for stochastic partial differential equations," *Stochastics and Dynamics* Vol.8, No.3 p. 505-518 **2008**

Tomás Caraballo, Kening Lu, "Attractors for stochastic lattice dynamical systems with a multiplicative noise ," *Front. Math. China* Vol.3, No.3 p.317-335 **2008**

## Mathematics Education

## Physics and Astronomy

"Extreme-Ultraviolet Polarimeter Utilizing Laser-Generated High-Order Harmonics," N. Brimhall, M. Turner, N. Herrick, D. Allred, R. S. Turley, M. Ware, and J. Peatross, *Rev. Sci. Instrum.* 79, 103108 (2008).

## Statistics

Christensen WE, and Schauer JJ. (2008) "Quantifying and manipulating species influence in Positive Matrix Factorization," *Chemometrics and Intelligent Laboratory Systems*, **94**, 140-148.