

Ray Bai

LeConte College 207
1523 Greene St
Columbia, SC 29208

Email: rbai@mailbox.sc.edu
Homepage: <http://www.raybai.net>
Office phone: (803) 777-5893, Ext. 75893
Cell: (617) 953-3266

Academic Employment

- 2020- Assistant Professor of Statistics, University of South Carolina
- 2018-2020 Postdoctoral Researcher in Biostatistics and Informatics, University of Pennsylvania
Supervisors: Yong Chen and Mary Boland

Education

- 2014-2018 PhD in Statistics, University of Florida
Dissertation Title: Bayesian High-Dimensional Models with Scale-Mixture Shrinkage Priors
Advisor: Malay Ghosh
- 2014-2016 Master of Statistics, University of Florida
Project Title: Irrigation Forecasting in Southwest Florida
Advisor: Nikolay Bliznyuk
- 2010-2012 MS in Applied Mathematics, University of Massachusetts Amherst
- 2003-2007 BA in Economics and Government, Cornell University

Publications

(____ = student author, * = co-first author)

1. **Bai, R.***, Moran, G. E.* , Antonelli, J. L.* , Chen, Y., and Boland, M. R. (2022). Spike-and-slab group lassos for grouped regression and sparse generalized additive models. *Journal of the American Statistical Association*, **117**: 184-197.
2. Meeker, J. R., Burris, H. H., **Bai, R.**, Levine, L. D., and Boland, M. R. (2022). Neighborhood deprivation increases the risk of post-induction cesarean delivery. *Journal of the American Medical Informatics Association*, **29**: 329-334.
3. **Bai, R.**, Ročková, V., and George, E. I. (2021). Spike-and-slab meets LASSO: A review of the spike-and-slab LASSO. In Tadesse, M. G. and Vannucci, M. (Eds.), *Handbook of Bayesian Variable Selection*, 81-108. Chapman & Hall/CRC Press.
4. Meeker, J. R., Canelón, S. P., **Bai, R.**, Levine, L. D., and Boland, M. R. (2021). Individual- and neighborhood-level risk factors for severe maternal morbidity. *Obstetrics & Gynecology*, **137**: 847-854.
5. **Bai, R.** and Ghosh, M. (2021). On the beta prime prior for scale parameters in high-dimensional Bayesian regression models. *Statistica Sinica*, **31**: 843-865.

6. Boland, M. R., Liu, J., Balocchi, C., Meeker, J., **Bai, R.**, Mellis, I., Mowery, D. L., and Herman, D. (2021). Association of neighborhood-level factors and COVID-19 infection patterns in Philadelphia using spatial regression. *AMIA Annual Symposium Proceedings*, **2021**: 545-554.
7. **Bai, R.** and Ghosh, M. (2019). Large-scale multiple hypothesis testing with the normal-beta prime prior. *Statistics*, **53**: 1210-1233.
8. **Bai, R.** and Ghosh, M. (2018). High-dimensional multivariate posterior consistency under global-local shrinkage priors. *Journal of Multivariate Analysis*, **167**: 157-170.
9. Duerr, I., Merrill, H. R., Wang, C., **Bai, R.**, Boyer, M., Dukes, M. D., and Bliznyuk, N. (2018). Forecasting urban water demand with statistical and machine learning methods using large space-time data. *Environmental Modelling and Software*, **102**: 29-38.

