Sara A. Stoudt

Contact Information	Evans Hall 331, Department of Statistics Berkeley, CA 94720-3860 http://github.com/sastoudt http://www.stat.berkeley.edu/∼sstoudt/ http://sastoudt.github.io/	724-464-3179 sstoudt@berkeley.edu @sastoudt

Research Interests: applied and computational statistics with applications to environmental data

Education

Smith College, Northampton, MA

B.A., Mathematics and Statistics , 2015
Magna Cum Laude with Highest Honors 3.95/4.0
Major GPA: 4.0/4.0

University of California, Berkeley, Berkeley, CA

Ph.D., Statistics, August 2015 - expected 2020
Advisors: Will Fithian and Perry de Valpine

Research Experience

Research Assistant May 2013 to August 2013, May 2014 to August 2014, May 2015 to August 2015, May 2016 to July 2016, May 2017 to July 2017

- Statistical Engineering Division, National Institute of Standards and Technology
  - Supervisor: Antonio Possolo, Ph.D
  - Measuring Optical Apertures for Solar Irradiance Monitoring (Circles of Best Fit, Uncertainty Quantification)
  - Homogenization of Surface Temperature Records (Time Series Analysis, Uncertainty Quantification)
  - Errors in Variables Modeling for Force Calibrations (Errors in Variables, Uncertainty Quantification)
  - Errors in Variables Modeling for Force Calibrations, Size Measurement of Nanoparticles, Implementations for Easy Use by Scientists (Shiny applications)
  - Interpolation of Atmospheric Greenhouse Gas Fluxes (Lattice Kriging), Evaluation of the accuracy, consistency, and stability of measurements of the Planck constant (meta-analysis), Gas Standard Reference Material Analysis (Shiny application)

Research Project Fall 2016- ongoing

- UC Berkeley
  - Supervisors: Will Fithian (Department of Statistics) and Perry de Valpine (Department of Environmental Science, Policy and Management)
  - Identifiability in Species Occupancy Models
  - parametric v. non-parametric identifiability in presence-only and presence-absence data regimes

Collaborations in Ecology Fall 2016-ongoing

- Fitting Models with Phylogenetic and Measurement Errors with Soorim Song (GitHub repository pending move from private to public)
- Hierarchical Modeling of Chronic Wasting Disease in Canadian Deer with Dana Seidel

Research Project August 2016

- San Francisco Estuary Institute (SFEI)
  - Mentors: David Senn (SFEI), Erica Spotswood (SFEI), Perry de Valpine (UC Berkeley, ESPM), Marcus Beck (EPA)
  - building Generalized Additive Models for understanding the variability in chlorophyll over space and time
  - comparing the use of Generalized Additive Models and Weighted Regression on Time, Discharge, and Season
  - https://github.com/sastoudt/DS421_summerProject/

Research Project February 2014 to May 2014, September 2014-April 2015

- Mathematics and Statistics Department, Smith College
  - Supervisor: Ben Baumer, Ph.D
  - Using Machine Learning to Predict March Madness (KNN, SVM, Neural Nets, Random Forests)
  - Honors Project: Geostatistical Models for the Spatial Variability of the Abundance of Uranium in Soils and Sediments of the Continental United States (Local Regression, Generalized Additive Models, Gaussian Processes/Kriging)

Research Project January 2015 to May 2015

- Mathematics and Statistics Department, Smith College
  - Supervisor: Nessy Tania, Ph.D.
  - Traffic Generation Model for Telecom Applications (Empirical Copulae)

Research Assistant February 2013 to March 2014

- Mathematics and Statistics Department, Smith College
  - Supervisor: Nicholas Horton, Ph.D
  - Roadless America: An Activity for Introduction to Sample Proportions and Their Confidence Intervals (Statistics Education)

Research Project September 2013 to December 2013

- Mathematics and Statistics Department, Smith College
  - Supervisor: Katherine Halvorsen, Ph.D.
  - Factors Associated With Changes in Academic Performance During the Transition from Elementary to Middle School (Multiple Regression and Analysis of Variance)

