

## Sara A. Stoudt, PhD

---

CONTACT INFORMATION	Department of Mathematics, Bucknell University Lewisburg, PA <a href="http://sastoudt.github.io/">http://sastoudt.github.io/</a> <a href="http://github.com/sastoudt">http://github.com/sastoudt</a>	724-464-3179 <a href="mailto:sas072@bucknell.edu">sas072@bucknell.edu</a> @sastoudt
RESEARCH INTERESTS	applied and computational statistics in ecology, environmental science, and metrology, teaching the communication of statistics, data journalism	
EDUCATION	<b>University of California, Berkeley</b> , Berkeley, CA  Ph.D., Statistics, Fall 2015 - Summer 2020 Advisors: Will Fithian (Department of Statistics) and Perry de Valpine (Department of Environmental Science, Policy and Management) <i>National Physical Science Consortium Fellow</i> <i>National Science Foundation Research Traineeship: Data Sciences for the 21st Century, Environment and Society</i>  <b>Smith College</b> , Northampton, MA  B.A., Mathematics and Statistics, 2015	
DISSERTATION RESEARCH	<ul style="list-style-type: none"><li>• identifiability in species distribution and abundance models</li><li>• community metric estimation in latent factor joint species distribution models under model mis-specification</li><li>• statistics communication: <i>Communicating with Data: The Art of Writing for Data Science</i> (book written with Deborah Nolan, Oxford University Press, 2021)</li></ul>	
JOB EXPERIENCE	<b>Bucknell University</b> Starting Fall 2021 <ul style="list-style-type: none"><li>• tenure-track Assistant Professor in Mathematics Department</li></ul> <b>Berkeley Institute for Data Sciences Fellow</b> Fall 2018 - Spring 2020 <ul style="list-style-type: none"><li>• Diversity and Inclusion Working Group</li><li>• Best Practices Working Group</li></ul> <b>Data Desk Intern - Los Angeles Times</b> Summer 2019 <ul style="list-style-type: none"><li>• Census data aggregator</li><li>• Governor Newsom social media analysis</li></ul> <b>Data Science Intern - Farmers Business Network</b> Summer 2018 <ul style="list-style-type: none"><li>• Crop yield prediction</li></ul> <b>Summer (Undergraduate/Graduate) Research Fellow</b> Summers 2013-2017 <ul style="list-style-type: none"><li>• Statistical Engineering Division, National Institute of Standards and Technology<ul style="list-style-type: none"><li>– Measuring optical apertures for solar irradiance monitoring</li><li>– Homogenization of surface temperature records</li><li>– Errors-in-variables modeling for force calibrations</li><li>– Interpolating atmospheric greenhouse gas fluxes</li><li>– Evaluating the accuracy, consistency, and stability of measurements of the Planck constant</li></ul></li></ul>	
TEACHING EXPERIENCE	<b>Lecturer, Statistical and Data Sciences Program, Smith College</b> <ul style="list-style-type: none"><li>• Communicating with Data (remote, almost 70 students) Fall 2020</li><li>• Introduction to Probability and Statistics (remote, about 45 students) Fall 2020, Spring 2021</li></ul>	

### Graduate Student Instructor (GSI), Statistics, UC Berkeley

- Co-developed and co-instructed new writing in statistics course (about 15 students) Fall 2017
- Co-developed and co-instructed new blogging for data science independent study (about 5 students) Spring 2018
- Outstanding GSI award

### Miscellaneous Primary Instructor

- Data Storytelling Workshop (Correlation One's Women's Summit, 2020) (cancelled due to COVID-19)
- [Introduction to Statistics in R Workshop](#) (NICAR, 2020)
- [Unleashing the power of biodiversity data with data science](#) (Lewis and Clark workshop 2019, co-instructed with Ciera Martinez)

### Miscellaneous Teaching Assistant

- Linear Regression Workshop (NICAR, 2020)
- Peer Data Review coach (Open News, 2019)
- R bootcamp (D-Lab/UC Berkeley Statistics Department, 2018)
- Software Carpentry (Data Science for the 21st Century Training Program, 2018)

### PROGRAMMING LANGUAGES

- Advanced: R (and Shiny)
- Experience With: Python, SQL, HTML, CSS, D3, JavaScript, bash, Matlab, Java, GIS, AMPL, Mathematica, NIMBLE, WinBUGS, LaTeX

