

Abhirup Datta

Johns Hopkins Bloomberg School of Public Health
615 N Wolfe Street, Baltimore, MD 21205

☎ (410) 502 2988
✉ abhidatta@jhu.edu
<http://abhidatta.com>

Academic Appointments

2016– Present Assistant Professor, *Department of Biostatistics*, Johns Hopkins Bloomberg School of Public Health.

Education

2012–2016 Ph.D., Biostatistics, *University of Minnesota*, Minneapolis.
Advisers: [Dr. Sudipto Banerjee](#) and [Dr. Hui Zou](#)

2008–2010 M.Stat., *Indian Statistical Institute*, Kolkata, India.
Specialization in Math-Stat-Probability

2005–2008 B.Stat., *Indian Statistical Institute*, Kolkata, India.

Research Interests

Air pollution, Bayesian statistics, Biostatistics, High dimensional data, Ecology, Environmental health, Epidemiology, Small area estimation, Spatial statistics

Papers

* indicates equal contributions

† indicates corresponding author or project leader

indicates first author is a student advisee

1. **Datta, A.** and Lin, W. and Rao, A. and Diouf, D. and Kouame, A. and Edwards, J. K. and Bao, L. and Louis, T. A. and Baral, S., *Bayesian estimation of MSM population in Côte d'Ivoire*, *Statistics and Public Policy* (Accepted)
2. Taylor-Rodriguez, D., Finley, A. O., **Datta, A.**, Babcock, C., Andersen, H., Cook, B. D., Morton, D. C., and Banerjee, S. *Spatial Factor Models for High-Dimensional and Large Spatial Data: An Application in Forest Variable Mapping*, *Statistica Sinica* (Accepted)
3. **Datta, A.**, Zou, H. and Banerjee, S. *Bayesian Inference for High-dimensional Changing Linear Regression with Application to Minnesota House Price Index Data*, *Statistics and Its Interface* (Accepted)
4. Finley, A.O., **Datta, A.**, Cook, B.C. , Morton, D.C., Andersen, H.E. and Banerjee, S. *Efficient algorithms for Bayesian Nearest Neighbor Gaussian Processes*, *Journal of Computation and Graphical Statistics* (Accepted)
5. Edwards, J. K., Hileman, S., Donastorg, Y., Sanchez, R., Zadrozny, S., Baral, S., Hargreaves, J., Fearon, E., Zhao, J., **Datta, A.** and Weir, S. S., [Estimating sizes of key populations](#)

- at the national level: considerations for study design and analysis*, *Epidemiology*, vol. 29, no. 6, pp. 795–803, 2018
6. #† Saha, A. and **Datta, A.**, *BRISC: Bootstrap for rapid inference on spatial covariances* *Stat*, e184. 2018
 7. **Datta, A.** and Zou, H. *CoCoLasso for High-dimensional Error-in-variables Regression* *Annals of Statistics*, vol. 45, no. 6, pp. 2400-2426, 2017.
 8. * Butler, E. E. and * **Datta, A.** and Flores-Moreno, H. and Chen, M. and Wythers, K. R. and Fazayeli, F. and Banerjee, A. and Atkin, O. K. and Kattge, J. and Amiaud, B. and Blonder, B. and Boenisch, G. and Bond-Lamberty, B. and Brown, K. A. and Byun, C. and Campetella, G. and Cerabolini, B. E. L. and Cornelissen, J. H. C. and Craine, J. M. and Craven, D. and de Vries, F. T. and Diaz, S. and Domingues, T. F. and Forey, E. and Gonzalez-Melo, A. and Gross, N. and Han, W. and Hattingh, W. N. and Hickler, T. and Jansen, S. and Kramer, K. and Kraft, N. J. B. and Kurokawa, H. and Laughlin, D. C. and Meir, P. and Minden, V. and Niinemets, U. and Onoda, Y. and Penuelas, J. and Read, Q. and Sack, L. and Schamp, B. and Soudzilovskaia, N. A. and Spasojevic, M. J. and Sosinski, E. and Thornton, P. E. and Valladares, F. and van Bodegom, P. M. and Williams, M. and Wirth, C. and Reich, P. B., *Mapping local and global variability in plant trait distributions* *Proceedings of the National Academy of Sciences*, vol. 114, no. 51, pp. E10937–E10946, 2017.
 9. **Datta, A.**, Banerjee, S., Finley, A.O., Hamm, N.A.S. and Schaap, M. *Non-separable Dynamic Nearest Neighbor Gaussian Process Models for Large Spatio-temporal Data with Application to Particulate Matter Analysis*, *Annals of Applied Statistics*, vol. 10, no. 3, pp. 128-1316, 2016.
 10. **Datta, A.**, Banerjee, S., Finley, A.O. and Gelfand, A.E. *On nearest-neighbor Gaussian process models for massive spatial data*, *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 8, no. 5, pp. 162-171, 2016.
 11. **Datta, A.**, Banerjee, S., Finley, A.O. and Gelfand, A.E. *Hierarchical Nearest Neighbor Gaussian Process models for Large Geostatistical Datasets*, *Journal of the American Statistical Association*, vol. 111, no. 514, pp. 800-812, 2016.
 12. Zhang, L., **Datta, A.**, and Banerjee, S. (Under review) *Practical Bayesian Inference for Massive Spatial Data on Modest Computing Environments*, [\[ArXiv\]](#)
 13. Heaton, M. J. and **Datta, A.** and Finley, A. and Furrer, R. and Guhaniyogi, R. and Gerber, F. and Gramacy, R. B. and Hammerling, D. and Katzfuss, M. and Lindgren, F. and Nychka, D. W. and Sun, F. and Zammit-Mangion, A. (Under review), *Methods for Analyzing Large Spatial Data: A Review and Comparison*, [\[ArXiv\]](#)
 14. **Datta, A.**, Banerjee, S. and Hodges, J. S. (Under review) *Spatial disease mapping using Directed Acyclic Graph Auto-Regressive (DAGAR) models* [\[ArXiv\]](#)