Awards

- National Physical Science Consortium Fellow 2015-2018
- Data Sciences for the 21st Century: Environment and Society Graduate Training Program 2015-2017
- Gertrude M. Cox Scholarship 2015
- 2nd Place USPROC Undergraduate Research Project Competition 2015
- Geocomputation Conference Best Poster Award 2015
- Elected to Mu Sigma Rho 2015
- Goldwater Scholar 2014
- Elected to Phi Beta Kappa Society 2014
- First Place: Statistics in Sports Undergraduate Research Competition at JSM 2014
- Honorable Mention Undergraduate CLASS Project Competition 2014
- Best in Show- Five College Data Fest 2014,2015
- Elected to Sigma Xi 2013

Smith College

- Ann Kirsten Pokora Prize for excellence in mathematics 2015
- John Everett Brady Prize for best performance in the beginning Latin course 2013
- Suzan Rose Benedict Prize for excellence in mathematics 2013
- Smith College Dean's List 2011-15
- Smith College STRIDE (Student Research in Departments) Scholarship 2011-15

Publications

- **Stoudt, S.** "Geostatistical Models for the Spatial Distribution of Uranium in the Continental United States" *Advances in Geocomputation: Geocomputation 2015 - The 13th International Conference* Springer Advances in Geographic Information Science, 2017, pp. 325-334. http://www.springer.com/us/book/9783319227856
- **Stoudt, S.**, Badian-Pessot, P., Mahop, B. N., Earley, E., Menter, J., Flores, Y., Williams, D., Zhang, W., Maharajan, L., Bao, Y., Rosenbauer, L., Nguyen, V., Mendiratta, V., Tania, N. "Modeling Internet Traffic Generations Based on Individual Users and Activities for Telecommunication Applications" *American Journal of Undergraduate Research* Volume 13, Issue 3, August 2016, pp. 53-65. http://www.ajuronline.org/uploads/Volume_13_3/AJUR%20Vol%2013%20Issue%203 %2008.25.16%20pp.53-65.pdf
- Possolo, A., Schlamminger, S., **Stoudt, S.**, Pratt, J. R., and Williams, C. J. "Evaluation of the accuracy, consistency, and stability of measurements of the Planck constant used in the redefinition of the International System of Units" *Metrologia* http://iopscience.iop.org/article/10.1088/1681-7575/aa966c
- Bartel, T., Possolo, A., and **Stoudt, S**. "Force Calibrations using Errors-in-Variables Regression and Monte Carlo Uncertainty Evaluations" *Metrologia* Volume 53, Number 3, June 2016, pp. 965-980(16). http://iopscience.iop.org/article/10.1088/0026-1394/53/3/965/meta
- **Stoudt, S.**, Cao, Y., Udwin, D., and Horton, N. J. "What Percent of the Continental US is Within One Mile of a Road?" *Statistics Education Web,* 2014. http://www.amstat.org/education/stew/pdfs/PercentWithinMileofRoad.pdf
- **Stoudt, S.**, Santana, L., and Baumer, B. "In Pursuit of Perfection: An Ensemble Method for Predicting March Madness Match-Up Probabilities" *JSM 2014 Proceedings* http://www.science.smith.edu/~bbaumer/w/pub/jsm2014_perfect_bracket.pdf

Papers in Preparation

- Hong, J., **Stoudt, S.**, and de Valpine, P. "Sampling-Based Approaches to Maximum Likelihood Estimation for Latent Variable Models" (submitted to the Journal of Computational and Graphical Statistics)
- Kerlo, A. M., **Stoudt, S.**, Hajni, K., Lavoie, T., Shepson, P. B., and Possolo, A. "Improved Interpolation of Atmospheric Greenhouse Gas Fluxes Using Statistical Models"
- **Stoudt, S.**, Udwin, D. , and Francis, K. " 'Cold Enough for You?': Developing Data Science Skills through Exploration of Weather Perceptions on Twitter"
- Calderon Jimenez, B., Sarmanho, G., **Stoudt, S.**, Montoro Bustos, A., Strenge, I.H., Johnson, M.E., Murphy, K.E., Possolo, A., ad Winchester, M.R. "*nanoICP-MS:* a new statistical and interactive web application for the determination and quantification of metallic nanoparticles by single particle ICP-MS"

Programming Languages

- Proficient: R, Matlab, LaTeX
- Experience With: Python, SQL, Mathematica, WinBUGS, Java, GIS, AMPL, NIMBLE, D3, JavaScript, bash