### PEER-REVIEWED PUBLICATIONS

- Nolan, D., and **Stoudt, S.** "The Promise of Portfolios: Training Modern Data Scientist" (accepted *Harvard Data Science Review*)
- **Stoudt, S.**, and Possolo, A. "Statistical and Computational Tools for Metrologists" (in press) chapter for *Advanced Mathematical and Computational Tools in Metrology and Testing XII*
- **Stoudt, S.\***, Vasquez, V.\*, and Martinez, C\*. "Principles for data analysis workflows" (\*equal authorship) *PLOS Computational Biology* Volume 17, Number 3, 2021.
- **Stoudt, S.**, Pintar, A., and Possolo, A. "Coverage Intervals" *Journal of Research of National Institute of Standards and Technology* Volume 126, Number 126004, March 2021
- **Stoudt, S.**, Pintar, A., and Possolo, A. "Uncertainty Evaluations from Small Datasets" *Metrologia* Volume 58, Number 1, January 2021
- Possolo, A., Schlamming, 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* Volume 55, Number 1, December 2017
- **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.
- **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., and 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.
- 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).
- **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.

PUBLICATIONS IN  
PROGRESS

- **Stoudt, S.\***, Goldstein, B. R.\*, and de Valpine, P. (\*equal authorship) “Identifying Charismatic Bird Species and Traits with Community Science Observations” (submitted, [preprint on bioRxiv](#))
- Nolan, D. and **Stoudt, S.** “Captions: The Unsung Heroes of Data Communication” (submitted for IASE 2021)
- **Stoudt, S.**, de Valpine, P., and Fithian, W. “Semi-parametric Methods for Assessing Identifiability of Species Distribution and Abundance Models” (submitted to *Ecology*)
- Nolan, D., and **Stoudt, S.** “Storyboarding as an Exploratory Data Analysis Companion” (in preparation)
- Hong, J., **Stoudt, S.**, and de Valpine, P. “Fast maximum likelihood estimation for general hierarchical models” (in revision)
- **Stoudt, S.**, de Valpine, P. and Fithian, W. “Stress Testing Latent Factor Approaches to Joint Species Distribution Models” (in preparation)

OTHER  
PUBLICATIONS

- Goeva, A., Jones, P., **Stoudt, S.**, and Trisovic, A. “[Recipes for Connector Courses from the Early-Career Board Kitchen](#)” *Harvard Data Science Review*, June 2021.
- Frost, S\*., Goeva, A\*., Pombra, J\*., Seaton, W\*. **Stoudt, S\*.**, Trisovic, A\*, Wang, C\*., and Zucker, C\*.. “[Kaleidoscopic Perspectives on Practicum-based Data Science Education](#)” (\*equal authorship) *Harvard Data Science Review*, February 2021.
- Nolan, D., and **Stoudt, S.** “[Reading to Write](#)” *Significance*, December 2020.
- Goeva, A\*., **Stoudt, S\*.**, and Trisovic, A\*. “[Toward Reproducible and Extensible Research: from Values to Action](#)” (\*equal authorship) *Harvard Data Science Review*, December 2020.
- Frost, S\*., Goeva, A\*., Seaton, W\*. **Stoudt, S\*.**, and Trisovic, A\*. “[Early-Career View on Data Science Challenges: Responsibility, Rigor, and Accessibility](#)” (\*equal authorship) *Harvard Data Science Review*, September 2020.
- Sholler, D., **Stoudt, S.**, Kennedy, C., Hoces de la Guardia, F., Lanusse, F., Ram, K., Ottoboni, K., Stuart, M., Vareth, M., Varoquaux, N. , Barter, R. , Geiger, R. S., Peterson, S., and van der Walt, S. “[Resistance to Adoption of Best Practices: a report from the Berkeley Institute for Data Science’s Best Practices in Data Science Series](#)”
- Geiger, R. S., Sholler, D., Culich, A., Martinez, C., Hoces de la Guardia, F., Lanusse, R., Ottoboni, K., Stuart, M., Vareth, M., Varoquaux, N., **Stoudt, S.**, and van der Walt, S. “[Challenges of Doing Data-Intensive Research in Teams, Labs, and Groups: Report from the BIDS Best Practices in Data Science Series](#)”
- Geiger, R. S., DeMasi, O., Culich, A. Zoglauer, A., Das, D. Hoces de la Guardia, F., Ottoboni, K., Fenner, M., Varoquaux, N., Barter, R., Barnes, R., **Stoudt, S.**, Dorton, S., and van der Walt, S. “[Best Practices for Fostering Diversity and Inclusion in Data Science: Report from the BIDS Best Practices in Data Science Series](#)”
- Sholler, D., Das, D., Hoces de la Guardia, F., Hoffmann, C., Lanusse, F., Varoquaux, N., Garcia, R., Geiger, R. S., McDevitt, S., Peterson, S., and **Stoudt, S.** “[Best Practices for Managing Turnover in Data Science Groups, Teams, and Labs](#)”
- **Stoudt, S.**, Santana, L., and Baumer, B. “[In Pursuit of Perfection: An Ensemble Method for Predicting March Madness Match-Up Probabilities](#)” *JSM 2014 Proceedings*