15. Flores-Moreno, H., Fazayeli, F., Banerjee, A., **Datta, A.**, Kattge, J., Butler, E., Atkin, O., Wythers, K., Chen, M., Anand, M., Bahn, M., Burrascano, S., Byun, C., Cornelissen, J., Craine, J., Gonzalez-Melo, A., Hattingh, W., Jansen, S., Kraft, N., Kramer, K., Laughlin, D., Minden, V., Niinemets, U., Onipchenko, V., Penuelas, J., Soudzilovskaia, N. and Reich, P. (Under review) *Robustness of trait connections between multiple plant organs across environmental gradients and growth forms*
16. **Datta, A.**, Fiksel, J., Amouzou, A. and Zeger, S. (Under review) *Local calibration of verbal autopsy algorithms* [\[ArXiv\]](#)

Software

- R-package *calibratedVA: Locally calibrated cause specific mortality fractions using verbal autopsy data*, Fiksel J., **Datta, A.**, (2018).
<https://github.com/jfiksel/CalibratedVA>
- R-package *spNNGP: Spatial Regression Models for Large Datasets using Nearest Neighbor Gaussian Processes*, Finley A.O., **Datta, A.**, Banerjee, S. (2017), R package version 0.1.0.
<https://CRAN.R-project.org/package=spNNGP>
- R-package *BRISC: Bootstrap for rapid inference on spatial covariances*, Saha, A., **Datta, A.**, Nosedal, J., Okazaki, N. (2018), R package version 0.1.0.
<https://CRAN.R-project.org/package=BRISC>

Funding

- July 2018 to June 2019 Statistical Maps of Air Quality in Baltimore City Using Low-Cost Monitoring Data, *Bloomberg American Health Initiative Spark Award*, Role: PI (Co-PI: Kirsten Koehler), \$71,000 (including 20% effort).
- Oct 2016 to Mar 2019 Project SOAR (Supporting Operational AIDS Research), *USAID SH1420* (PI: Deanna Kerrigan), Role: Statistical Consultant, 15% effort.
- Jan 2017 to Dec 2019 Comprehensive Mortality Surveillance for Action (COMSA), *Bill & Melinda Gates Foundation* (PI: Agbessi Amouzou), Role: Col, 20% effort.
- 2018 to 2022 NIH 1R01EB026549-01, (PI: Martin Lindquist), Role: Col, 10% effort.
- Aug 2018 - July 2019 NIH 1R21DA046188-01A1, PI: Saonli Basu), Role: PI on subaward, 10% effort.

Awards

- 2018 Top 1% of reviewers in Mathematics, *Publons*.
- 2018 Honorable mention: Savage Award Applied Methodology, *International Society for Bayesian Analysis (ISBA)*.

