# Lu Zhang

*Curriculum Vitae*

*Division of Biostatistics*
*Department of Population and Public Health Sciences*
*University of Southern California*
*1845 N. Soto Street, Los Angeles, CA, 90089-9239, USA*
☎ *+1 (313) 442 1812*
✉ *lzhang63@usc.edu*
*Homepage: https://luzhangstat.github.io/*

## Employment

2022–current **Assistant Professor**, *University of Southern California*, USA.

2020–2022 **Postdoctoral Researcher**, *Columbia University*, USA.
Supervisor: Bob Carpenter, Andrew Gelman

## Education

2014–2020 **Ph.D. in Biostatistics**, *University of California, Los Angeles*, USA.
Advisor: Sudipto Banerjee

2010–2014 **B.S. in Mathematics and Applied Mathematics**, *Fudan University*, China.

## Research Interests

Spatial analysis, Bayesian statistics, high dimensional inference, computational statistics and open-source software development

## Papers (* co-first author, † students mentored by me)

### Publications and Accepted Manuscripts

1. Shengjie Liu†, **Lu Zhang**. Deep Feature Gaussian Processes for Single-Scene Aerosol Optical Depth Reconstruction. (2024) *Accepted by* ***IEEE Geoscience and Remote Sensing Letters***
2. **Lu Zhang***, Wenpin Tang*, Sudipto Banerjee, Fixed-Domain Asymptotics Under Vecchia's Approximation of Spatial Process Likelihoods (2023). ***Statistica Sinica***. `http://www3.stat.sinica.edu.tw/ss_newpaper/SS-2021-0428_na.pdf`
3. **Lu Zhang**, Bob Carpenter, Andrew Gelman, Aki Vehtari (2022). Pathfinder: Parallel quasi-Newton variational inference. ***Journal of Machine Learning Research***. . `https://www.jmlr.org/papers/volume23/21-0889/21-0889.pdf`
4. **Lu Zhang** (2022). Applications of Conjugate Gradient in Bayesian computation. ***Wiley StatsRef-Statistics Reference Online.*** `https://doi.org/10.1002/9781118445112.stat08411`
5. Wenpin Tang*, **Lu Zhang***, Sudipto Banerjee (2021). On identifiability and consistency of the nugget in Gaussian spatial process models. ***Journal of the Royal Statistical Society Series B***, *(Statistical Methodology)*, `https://rss.onlinelibrary.wiley.com/doi/10.1111/rssb.12472`

6. **Lu Zhang**, Sudipto Banerjee, (2021) Spatial Factor Modeling: A Bayesian Matrix-Normal Approach for Misaligned Data. ***Biometrics***, *78(2), 560-573.*. http://doi.org/10.1111/biom.13452
7. **Lu Zhang**, Sudipto Banerjee, Andrew O. Finley (2021). High-dimensional multivariate geostatistics: A Bayesian matrix-normal approach. ***Environmetrics***, *32(4), e2675.* **Selected for the 2021 Wiley-TIES Best Paper Award** https://onlinelibrary.wiley.com/doi/10.1002/env.2675
8. Gregory L. Watson, Di Xiong, **Lu Zhang**, Joseph A. Zoller, John Shamshoian, Phillip Sundin, Teresa Bufford, Anne W. Rimoin, Marc A. Suchard, Christina M. Ramirez (2021). Pandemic velocity: forecasting COVID-19 in the US with a machine learning & Bayesian time series compartmental model. ***PLOS Computational Biology***, 17(3), e1008837.
9. Di Xiong*, **Lu Zhang***, Gregory L. Watson, Phillip Sundin, Teresa Bufford, Joseph A. Zoller, John Shamshoian, Marc A. Suchard, Christina M. Ramirez, (2020). Pseudo-likelihood based logistic regression for estimating COVID-19 infection and case fatality rates by gender, race, and age in California. ***Epidemics***, *33, 100418.* https://www.sciencedirect.com/science/article/pii/S1755436520300396
10. **Lu Zhang**, Abhirup Datta, Sudipto Banerjee. (2019). Practical Bayesian modeling and inference for massive spatial data sets on modest computing environments. ***Statistical Analysis and Data Mining: The ASA Data Science Journal***, *12(3), 197-209.* https://onlinelibrary.wiley.com/doi/full/10.1002/sam.11413

