





Debangon Dey, Ph.D.




✉ debangan@stat.tamu.edu  Ddey07

 <https://debangandey.rbind.io/>






Research Interests

- Summary  A statistician focused on developing methods to analyze intensive multilevel multimodal longitudinal data and big geospatial data. Applications focus on Digital Health Technologies (DHTs) and Environmental Sciences.
- Areas  Functional data analysis, longitudinal data, spatial statistics, stochastic processes, semiparametric statistics.

Education




- 2017 – 2022  **Ph.D., Johns Hopkins Bloomberg School of Public Health** in Biostatistics.
Thesis title: *Topics in Modeling of Multivariate Mixed Data Types and Highly Multivariate Spatial Data*
- 2015 – 2017  **Masters, Indian Statistical Institute** in Statistics.
Thesis title: *On Investigation of Genotypic Association with Longitudinal Observations of Traits.*
- 2012 – 2015  **Bachelors, Indian Statistical Institute** in Statistics.

Professional Experience

- 07/2025 – present  **Assistant Professor**, Department of Statistics, Texas A&M University.
- 05/2022 – 05/2025  **Postdoctoral Fellow**, National Institute of Mental Health (NIMH), NIH. Lead analyst for the Motor Activity Research Consortium for Health (mMARCH). Advisor: Dr. Kathleen R. Merikangas.
- 05/2021 – 08/2021  **Applied Scientist Intern**, Amazon. Developed heterogeneous treatment effect methods to reduce A/B testing trial length in the Amazon Search team.
- 2020  **Analyst**, Baltimore City Health Department – Johns Hopkins Data Analysis Support Team for COVID-19 Response.
- 2015, 2016  **Summer Intern**, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health.

Research Publications and Repositories


Journal Articles

-  **D. Dey**, S. Banerjee, M. A. Lindquist, and A. Datta, “Graph-constrained analysis for multivariate functional data,” *Journal of Multivariate Analysis*, p. 105 428, 2025.
-  **D. Dey**, H. A. Lateef, A. Leroux, V. Zipunnikov, and K. Merikangas, “Associations between daily outdoor temperature and subjective real-time ratings of emotional states and sleep in mood disorder subtypes,” *Journal of Affective Disorders*, p. 120 918, 2025.
-  **D. Dey**, R. Ghosal, K. Merikangas, and V. Zipunnikov, “Functional principal component analysis for continuous non-gaussian, truncated, and discrete functional data,” *Statistics in Medicine*, 2024.

- 4 T. M. Lateef, **D. Dey**, A. Leroux, L. Cui, M. Xiao, V. Zipunnikov, and K. R. Merikangas, "Association between electronic diary-rated sleep, mood, energy, and stress with incident headache in a community-based sample," *Neurology*, vol. 102, no. 4, e208102, 2024, **Article featured at CNN, Nat Geo, NIH Director's blog and others.**
- 5 **D. Dey**, A. Datta, and S. Banerjee, "Modeling multivariate spatial dependencies using graphical models," *The New England Journal of Statistics in Data Science*, pp. 1–13, 2023.
- 6 J. Glaus, S. J. Kang, W. Guo, F. Lamers, M.-P. F. Strippoli, A. Leroux, **D. Dey**, K. J. Plessen, J. Vaucher, P. Vollenweider, *et al.*, "Objectively assessed sleep and physical activity in depression subtypes and its mediating role in their association with cardiovascular risk factors," *Journal of psychiatric research*, vol. 163, pp. 325–336, 2023.
- 7 E. K. Stapp, V. Zipunnikov, A. Leroux, L. Cui, M. M. Husky, **D. Dey**, and K. R. Merikangas, "Specificity of affective dynamics of bipolar and major depressive disorder," *Brain and Behavior*, e3134, 2023.
- 8 **D. Dey**, A. Datta, and S. Banerjee, "Graphical gaussian process models for highly multivariate spatial data," *Biometrika*, vol. 109, no. 4, pp. 993–1014, 2022.
- 9 M. R. Kaufman, **D. Dey**, C. Crainiceanu, and M. Dredze, "# Metoo and google inquiries into sexual violence: A hashtag campaign can sustain information seeking," *Journal of interpersonal violence*, vol. 36, no. 19–20, pp. 9857–9867, 2021.
- 10 **D. Dey** and V. Zipunnikov, "Discussion of "an epidemiological forecast model and software assessing interventions on the covid-19 epidemic in China"," *Journal of Data Science*, vol. 18, no. 3, pp. 433–436, 2020.
- 11 S. Deb and **D. Dey**, "Spatial modeling of shot conversion in soccer to single out goalscoring ability," *Journal of Sports Analytics*, vol. 5, no. 4, pp. 281–297, 2019.
- 12 V. R. Varma, **D. Dey**, A. Leroux, J. Di, J. Urbanek, L. Xiao, and V. Zipunnikov, "Total volume of physical activity: TAC, TLAC or TAC (λ)," *Preventive medicine*, vol. 106, pp. 233–235, 2018.
- 13 A. Spira, V Zipunnikov, M. Wu, **D Dey**, E. Simonsick, L Ferrucci, C Davatzikos, and S Resnick, "Association of circadian rest and activity rhythms with brain volumes in cognitively normal older adults," *Innovation in Aging*, vol. 1, no. suppl_1, pp. 866–866, 2017.
- 14 V. R. Varma, **D. Dey**, A. Leroux, J. Di, J. Urbanek, L. Xiao, and V. Zipunnikov, "Re-evaluating the effect of age on physical activity over the lifespan.," *Preventive medicine*, vol. 101, pp. 102–108, 2017, **Article featured at TIME, Washington Post, WSJ, BBC, WPYR and others.**
- 15 V Zipunnikov, **D Dey**, A Leroux, J Di, J Urbanek, J Schrack, and C Crainiceanu, "Total physical activity and its circadian allocation are independent predictors of mortality," *Innovation in Aging*, vol. 1, no. Suppl 1, p. 1239, 2017.