- 2018 Excellence in Teaching, *Johns Hopkins Bloomberg School of Public Health*, Fourth Quarter, Probability IV.
- 2017 ASA Outstanding Statistical Application Award, *American Statistical Association*.
- 2017 Top 1% of reviewers in Mathematics, *Publons*.
- 2016 ENAR Distinguished Student Paper Award, *International Biometric Society*, Austin, TX.
- 2016 Delta Omega Honorary Society Student Inductee (Pi Chapter), Minneapolis, MN.
- 2016 Best Student Seminar Presentation Award, *Division of Biostatistics*, University of Minnesota, Minneapolis, MN.
- 2015 Inter-disciplinary Doctoral Fellowship 2015-16, *Division of Biostatistics*, University of Minnesota Graduate School, Minneapolis, MN.
- 2014 JSM Student Paper Award, *American Statistical Association, Section on Bayesian Statistical Science*, Boston, MA.
Also selected for best paper award in Statistics and the Environment Section
- 2014 Outstanding Teaching Assistant Award, *Division of Biostatistics, University of Minnesota*, Minneapolis, MN.
- 2014 Best Paper Award, *Division of Biostatistics, University of Minnesota*, Minneapolis, MN.

Conference Travel Awards

- 2018 ISBA World Meeting, *International Society for Bayesian Analysis*, Edinburgh, UK.
- 2015 G70 Conference, *Duke University*, Durham, NC.
- 2014 ISBA World Meeting, *International Society for Bayesian Analysis*, Cancun, Mexico.
- 2014 Conference on Non-parametric Statistics for Big Data and Celebration to Honor Professor Grace Wahba, *University of Wisconsin*, Madison, WI.
- 2014 Pan-American Advanced Study Institute on Spatio-Temporal Statistics, *Travel grant from National Science Foundation*, Buzios, Brazil.

Teaching

Thesis advisor

- In progress Arkajyoti Saha, *Ph.D. student in Department of Biostatistics*, Co-advised with Nilanjan Chatterjee.
- In progress Andrew Pita, *Sc.M. student in Department of Biostatistics*.

Final Oral Examination / Thesis Reader

- 2018 Josh Colston, *Department of International Health*.

Short Course

July 2017 Full day short course on Bayesian models for high dimensional spatial data, *Joint Statistical Meetings*, Baltimore, MD, [Course webpage](#).

Courses at Hopkins

Mar-May, 2018 Probability Theory IV.

May, 2018 Guest lecture in Environmental Epidemiology: Time Series.

Jan-Mar, 2018 Special topics: Advanced spatial statistics, [Github page](#).

Mar-May 2017 Special topics: Scalable methods for large spatial data, [Github page](#).

Teaching Assistant, Division of Biostatistics, University of Minnesota

Spring 2015 Advanced Statistical Inference.
Instructors: Dr. Cavan Reilly and Dr. David Vock

Fall 2014 Probability Models for Biostatistics.
Instructor: Dr. Baolin Wu

Spring 2014 Bayes Decision Theory and Data Analysis.
Instructor: Dr. Sudipto Banerjee

Spring 2014 Advanced Statistical Inference.
Instructors: Dr. Julian Wolfson and Dr. David Vock

Fall 2013 Probability Models for Biostatistics.
Instructor: Dr. Baolin Wu

Spring 2012 Advanced Regression.
Instructor: John Hughes, Ph.D

Fall 2012 Statistical Methods for Correlated Data.
Instructor: Dr. Julian Wolfson

Professional Activities

Reviewer for *Advances in Statistical Climatology, Meteorology and Oceanography*
Annals of Applied Statistics
Biometrics
Computational Statistics and Data Analysis (CSDA)
Environmetrics
Journal of the American Statistical Association (JASA)
Journal of Computation and Graphical Statistics (JCGS)
Journal of Multivariate Analysis (JMVA)
Journal of the Royal Statistical Society Series C (JRSSC)

Sankhya A
Spatial Statistics
Statistica Sinica
Statistics in Medicine
Transactions on Pattern Analysis and Machine Intelligence

Other services *Faculty search committee, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 2017-18*

Seminar organizer, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 2017

Session organizer "Statistical Advancements in Forestry, Ecology and Climate Modeling", Joint Statistical Meetings, 2018, Vancouver, CA

Session Chair, Joint Statistical Meetings, 2014, Boston, MA

Invited Presentations

- Aug 2018 *Joint Statistical Meetings, Vancouver, Canada.*
- Jun 2018 *ISBA World Meeting, Edinburgh, UK.*
- Mar 2018 *Eastern North American Region Meetings (ENAR), International Biometric Society, Atlanta, GA.*
- Dec 2017 *International Indian Statistical Association Conference, Hyderabad, India.*
- Dec 2017 *10th International Conference of the ERCIM WG on Computational and Methodological Statistics, London, UK.*
- Nov 2017 *American Public Health Association Annual Meeting, Atlanta, GA.*
- Oct 2017 *UNAIDS Reference Group Fall Meeting 16-18 October 2017, London, UK.*
- Feb 2017 *Department of Mathematics and Statistics, University of Maryland , Baltimore County, MD.*
- Feb 2017 *CDC Consultation Conference on Key Populations, CDC, Atlanta, GA.*
- Dec 2016 *Platinum Jubilee International Conference on Applications of Statistics, Calcutta University, Kolkata, India.*
- Nov 2016 *President's Emergency Plan for Aids Relief, Washington DC.*
- Feb 2016 *Department of Statistical Science, Duke University, Durham, NC.*
- Feb 2016 *Department of Biostatistics, University of Michigan, Ann Arbor, MI.*
- Feb 2016 *Department of Biostatistics, Johns Hopkins University, Baltimore, MD.*
- Feb 2016 *Department of Statistics, University of California, Irvine, CA.*
- Feb 2016 *Department of Biostatistics, University of North Carolina, Chapel Hill, NC.*
- Jan 2016 *Division of Biostatistics Student Seminar, University of Minnesota, Minneapolis, MN.*
- Dec 2015 *International Indian Statistical Association Conference, Pune, India.*