Technical Reports and Manuscripts in Preparation

(____ = student author, * = co-first author)

10. Wang, S., Shin, M., and **Bai, R.** (2023+). Generative quantile regression with variability penalty. arXiv:2301.03661.
11. Liu, Q., Huang, X., and **Bai, R.** (2023+). Bayesian modal regression based on mixture distributions. arXiv:2211.10776.
12. Zgodic, A., **Bai, R.**, Zhang, J., Wang, Y., Bondell, H., and McLain, A. C. (2023+). Sparse high-dimensional linear mixed modeling with a partitioned empirical Bayes ECM algorithm. *Preprint*.
13. Wang, S.-H., **Bai, R.**, and Huang, H.-H. (2023+). Ultrahigh dimensional variable selection for mixed-type multivariate response Bayesian generalized linear models. arXiv:2201.12839.
14. Balocchi, C.*, **Bai, R.***, Liu, J., Canelón, S. P., George, E. I., Chen, Y., and Boland, M. R. (2023+). Uncovering patterns for adverse pregnancy outcomes with spatial analysis: Evidence from Philadelphia. arXiv:2105.04981.
15. **Bai, R.**, Boland, M. R., and Chen, Y. (2023+). NVC-SSL: Nonparametric varying coefficient spike-and-slab lasso for high-dimensional Bayesian varying coefficient models. arXiv:1907.06477.
16. **Bai, R.**, Lin, L., Boland, M. R., and Chen, Y. (2023+). A robust Bayesian Copas selection model for correcting and quantifying the impact of publication bias. arXiv:2005.02930.
17. **Bai, R.***, Liu, X.*, Lin, L., Liu, Y., Kimmel, S. E., Chu, H., and Chen, Y. (2023+). A Bayesian selection model for correcting outcome reporting bias with application to a meta-analysis on heart failure interventions. arXiv:2110.08849.
18. Deshpande, S. K., **Bai, R.**, Balocchi, C., Starling, J. E., and Weiss, J. (2023+). VCBART: Bayesian trees for varying coefficients. arXiv:2003.06416.
19. **Bai, R.** (2023+). Spike-and-slab group lasso for consistent estimation and variable selection in non-Gaussian generalized additive models. arXiv:2007.07021.
20. **Bai, R.** (2023+). Contraction for Bayesian varying coefficient models with functional random effects and continuous shrinkage priors. *In preparation*.

Funding

National Science Foundation (May 2022-July 2023). Collaborative Research: Bayesian and Semi-Bayesian Methods for Detecting Relationships in High Dimensions (DMS-2015528). Role: **PI**¹ (in collaboration with Jun Liu at Harvard University). Amount: \$93,831.

University of South Carolina ASPIRE-I, Track 1 grant (Jul 1, 2022 - Sep 30, 2023). Scalable Bayesian Survival Analysis with Informative Cluster Size. Role: **PI**. Amount: \$13,583.

National Science Foundation (August 2021-July 2026). RTG: Mathematical Foundation of Data Science at University of South Carolina (DMS-2038080). PI: Linyuan Lu, University of South Carolina. Role: **Senior Personnel**. Amount awarded to RB: \$10,932.

Grant Submissions

National Institute of Health R01 (July 2023-June 2028). DREAM: Data-driven Drug Repurposing for AD/ADRD Medications. Co-PIs: Cui Tao, University of Texas Health Science Center at Houston, and Yong Chen, University of Pennsylvania. Role: **Subcontract PI**. Amount requested for RB: \$250,000.

Advising

PhD Advisees

Shijie Wang, expected 2024

Dissertation Title: New Deep Learning Approaches to Classical Statistical Problems

Zile Zhao, expected 2024

Sijian Fan, expected 2025

Other Student Supervisions

(* I mentored these students on research projects but was not their primary supervisor)

Qingyang Liu, PhD candidate in Statistics at University of South Carolina

Xin Zhi, PhD candidate in Statistics at University of South Carolina

Hung-Tien Huang, Undergraduate student in Computer Science at University of South Carolina

Liyan Xiong, PhD student in Biostatistics at University of South Carolina

Dayuan Wang, PhD student in Biostatistics at University of Florida

Anja Zgodic, PhD student in Biostatistics at University of South Carolina

Member of Doctoral Committee

Department of Statistics

Shan Zhong (PhD 2022), Zhen Yang (PhD 2022), Yang He (PhD 2023), Qingyang Liu (PhD expected 2023), Yuchen Mao (PhD expected 2023), Zehao Yu (PhD expected 2024)

¹I replaced Minsuk Shin as the Principal Investigator of this grant in May 2022.