Preprints and Under Revision

- 1 A. Burman, S. Choudhury, and **D. Dey**, *Doubly-unlinked regression for dependent data*, 2026. arXiv: 2603.19506 [math.ST]. [URL: https://arxiv.org/abs/2603.19506](https://arxiv.org/abs/2603.19506).
- 2 **D. Dey**, R. Ghosal, K. Merikangas, and V. Zipunnikov, *Multivariate functional principal component analysis for mixed-type mhealth data: An application to mood disorders*, 2026. arXiv: 2603.11385 [stat.ME]. [URL: https://arxiv.org/abs/2603.11385](https://arxiv.org/abs/2603.11385).
- 3 **D. Dey** and V. Zipunnikov, *Regression and dimension reduction for multivariate mixed-type data via semiparametric gaussian copula*, 2026. arXiv: 2205.06868 [stat.ME]. [URL: https://arxiv.org/abs/2205.06868](https://arxiv.org/abs/2205.06868).
- 4 M. M. Husky, L. Hunt, A. Leroux, **D. Dey**, K. Conway, C. Vetter, V. Zipunnikov, R. Choudhari, M. Kass, M. Leyden, Y. Kotani, T. Lateef, G. Salum, A. Swaminathan, Y. Xiao, A. Klein, M. Milham, and


K. R. Merikangas, *Real-time Ecological Assessment of the Context of mental and physical Health (REACH) on Curious*, Research Square [Preprint]. PMID: 41472691, 2025.  URL: <https://doi.org/10.21203/rs.3.rs-8171682/v1>.

- 5 D. Dey and V. Zipunnikov, *Connecting population-level auc and latent scale-invariant R^2 via semiparametric gaussian copula and rank correlations*, 2019. arXiv: 1910.14233 [stat.ME].

Packages and Repositories



- 1 D. Dey and R. Ghosal, *R package – SGCTools: Analysis tools for multivariate mixed data and continuous/truncated/discrete functional data using Semiparametric Gaussian Copula (SGC)*, <https://github.com/Ddey07/SGCTools>, 2023.
- 2 D. Dey and A. Pita, *The Good, The Bad, and The Ugly of the Beautiful Game — Analyzing FIFA World Cup 2018 data*, <https://ddey07.github.io/open-data/>, 2018.

Books and Chapters

- 1 S. Deb and D. Dey, *Mathematical Techniques for Competitive Examinations*, English, Paperback, ISBN: 978-9393330109.  URL: <https://lead.to/amazon.com/?op=bt&la=en&cu=usd&key=9393330107>.





Grants and Funding

Funded Support

- 2026–2027  **NIH Director’s Challenge Innovation Award** — *Biologic Rhythms and Environmental Contexts in Human Health: Integrating Epidemiology and Circadian Science*. PI: K. Merikangas (NIMH). Total: \$192,000 (FY2026) + \$213,064 (FY2027). Role: Biostatistician Consultant (IPA, Texas A&M); leading data management and analytic strategies for multilevel circadian and metabolic data integration.
- 2026  Intergovernmental Personnel Act (IPA) Assignment with National Institute of Mental Health (NIMH), NIH. \$35,147.54 (summer appointment). Role: Appointee. Organizing and conducting collaborative research on real-time ecological assessment of mental and physical health outcomes.

Talks and Presentations

Invited Talks

- 2025  “Digital Phenotyping for Mixed-Type mHealth Data.” Interdisciplinary Statistical Research Unit (ISRU), Indian Statistical Institute, Kolkata, India. December 2025.
-  “Innovations in Multivariate Stochastic Processes: Bridging Environment and Digital Health.” STATCafe Graduate Seminar Series, Department of Statistics, Texas A&M University. September 2025.
-  “Innovations in Multivariate Stochastic Processes: Bridging Environment and Digital Health.” Precise Advanced Technologies and Health Systems for Underserved Populations (PATHS-UP), Texas A&M University. September 2025.
-  “Graphical Methods for Multivariate Spatial Data.” Interdisciplinary Hydrology Seminar Series, Departments of Biological and Agricultural Engineering and Water Management and Hydrologic Sciences, Texas A&M University. September 2025.

