Ray Bai

216 LeConte College
1523 Greene St
Columbia, SC 29208

Cell: (617) 953-3266
Email: rbai@mailbox.sc.edu
Homepage: http://www.raybai.net

Research Interests

high-dimensional statistics, scalable algorithms for large and complex data sets, Bayesian methodology and theory, semiparametric and nonparametric modeling, spatial statistics, meta-analysis, analysis of electronic health records

Academic Employment

2020-   Assistant Professor, University of South Carolina, Department of Statistics

2018-2020   Postdoctoral Researcher, University of Pennsylvania, Perelman School of Medicine
             Supervisors: Yong Chen and Mary Boland

Education

2014-2018   PhD in Statistics, University of Florida
             Dissertation: Bayesian High-Dimensional Models With Scale-Mixture Shrinkage Priors
             Advisor: Malay Ghosh

2014-2016   Master of Statistics, University of Florida
             Project: Irrigation Forecasting in Southwest Florida with Generalized Additive Models
             Advisor: Nikolay Bliznyuk

2010-2012   MS in Applied Mathematics, University of Massachusetts Amherst

2003-2007   BA in Economics and Government, Cornell University

Honors and Awards

Graduate School Fellowship, 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

Professional Society Memberships

American Statistical Association

International Society for Bayesian Analysis

International Biometric Society

Publications

__ = *student advised by RB*
† = *co-first author*
∗ = *alphabetical order*

Preprints

1. Deshpande, S. K., **Bai, R.**, Balocchi, C., Starling, J. E., and Weiss, J. (2021+). VCBART: Bayesian trees for varying coefficients. Invited revision at *Journal of the American Statistical Association*.

2. **Bai, R.** (2021+). A unified computational and theoretical framework for high-dimensional Bayesian additive models. Invited revision at *Statistica Sinica*.

3. **Bai, R.**, Lin, L., Boland, M. R., and Chen, Y. (2021+). A robust Bayesian Copas selection model for quantifying and correcting publication bias. Under review at *Annals of Applied Statistics*.

4. **Bai, R.**, Boland, M. R., and Chen, Y. (2021+). Fast algorithms and theory for high-dimensional Bayesian varying coefficient models. Under review at *Journal of Machine Learning Research*.

5. Meeker, J. R., Burris, H. H., **Bai, R.**, Levine, L. D., and Boland, M. R. (2021+). Neighborhood deprivation increases the risk of post-induction cesarean delivery. Under review at *Obstetrics & Gynecology*.

6. Balocchi, C.†, **Bai, R.**†, Liu, J., Canelón, S., George, E. I., Chen, Y., and Boland, M. R. (2021+). A Bayesian hierarchical modeling framework for geospatial analysis of adverse pregnancy outcomes.

Peer-Reviewed Articles

7. **Bai, R.** and Ghosh, M. (2021). On the beta prime prior for scale parameters in high-dimensional Bayesian regression models. *Statistica Sinica*, **31**.

8. 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* (in press).

9. Boland, M. R., Liu, J., Balocchi, C., Meeker, J., **Bai, R.**, Mowery, D., and Herman, D. (2021). A method to link neighborhood-level covariates to COVID-19 infection patterns in Philadelphia using spatial regression. *AMIA 2021 Virtual Informatics Summit* (in press).

10. **Bai, R.**†, Moran, G. E.†, Antonelli, J. L.†, Chen, Y., and Boland, M. R. (2020+). Spike-and-slab group lassos for grouped regression and sparse generalized additive models. *Journal of the American Statistical Association* (in press).

11. **Bai, R.** and Ghosh, M. (2019). Large-scale multiple hypothesis testing with the normal-beta prime prior. *Statistics*, **53**: 1210-1233.

12. **Bai, R.** and Ghosh, M. (2018). High-dimensional multivariate posterior consistency under global-local shrinkage priors. *Journal of Multivariate Analysis*, **167**: 157-170.

13. Duerr, I., Merrill, H. R., Wang, C., **Bai, R.**, Boyer, M. J., 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.

Book Chapters

14. **Bai, R.**, Ročková, V., and George, E. I. (2021+). Spike-and-slab meets LASSO: A review of the spike-and-slab LASSO. *Handbook of Bayesian Variable Selection*, Tadesse, M. and Vannucci, M. eds. Chapman & Hall/CRC Press (in press).

Selected Works in Progress

Bai, R., Jeong, S., and Ročková, V. (2021+). Minimax rates and adaptive procedures for generalized nonparametric regression.

Bai, R. (2021+). Uncertainty quantification for Bayesian vector autoregressive models.

Bai, R., Liu, X., Lin, L., Chu, H., and Chen, Y. (2021+). ABSORB: A Bayesian Selection model for correcting and quantifying Outcome Reporting Bias in multivariate meta-analysis.

Bai, R.[*] and Qin, Q.[*] (2021+). Analysis of MCMC algorithms for Gaussian process regression with automatic relevance determination kernels.

Students Advised

PhD Students

Zile Zhao (in progress). Studying high-dimensional regression and variational Bayes machine learning algorithms

Member of PhD committee

Shan Zhong (in progress)

Grants Funded

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.

National Science Foundation (September 2020-August 2022). RII Track 1: Materials Assembly and Design Excellence in South Carolina: MADE in SC (OIA-1655740). PI: Prakash Nagarkatti, University of South Carolina. Role: Co-PI. Amount awarded to RB: $30,000.

Teaching Experience

University of South Carolina

STAT 718: High-Dimensional Data (special topics course), Spring 2021

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

University of Florida

STA 3024: Introduction to Statistics II, Spring 2016

Presentations

Invited Talks

1. Seminar, Department of Statistics, University of California, Davis, April 2021

2. Seminar, School of Mathematical and Statistical Sciences, Arizona State University, January 2020

3. Seminar, Department of Statistics, Florida State University, January 2020

4. Seminar, Department of Mathematics & Statistics, San Diego State University, January 2020

5. Seminar, Department of Statistics, University of California, Santa Cruz, January 2020

6. Seminar, Department of Statistics, University of South Carolina, January 2020

7. Guest lecture, GCB533: Statistics for Genomics and Biomedical Informatics, University of Pennsylvania, December 2019

8. Statistics Student Seminar, University of Florida, April 2019

9. Invited talk, Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, March 2018

Contributed Talks

1. Joint Statistical Meetings, Denver, CO, July 2019

2. Joint Statistical Meetings, Vancouver, BC, Canada, July 2018

3. Statistics Student Seminar, University of Florida, March 2018

4. 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

University of South Carolina

Member of the PhD Qualification Exam Committee, 2021

Member of the Hiring Committee, 2021

University of Florida

Organizer of Statistics Student Seminar Series, 2016-2017

Professional Service

Grant reviewer for the National Science Foundation

Journal reviewer for: *The American Statistician* (1), *Annals of the Brazilian Academy of Sciences* (1), *Computational Statistics & Data Analysis* (1), *Journal of Statistical Planning and Inference* (2), *Journal of the American Statistical Association* (3), *Lifetime Data Analysis* (1), *Research Synthesis Methods* (1), *Scandanavian Journal of Statistics* (1), *Science China Mathematics* (1), *Statistical Methods in Medical Research* (1)

Judge for 2020 SBSS Student Paper Competition

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: February 26, 2021