Presentations

- Interdisciplinary Graduate Education in Data Science: DS421 NRT
  Poster
  Berkeley Institute for Data Science Data Science Faire, UC Berkeley May 2017 With DS421 Cohort 1

- Streamlining Climate Model Accessibility for Integration into Site-Specific Life Science Research

Talk
Data Science for the 21st Century Annual Symposium, UC Berkeley May 2017 With Jenna Baughman

- Sampling-Based Approaches to Maximum Likelihood Estimation for Latent Variable Models
  Poster
  BSTARS, UC Berkeley March 2017
  With Johnny Hong and Perry de Valpine

- Uncertainty Quantification and Statistics
  Talk (Invited)
  NIST Presentation to SPIRAL students, July 2016 and July 2017

- Geostatistical Models for the Spatial Distribution of Uranium in the Continental United States
  Plenary Talk (Invited)
  First Electronic Undergraduate Statistics Research Conference, October 2015

- Internet Traffic Generation
  Talk (Invited)
  MAA Mathfest, August 2015
  With Erika Earley, Yadira Flores, and Jordan Menter

- "Big Force" Calibrations: An Errors in Variables Approach
  Talk
  Summer Undergraduate Research Fellow Colloquia, National Institute of Standards and Technology, August 2015
  With Antonio Possolo and Tom Bartel

- Geostatistical Models for the Spatial Distribution of Uranium in the Continental United States
  Poster
  Geocomputation, May 2015

- Correcting Temperature Records for Biases Unrelated to the Climate
  Summer Undergraduate Research Fellow Colloquia, National Institute of Standards and Technology, August 2014
  With Antonio Possolo

- The Perfect Bracket: Machine Learning in NCAA Basketball
  SPEED poster and presentation
  Joint Statistical Meetings, August 2014
  With Loren Santana and Ben Baumer

- Taking a Closer Look at Learning: Factors Associated with Changes in Academic Performance During the Transition from Elementary to Middle School
  Poster
  Women in Statistics Conference, May 2014
  With Dana Hsu, Anna Rockower, and Katherine Halvorsen

- Measuring Optical Apertures for Solar Irradiance Monitoring *Plenary Talk*
  Summer Undergraduate Research Fellow Colloquia, National Institute of Standards and Technology, August 2013

With Maritoni Litorja, and Antonio Possolo

Smith College

- Geostatistical Models for the Spatial Distribution of Uranium in the Continental United States
  Poster
  Collaborations, April 2015

- Geostatistical Models for the Spatial Distribution of Uranium in the Continental United States
  Talk
  Thesis Defense, April 2015

- Traffic Generation Model for Telecom Applications: Simulating Internet Traffic Using the Empirical Copula Method
  Poster
  Collaborations, April 2015

- Traffic Generation Model for Telecom Applications: Simulating Internet Traffic Using the Empirical Copula Method
  Talk
  Center for Women in Mathematics, April 2015

- Beyond Calculations: Communicating through Statistics
  Talk (Invited)
  Smith in the World Conference, November 2014

- Correcting Temperature Records for Biases Unrelated to the Climate
  Talk
  WIMIN Conference, September 2014
  With Antonio Possolo

- The Perfect Bracket: Machine Learning in NCAA Basketball
  Talk
  Center for Women in Mathematics, April 2014
  With Loren Santana, and Ben Baumer

- WeathR - Social Media: Changing the Way We Talk About the Weather
  Talk
  Collaborations, April 2014
  With Dana Udwin and Kelly Francis

- The Perfect Bracket: Machine Learning in NCAA Basketball
  Poster
  Collaborations, April 2014
  With Loren Santana, Ben Baumer

- An Objective Ranking of Best College Sports Coaches of all Time
  Poster
  Collaborations, April 2014
  With Maja Milosavljevic and Yu Jin

- Measuring Optical Apertures for Solar Irradiance Monitoring
  Talk

SMATH Conference, September 2013
With Maritoni Litorja and Antonio Possolo

- Sampling With Google Maps and R to Estimate Population Proportions
  Poster
  Collaborations, April 2013
  With Yue Cao, Dana Udwin, and Nicholas Horton

Teaching Experience

Graduate Student Instructor, Statistics, UC Berkeley Fall 2017

- Communicating with Data: The Art of Writing for Data science with Deborah Nolan
- Assist with course development and instruction of new course