### Preprints

11. **Lu Zhang**, Wenpin Tang, Sudipto Banerjee. Bayesian Geostatistics Using Predictive Stacking, https://arxiv.org/abs/2304.12414
12. Shane Sparkes†, **Lu Zhang**. Properties and Deviations of Random Sums of Densely Dependent Random Variables, https://arxiv.org/abs/2310.11554
13. **Lu Zhang**, Andrew Finley, Arne Nothdurft, Sudipto Banerjee. Bayesian Modeling of Incompatible Spatial Data: A Case Study Involving Post-Adrian Storm Forest Damage Assessment, https://arxiv.org/abs/2311.11256
14. Shane Sparkes†, Erika Garcia, **Lu Zhang**. The functional average treatment effect, https://arxiv.org/abs/2312.00219
15. Soumyakanti Pan, **Lu Zhang**, Jonathan R. Bradley, Sudipto Banerjee. Bayesian Inference for Spatial-temporal Non-Gaussian Data Using Predictive Stacking, hhttps://arxiv.org/abs/2406.04655

## Packages

1. **Lu Zhang** and Jun Yin (2018). *phase1PRMD: Personalized Repeated Measurement Design for Phase I Clinical Trials. R package version 1.0.2. CRAN:* https://cran.r-project.org/web/packages/phase1PRMD/index.html
2. Xiang Chen, **Lu Zhang**, Sudipto Banerjee (2018). *JAMAJniLite: A JAVA package providing a java interface for lapack and blas libraries and using the classes defined by JAMA Package Github:* https://github.com/JAMAJni/JAMAJniLite

3. **Lu Zhang**, LiZhen Nie, Sudipto Banerjee (2017). *JALAJni: A JAVA package providing a java interface for lapack and blas library Github:* `https://github.com/JaLAJni/JaLAJni`

## Grants

03/2023-03/2024 **Principal Investigator**, *Assessing Particulate Matter Exposures Based On Multi-Source Satellite Data Using Scalable Gaussian Process Models*, Southern California Environmental Health Sciences Center (P30ES007048) Pilot Projects Program. Total Direct Costs $47,500

## Teaching Experience

### Instructor at USC

Summer 2024 PM 511a: **Data Analysis (a)**

Fall 2023 PM 569: **Spatial Statistics**

### Graduate Teaching Assistant at UCLA

2015-2020 Biostat 100A: **Introduction to Biostatistics**
(Summer 2015, Fall 2015, Spring 2016, Summer 2017, Fall 2019)

2016-2020 Biostat 100B: **Introduction to Biostatistics**
(Winter 2016, Winter 2017, Winter 2018, Winter 2020)

Fall 2016 Biostat 200A: **Basic Biostatistics**

Spring 2017 Biostat 411: **Analysis of Correlated Data**

Fall 2017 Biostat 255A: **Advanced Topics & Probability in Biostatistics**

Winter 2017 Biostat 255B: **Advanced Topics & Probability in Biostatistics**

Spring 2018 Biostat 257: **Statistical Computing**

Spring 2019 Biostat 241: **Spatial modeling**

Fall 2019 Public Health 200: **Foundations in Public Health**

Spring 2020 Biostat 214: **Finite Population Sampling**

## Working Experience

Jun. - Sep. 2018 **Internship in Biostatistics**, *Mayo Clinic*, Rochester, Minnesota USA,
Sponsor: Yin Jun, Ph.D.