NON-ACADEMIC  
WRITING

- [COVID-19 Case Fatality Rate Bias Visual Explainer](#) with Aleksandrina Goeva, *Harvard Data Science Review*, February 2021.
- “[Data Visualization Beyond the Screen](#)”, Northeast Big Data Innovation Hub Blog, January 2021.
- scripts for 15 episodes of Study Hall: Data Literacy (produced by Arizona State University and the Crash Course team at Complexly) (released weekly throughout Fall 2020)
- [Communicating with Data](#) newsletter, jointly written with students in my Fall '20 course.

- Ecology for the Masses Stats Corner
- Interviewed for Music Journalism Insider newsletter
- Just how does Kidz Bop censor songs? (The Pudding)
- You Know Karen (The Pudding)
- Berkeley Science Review
- Logic Magazine
- Tidy Tuesday and #rstats blog
- What does probability mean anyway?
- Fixed, mixed, and random effects

#### AWARDS

- RStudio Diversity Scholar 2020
- Berkeley Institute for Data Science Fellow 2018-2020
- Outstanding Graduate Student Instructor award 2018
- 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
- Best Poster Award: Geocomputation Conference 2015
- Elected to Mu Sigma Rho 2015
- Ann Kirsten Pokora Prize for excellence in mathematics at Smith College 2015
- Goldwater Scholar 2014
- Elected to Phi Beta Kappa Society 2014
- First Place: Statistics in Sports Undergraduate Research Competition at Joint Statistical Meetings 2014
- Honorable Mention Undergraduate CLASS Project Competition 2014
- Best in Show: Five College Data Fest 2014, 2015
- Elected to Sigma Xi (student) 2013
- Suzan Rose Benedict Prize for excellence in mathematics at Smith College 2013
- Smith College STRIDE (Student Research in Departments) Scholarship at Smith College 2011-15

#### GRANTS

- awarded (co-PI) Code for Science and Society Virtual Event Grant, 2021.
- awarded (co-PI) Academic Data Science Alliance Career Development Network Seed Grant, 2021.
- awarded Curriculum Enhancement Grant from Smith College's Design Thinking Initiative to fund mailing maker kits to Communicating with Data students, Fall 2020.
- awarded grants from UC Berkeley Wellness Fund, UC Berkeley Chancellor's Advisory Committee on Student Services and Fees, and UC Berkeley Student Technology Fund Committee to fund the "Fostering diverse and inclusive data science at Berkeley" series at BIDS (2019)

#### SERVICE

- co-organizer of Data Science by Design (DSxD) initiative, including Creator Conf in May 2021
- organizer of graduate school panel for Smith College Statistical and Data Sciences Program and Five College Statistics students, November 2020
- member of Harvard Data Science Review Early Career Board, Fall 2020-ongoing
- Women in Statistics and Data Science Rotating Twitter Curator, August 2020
- Co-organizer of Code and Coffee East Bay for the Bay Area Women in Machine Learning & Data Science Meetup group, Spring 2019 - Spring 2020
- Member of BIDS Diversity and Inclusion Working Group, Fall 2018 - Spring 2020
- We Are RLadies Rotating Twitter Curator, October 2018

- Co-president of the Statistics Graduate Student Association (SGSA), Fall 2017 - Spring 2018
- Co-organizer of UC Berkeley DataFest, Springs 2016 - 2018
- Co-organizer of [SGSA Gender Issues Roundtable Discussion](#), Fall 2016
- Co-organizer of Statistics Graduate Student Association Diversity Discussion, Spring 2017
- Statistics Graduate Student Association Diversity Affairs Member, Fall 2016 - Spring 2017
- Graduate Student Volunteer: “Roadless America” Interactive Activity for Cal Day 2016 and 2017

#### RELEVANT ACTIVITIES

- [Blogdown website production](#) for the Murmuration project
- Author of [base R to dplyr vignette](#) 2020
- Author of [base R to stringr vignette](#) 2019
- Developer of interactive visualization for the [Wealth Tax Simulator](#) 2019
- NCAR Graduate Workshop on Environmental Data Analytics (by application) 2016
- [San Francisco Estuary Institute consulting](#) 2016
- SAMSI International Temperature Initiative (by application) 2014