Teaching Assistant, Mathematics and Statistics, Smith College September 2013– May 2014, September 2014–May 2015

- available for assistance with Calculus, Discrete Mathematics, Linear Algebra, Introductory and Intermediate Statistics
- Introduction to Statistics Grader (Fall 2013)
- Data Science Grader (Fall 2014)
- Mathematical Statistics Grader (Spring 2015)

Relevant Activities

- Attended the 3rd Annual Graduate Workshop on Environmental Data Analytics 2016
- Attended the Google Geospatial Technology Workshop 2015
- Attended the Statistical and Applied Mathematical Science Institute's Summer Program: The International Temperature Initiative 2014
- Editor in Chief of Smith College's Science Journal, Smith Scientific, aimed towards raising awareness of on-campus student and faculty research as well as making the sciences accessible to those in other fields 2013-4
- Participant in the Mathematical Contest in Modeling 2014
- Participant in the Kaggle March Machine Learning Madness Competition
  top 20 in stage one 2014
- Head Writer for Mathematics Section of Smith College's Science Journal 2012
- Writer for Mathematics Section of Smith College's Science Journal 2011

R Packages in Progress: Available Upon Request

- **RegionalHomogenization** *(includes a vignette)*: This package provides a way to access data from the International Surface Temperature Initiative (ISTI) and format it in a way that can be used with an algorithm for the homogenization of surface temperature series. This method does not explicitly find change-points but aims to remove the spurious bias that is not related to climate change as well as quantify the uncertainty surrounding this adjustment.
- **ciRcleFits**: Following Nikolai Chernov's MATLAB implementation of various circle fits (http://people.cas.uab.edu/ mosya/cl/MATLABcircle.html), this package contains R implementations of these fits. This package also contains functions for visualizing radial residuals and an application of these visualization techniques for time series data.

Final Class Projects

- Spatio-temporal Relationships of Water Quality Measurements in the Bay Delta: A Time and Frequency Domain Approach, Spring 2017
  https://github.com/sastoudt/stat248project_BayDeltaTS
- Streamlining Climate Model Accessibility for Integration into Site-Specific Life Science Research, Spring 2017
  https://github.com/sastoudt/UCB_DS421_NEX_partnerProject
- Maximum Likelihood Estimation with Latent Variables, Fall 2016 (GitHub repository

pending move from private to public)
- Graphical Model Structure Finding, Fall 2016
- Drought Vulnerability in the Continental United States, Spring 2016
  https://github.com/bolliger32/US_drought_vulnerability
- Cluster Randomization, Spring 2016
- Sensitivity Analysis: Random Walker Segmentation, Spring 2014
- Network Centrality, Spring 2014
- Weather Analysis of Smith's Field Station Data, January Term 2014
- WeathR- Perceptions of Weather in Social Media, Fall 2013
- Age Specific Population Dynamics, Fall 2013
- Periodicity and Climate Data: A Look Into the 100,000 Year Problem, Spring 2013
- Implementing the Graph Data Structure in Java, Spring 2013
- What is Associated with United Nation Members' Gross Domestic Product per Capita? Examining Geographical, Environmental, Political, and Social Factors, Fall 2012

Departmental Service

- Co-president of the Statistics Graduate Student Association, Fall 2017-Spring 2018
- Co-organizer of UC Berkeley DataFest, Spring 2016 and 2017
- Co-organizer of Statistics Graduate Student Association Gender Issues Roundtable Discussion, Fall 2016
  https://bids.berkeley.edu/news/gender-issues-roundtable-discussion-case-study-uncomfortable-conversations
- Co-organizer of Statistics Graduate Student Association Diversity Discussion, Spring 2017
- Statistics Graduate Student Association Hospitality Committee Member, Fall 2016-Spring 2017
- Statistics Graduate Student Association Diversity Affairs Member, Fall 2016- Sprint 2017
- Graduate Student Volunteer: "Roadless America" Interactive Activity for Cal Day 2016 and 2017

Review Service

- Journal of Statistics Education
- NIST Washington Review Editorial Board (Summer '17)

Professional Memberships

- American Statistical Association
- Sigma Xi
- Phi Beta Kappa
- Mu Sigma Rho

References available upon request