- Statistical consultation to Physicians
- Experimental design (clinical trial design)
- Software development (develop R package)

## Selected Awards

2020 **Dean's Outstanding Student Award in Biostatistics**, Department of Biotatistics, UCLA

2018 **Celia G. and Joseph G. Blann Fellowship**, Department of Biotatistics, UCLA

2016 **Graduate Summer Research Mentorship,** Department of Biotatistics, UCLA

## Talks

### Invited

May. 2024 **Bayesian Geostatistics Using Predictive Stacking**.
NESS, Storrs, Connecticut, USA

May. 2023 **Pathfinder: A Parallel Quasi-Newton Algorithm for Reaching Regions of High Probability Mass**.
IRSA's 2023 conference, Institute for Research in Statistics and its Applications at University of Minnesota, Minneapolis, MN, USA

Apr. 2023 **Bayesian inference for high-dimensional latent spatial model: Why we should and how to avoid random walk in MCMC**.
DMS Colloquium, Department of Mathematics and Statistics at Auburn University, Auburn, AL, USA

Mar. 2023 **Pathfinder: A Parallel Quasi-Newton Algorithm for Reaching Regions of High Probability Mass**.
Bayes Comp 2023, Levi, Finland

Jan. 2023 **Bayesian inference for high-dimensional latent spatial model: Why we should and how to avoid random walk in MCMC**.
Purdue Research Colloquium, Statistics at Purdue University, West Lafayette, IN, USA

Sep. 2022 **Spatial Factor Modeling: A Bayesian Matrix-Normal Approach for Massive Spatial Data with Missing Observations**.
SIAM, Conference on Mathematics of Data Science, San Diego, CA, USA

Aug. 2022 **Spatial Factor Modeling: A Bayesian Matrix-Normal Approach for Massive Spatial Data with Missing Observations**.
JSM, Washington, DC, USA

Apr. 2022 **Pathfinder: A parallel quasi-Newton algorithm for reaching regions of high probability mass**.
SIAM Conference on Uncertainty Quantification (UQ22), Atlanta, Georgia, U.S.

Nov. 2021 **Pathfinder: Parallel quasi-Newton variational inference**.
Broad Institute, Remote

Sep. 2021 **Spatial Factor Modeling: A Bayesian Matrix-Normal Approach for Massive Spatial Data with Missing Observations**.
Mathematics and Applied Mathematics at Fudan University, Shanghai, China

Sep. 2021 **Spatial Factor Modeling: A Bayesian Matrix-Normal Approach for Massive Spatial Data with Missing Observations**.
School of Statistics and Management at Shanghai University of Finance and Economics, Shanghai, China

Jun. 2021 **Spatial Factor Modeling: A Bayesian Matrix-Normal Approach for Massive Spatial Data with Missing Observations**.
Biostatistics at Columbia University, New York, New York, USA

Dec. 2020 **Spatial Factor Modeling: A Bayesian Matrix-Normal Approach for Misaligned Data**.
Johns Hopkins University BLAST working group, Baltimore, Maryland, USA

Mar. 2020 **High-dimensional Multivariate Geostatistics: A Bayesian Matrix-Normal Approach**.
ENAR, Nashville, Tennessee, USA

### Contributed

Aug. 2022 **Bayesian Predictive Stacking Under Spatial Process Settings**.
IMSI Workshop, Chicago, IL, USA

Aug. 2021 **Pathfinder: A Parallel Quasi-Newton Algorithm for Reaching Regions of High Probability Mass**.
Joint Statistical Meetings

Aug. 2020 **Spatial Factor Modeling: A Bayesian Matrix-Normal Approach for Misaligned Data**.
Bernoulli-IMS One World Symposium 2020

Jul. 2019 **Bayesian Linear Model of Coregionalization (BLMC) for Large Scale Datasets with Accelerated Posterior Sampling Algorithm**.
Joint Statistical Meetings, Colorado, USA, poster presentation

Aug. 2017 **Practical Bayesian Inference Based on Nearest Neighbor Gaussian Processes Model for Massive Spatial Data**.
Joint Statistical Meetings, Baltimore, Maryland, USA

## Referee Experience

Journal of the Royal Statistical Society: Series B (1)
Biometrics (1)
Journal of Machine Learning (1)
Journal of Computational and Graphical Statistics (4)
Annals of Applied Statistics (1)
Bayesian Analysis (1)
Nature Communications (1)
Statistical Science (1)
Environmetrics (1)

## Professional Memberships

American Statistical Association
Eastern North American Region
International Chinese Statistical Association