#### PRESENTATIONS

- All Species Distribution and Abundance Models Are Wrong, Which Are Useful?  
*Talk*, TWS Annual Conference, November 2021, joint work with Perry de Valpine and Will Fithian
- [Storyboarding as Part of the Process of Statistical Investigation](#)  
*Lightning Talk and Live Discussion*, USCOTS Beyond Session, June 2021, with Deborah Nolan
- Communicating with Data: Practicing and Teaching  
*Talk*, Data Science for Social and Environmental Justice Virtual Writing and Research Development Group, June 2021, joint work with Deborah Nolan
- [Data Communication](#)  
*Panelist*, National Workshop on Data Science Education, UC Berkeley, June 2021, with Jeff Leek, Amelia McNamara, Trity Pourbahrami, and Deborah Nolan
- Creativity to Learn, Creativity to Teach  
*Session Co-lead*, Data Science by Design Creator Conf, May 2021, with Allison Horst and Sean Kross
- [Generalized Additive Models: Allowing for some wiggle room in your models](#)  
*Invited Talk*, Social Science Data Lab, Mannheim Centre for European Social Research at the University of Mannheim, Germany, March 2021
- From No Prereqs to Data Storytellers: Borrowing from the Journalist’s Toolbox to Teach Students to Communicate with Data  
*Talk*, Computation+Journalism Conference, February 2021
- The Art of Writing for Data Science  
*Guest Lecture*, Communicating Data and Statistics course, Columbia Journalism School, February 2021
- Ideas Swap on Using Kits for Hands-on Making  
*Panel*, Smith College, January 2021

- Species Distribution and Abundance Models: The Good, The Bad, and The Not Identifiable  
*Talk*, Smith College Sigma Xi Lectures, November 2020
- Joint Species Distribution Models: Are they walking the walk or just talking the talk?  
*Talk (Invited)*, Boston Women in Machine Learning and Data Science, October 2020
- Econometrics Meets Ecology  
*Talk (Invited)*, USFCA students at Women and Diversity in Economics Club, October 2020
- Communicating with Data: How and where does it fit in the data science curriculum?  
*Breakout Session Co-lead*, Academic Data Science Alliance Annual Meeting, October 2020, with Deborah Nolan
- Principles for data-intensive research workflows: Guidance for the classroom and the computational laboratory  
*Breakout Session Co-lead*, Academic Data Science Alliance Annual Meeting, October 2020, with Valeri Vasquez and Ciera Martinez
- An overview of and lessons learned from hosting a data science workshop series for undergraduate students from under-represented backgrounds  
*Co-presenter*, Academic Data Science Alliance Annual Meeting, October 2020, with Orianna Demasi and Stacey Dorton
- Saying “Yes”: A Data Memoir  
*Talk (Invited)*, RLadies Amherst, April 2020
- [Groove is in the Heart and the Data](#)  
*Lightning Talk*, RLadies San Francisco, December 2019
- Talking With the Public About Data Science  
*Panelist*, Moore-Sloan Data Science Environment Annual Summit, November 2019, with Joshua Tucker, Andrea Jones-Rooy, and moderator Meredith Broussard
- Diversity and Inclusion in Data Science  
*Talk and Discussion Lead*, Moore-Sloan Data Science Environment Annual Summit, November 2019, on behalf of BIDS Diversity and Inclusion Working Group
- Goodness-of-Fit Checks and Diagnostic Plots for Hierarchical Joint Species Distribution Models  
*Poster*, American Fisheries Society and The Wildlife Society Joint Annual Conference, September 2019, with Will Fithian and Perry de Valpine
- Species Distribution and Abundance Models: The Good, The Bad, and The Not Identifiable  
*Talk and Poster*, Berkeley Statistics Annual Research Symposium, UC Berkeley March 2019, with Will Fithian and Perry de Valpine
- Identifiability in the Wild: Econometrics Meets Ecology  
*Talk*, Third Annual Berkeley-Stanford Econometrics Jamboree, UC Berkeley, November 2018, with Will Fithian and Perry de Valpine

- Clarifying the Identifiability Controversy in Species Distribution Modeling  
*Poster*, Berkeley Statistics Annual Research Symposium, UC Berkeley March 2018, with Will Fithian and Perry de Valpine
- Sampling-Based Approaches to Maximum Likelihood Estimation for Latent Variable Models  
*Poster*, Berkeley Statistics Annual Research Symposium, UC Berkeley March 2017, with Johnny Hong and Perry de Valpine
- 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  
*Talk*, Summer Undergraduate Research Fellow Colloquia, National Institute of Standards and Technology, August 2014 (also given at WIMIN Conference in September 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 (Invited)*, Summer Undergraduate Research Fellow Colloquia, National Institute of Standards and Technology, August 2013, (also given at SMATH conference in September 2013) with Maritoni Litorja, and Antonio Possolo