Other Departments

Chrisogonas Odhiambo (PhD in Computer Science, 2022), Anja Zgodic (PhD in Biostatistics, 2023)

Honors and Awards

McCausland Faculty Fellow, University of South Carolina College of Arts and Sciences, August 2023-May 2026

Graduate School Fellow, University of Florida, August 2014-August 2018

Student Paper Competition Award, Section on Bayesian Statistical Science, Joint Statistical Meetings, January 2018

Travel Award, College of Liberal Arts and Sciences, University of Florida, October 2017, April 2018

Anderson Scholars Faculty Honoree, University of Florida, November 2016

Residential First-Year Experience Student Choice Award, University of Massachusetts, March 2011

Teaching

University of South Carolina

Undergraduate Courses Taught

STAT 517: Advanced Statistical Models, Fall 2022

Graduate Courses Taught

STAT 714: Linear Statistical Models, Fall 2020, Fall 2021, Fall 2022

STAT 718: High-Dimensional Data, Spring 2021, Spring 2023

STAT 721: Stochastic Processes, Spring 2022

Courses Developed and Revised

STAT 718: High-Dimensional Data

New special topics class covering statistical methodology and algorithms for high-dimensional data

STAT 721: Stochastic Processes

Course revised to place more emphasis on contemporary topics and applications such as Bayesian nonparametrics, spatial statistics, reinforcement learning, and financial mathematics

University of Florida

Courses Taught

STA 3024: Introduction to Statistics II, Spring 2016

Courses Served as Teaching Assistant

STAT 2023: Introduction to Statistics I, Fall 2015

Presentations

Invited Seminar Talks and Guest Lectures

1. South Carolina SmartState Center for Healthcare Quality, University of South Carolina, March 2023
2. Department of Statistics, University of Georgia, February 2023
3. School of Mathematics & Statistics, University of Glasgow, December 2022
4. Department of Statistics, Virginia Tech, November 2022
5. School of Mathematical and Statistical Sciences, Clemson University, October 2022
6. Department of Mathematical Sciences, University of Cincinnati, April 2022
7. Department of Biostatistics, Virginia Commonwealth University, November 2021
8. Department of Statistics, University of South Carolina, October 2021
9. Department of Statistics, University of Minnesota, October 2021
10. Department of Statistics, University of California, Davis, April 2021
11. School of Mathematical and Statistical Sciences, Arizona State University, January 2020
12. Department of Statistics, Florida State University, January 2020
13. Department of Mathematics & Statistics, San Diego State University, January 2020
14. Department of Statistics, University of California, Santa Cruz, January 2020
15. Department of Statistics, University of South Carolina, January 2020
16. Graduate Group in Genomics and Computational Biology, University of Pennsylvania, December 2019
17. Statistics Student Seminar, University of Florida, April 2019
18. PennCIL Lab, University of Pennsylvania, March 2018

Invited Conference Talks

1. 2023 Joint Statistical Meetings, Toronto, ON, Canada, August 2023
2. 6th International Conference on Econometrics and Statistics (EcoSta 2023), Tokyo, Japan, August 2023
3. Invited poster session, 2022 Joint Statistical Meetings, Washington, DC, August 2022
4. EURO 2022 Conference, Espoo, Finland, July 2022 (canceled due to illness)
5. ICSA 2022 Applied Statistics Symposium, Gainesville, FL, June 2022
6. UP-STAT 2022 Hybrid Conference, Buffalo, NY, May 2022

7. CFE-CMStatistics 2021, London, UK, December 2021
8. Fifth EAC-ISBA Conference: A Satellite Meeting of the 2020 ISBA World Meeting in Celebrating James O Berger's 70th Birthday (virtual), November 2021
9. 2021 ICSA Applied Statistics Symposium (virtual), September 2021