Aug 2010 Mahalanobis International Symposium on Statistics, *Kolkata, India*.

Contributed Presentations

July 2017 IMS New Researcher's Conference, *Baltimore, MD*.

July 2017 Spatial Statistics Conference, *Lancaster, UK*.

Aug 2016 Joint Statistical Meetings, *Chicago, IL*.

Mar 2016 Eastern North American Region Meetings (ENAR), International Biometric Society, *Austin, TX*.

Dec 2015 2015 Dow Sustainability Innovation Student Challenge Award (SISCA), *Minneapolis, MN*.

Mar 2015 Eastern North American Region Meetings (ENAR), International Biometric Society, *Miami, FL*.

Aug 2014 Joint Statistical Meetings, *Boston, MA*.

Jun 2014 Pan-American Advanced Study Institute on Spatio-Temporal Statistics, *Buzios, Brazil*.

Mar 2014 Eastern North American Region Meetings (ENAR), International Biometric Society, *Baltimore, MD*.

Poster Presentations

Apr 2015 G70: A Celebration of Alan Gelfand's 70th Birthday, *Durham, NC*.

Apr 2015 University of Minnesota School of Public Health Annual Research Day, *Minneapolis, MN*.

Jul 2014 ISBA World Meeting, *Cancun, Mexico*.

Jun 2014 Conference on Nonparametric Statistics for Big Data and Celebration to Honor Professor Grace Wahba, *Madison, WI*.

Apr 2014 University of Minnesota School of Public Health Annual Research Day, *Minneapolis, MN*.

Oct 2013 University of Minnesota U-Spatial Symposium, *Minneapolis, MN*.

Other Presentations at Hopkins

Nov 2017 Coherent extrapolation of linear regression using kriging, *Small Area Estimation and Spatial Statistics Working Group*.

Nov 2017 Overview of spatial statistics, *Small Area Estimation and Spatial Statistics Working Group*.

Apr 2017 Spatial disease mapping using Directed Acyclic Graph Autoregressive (DAGAR) models, *Small Area Estimation and Spatial Statistics Working Group*.

Mar 2017 Spatial Analysis with Big and Small Data, *Department of Biostatistics*.

Dec 2016 Large Scale Spatio-temporal Interpolation of Environmental Pollutants, *Department of Biostatistics*.

- Nov 2016 Approaches to population size estimation for key populations, *Small Area Estimation and Spatial Statistics Working Group*.
- Sept 2016 Bayesian Models for High Dimensional Spatial Data, *Department of Biostatistics*.
- Sept 2016 Bayesian Models for High Dimensional Spatial Data, *Small Area Estimation and Spatial Statistics Working Group*.

Other Research Experience

- 2015–2016 Inter-disciplinary Doctoral Dissertation Fellow, *Institute on the Environment, University of Minnesota*, Minneapolis, MN.
Mentor: Dr. Arindam Banerjee
- 2012–2015 Research Assistant, *Division of Biostatistics, University of Minnesota*, Minneapolis, MN.
Supervisor: Dr. Sudipto Banerjee
- Summers of 2015 & 2014 Research Consultant, *Geospatial Lab, Michigan State University*.
Supervisor: Dr. Andrew O. Finley
- Summer 2007 Visiting Student Research Program (Mathematics), *Tata Institute of Fundamental Research*, Mumbai, India.

Industry Appointments

- 2011–2012 Quantitative Analyst, *Morgan Stanley*, New York, NY.
- 2010–2011 Quantitative Analyst, *Morgan Stanley*, Mumbai, India.
- Summer 2009 Quantitative Analyst Intern, *Morgan Stanley*, Mumbai, India.
- Summer 2008 Analytics Intern, *SYSTAT Software Asia Pacific Ltd*, Bangalore, India.

Professional Memberships

- American Statistical Association
- Eastern North American Region International Biometric Society
- International Indian Statistical Association