Contributed Talks

1. Faculty lightning talk, Department of Statistics, University of South Carolina, September 2022
2. 2021 Joint Statistical Meetings (virtual), August 2021
3. 2021 ISBA World Meeting (virtual), June 2021
4. 2019 Joint Statistical Meetings, Denver, CO, July 2019
5. 2018 Joint Statistical Meetings, Vancouver, BC, Canada, July 2018
6. Statistics Student Seminar, University of Florida, March 2018
7. Statistics Student Seminar, University of Florida, October 2017

Contributed Conference Posters

1. Bayes Comp 2020, Gainesville, FL, January 2020
2. DBEI and CCEB Research Day, University of Pennsylvania, March 2019
3. O-Bayes 2017 Meeting, Austin, TX, December 2017

Departmental Service

Member of Data Science Degree Planning Committee, 2022

Member of Graduate Committee, 2021-2022, 2022-2023

Member of Hiring Committee, 2021, 2022

Member of PhD Qualification Exam Committee, 2021

Service to the University of South Carolina

Member, SPARC Grant Review Committee, 2022

Lunch Ambassador for New Faculty Orientation, 2022

Member, Top Scholar Selection Committee, 2021-2022, 2022-2023

Professional Service

Editorial Activities

Grant reviewer for the National Science Foundation in 2020, 2021, 2022

Judge for SBSS Student Paper Competition, 2020

Journal reviewer for 39 papers: *Journal of the American Statistical Association* (5), *Journal of the Royal Statistical Society: Series B* (1), *Biometrics* (2), *Annals of Applied Statistics* (2), *Journal of Computational and Graphical Statistics* (1), *Statistica Sinica* (1), *Bernoulli* (1), *IEEE Transactions on Information Theory* (1), *BMJ Open* (1), *Journal of Statistical Planning and Inference* (2), *Scandinavian Journal of Statistics* (3), *Statistics and Computing* (1), *Statistics in Medicine* (4), *Bayesian Analysis* (1), *Journal of the Royal Statistical Society: Series C* (1), *Statistical Methods in Medical Research* (3), *The American Statistician* (2), *Computational Statistics & Data Analysis* (1), *Lifetime Data Analysis* (1), *Metrika* (1), *The R Journal* (1), *Research Synthesis Methods* (1), *Science China Mathematics* (1), *Annals of the Brazilian Academy of Sciences* (1)

Conference and Panel Activities

Session chair, "Junior advances in Bayesian treed regression," ISBA 2022 World Meeting, Montreal, QC, Canada, July 2022

Panelist for the National Science Foundation Division of Mathematical Sciences LEAPS-B Panel (virtual), March 2022

Session chair, Section on Bayesian Statistical Science, Joint Statistical Meetings in 2021, 2022

Judge for student presentation competition at the SC-ASA Palmetto Symposium (virtual), April 2021

Member of Program Committee, "Your Model is Wrong: Robustness and misspecification in probabilistic modeling," NeurIPS 2021 Workshop (virtual), December 2021

Professional Society Memberships

American Statistical Association

International Society for Bayesian Analysis

International Biometric Society

Professional Development

Certificates

New Faculty Academy Certificate of Completion, University of South Carolina, 2022

Graduate Certificate in Business Administration, Northeastern University, 2009

Workshop Attendance

Propel Research Mentorship Program, University of South Carolina, 2022-2023

Perspectives in Statistical Modeling and Inference: A Workshop in Honor of Ed George's 70th Birthday, Philadelphia, PA, 2021

Industry Experience

2012-2014 Systems Engineer, General Dynamics Mission Systems

2007-2010 Business Analyst, State Street Bank & Trust

Computer Skills

R, C/C++, Python, MATLAB, Octave, Julia, JAGS, Stan, Windows, Linux

Last updated: January 23, 2023