Talks and Presentations (continued)

Contributed Talks

- 2024 ■ “Multivariate Principal Component Analysis for Mixed-Type Functional Data with Application to mHealth in Mood Disorders.” 19th International Conference on Computational and Methodological Statistics (CMStatistics 2025), Birkbeck, University of London, UK. December 2025.
- 2024 ■ “Multivariate Principal Component Analysis for Mixed-Type Functional Data with Application to mHealth.” Joint Statistical Meetings (JSM 2024), Portland, OR.
- 2023 ■ “Covariance Estimation and Functional Principal Component Analysis for Mixed-Type Functional Data.” Joint Statistical Meetings (JSM 2023), Toronto, Canada.
- 2021 ■ “Semiparametric Gaussian Copula Regression Modelling of Mixed Data Types.” ENAR Spring Meeting 2021.
- “Using Mobile Technologies to Investigate Impaired Sleep, Mood, and Energy as Real-Time Triggers of Migraine.” International Conference on Ambulatory Monitoring of Physical Activity and Movement (ICAMPAM 2021).
- “Modelling of Mixed-Type Intensive Longitudinal Data via Semiparametric Gaussian Copula and Its Application to Real-Time Mobile Monitoring of Daily Health Behaviors.” Mobile Apps and Sensors in Surveys (MASS 2021).
- “Graphical Gaussian Process Models for Highly Multivariate Spatial Data.” Joint Statistical Meetings (JSM 2021). *Travel Award Winner*.
- 2020 ■ “Connecting Population-Level AUC and Latent R^2 via Semiparametric Gaussian Copula.” ENAR Spring Meeting 2020.
- 2019 ■ “Connecting Population-Level AUC and Latent R^2 via Semiparametric Gaussian Copula and Rank Correlations.” Joint Statistical Meetings (JSM 2019). *Travel Award Winner*.
- “Joint Modeling of Binary and Continuous Measurements in Large Health Surveys and Its Application to Network Analysis, Frailty, and Mortality in NHANES 1999–2010.” ENAR Spring Meeting 2019.

Poster Presentations

- 2022 ■ “Influences of Mood, Energy and Sleep on Incident Headache Assessed with Prospective Real-Time Electronic Assessments.” 24th Annual NIMH IRP Scientific Training Day.

Session Organizing


- 2025 ■ Invited session proposal accepted: “Data-driven Advances in Mental Health Statistics: Novel Methods in mHealth and Neuroimaging.” 33rd International Biometric Conference (IBC 2026), Seoul, South Korea.
- Topic-contributed session proposal accepted: “Data-driven Advances in Mental Health Statistics: Novel Methods in mHealth and Neuroimaging.” Joint Statistical Meetings (JSM 2026).
- 2023 ■ Organized invited session: “Recent Developments in Methods for Digital Brain Health Data.” Joint Statistical Meetings (JSM 2023), Toronto, Canada.

Teaching



Courses Taught

- Spring 2026 ■ **STAT 632:** Statistical Methodology II — Bayesian Modeling and Inference, Texas A&M University.


Teaching (continued)

- Fall 2025  **STAT 641: The Methods of Statistics I** — Sections 600 (in-person) and 700 (web), Texas A&M University.

Teaching Assistantships



- 2019–2022  Statistical Methods in Public Health I–IV (PH.140.621–624), Johns Hopkins Bloomberg School of Public Health.
- 2018–2019  Probability Theory I–IV (PH.140.721–724), Johns Hopkins Bloomberg School of Public Health.

Other Teaching




- 2015–2017  Instructor, Ramanujan School of Mathematics. Taught advanced high school mathematics for two years.

Mentoring









PhD Student Mentoring

- 2026–  Gayun Kwon, Statistics PhD student, Texas A&M University.
-  Swapnonil Mondal, Statistics PhD student, Texas A&M University.

Other Mentoring

- 2022–2024  Ananya Swaminathan, PostBac, National Institute of Mental Health. Statistical components of research projects.
-  Rene Chaudhuri, PostBac, National Institute of Mental Health. Statistical components of research projects.
- 2020–2022  Final-term capstone project mentor for diverse cohorts of international students (25 countries) from non-statistical backgrounds, MPH program at Johns Hopkins Bloomberg School of Public Health.

Awards and Honors



- 2022  **Inducted into the Alpha chapter of Delta Omega**, national honorary society in public health.
- 2021  **Joint Statistical Meetings paper award**, *Section on Bayesian Statistical Science* of American Statistical Association.
-  **Selected among top six presenters**, Opta Pro Forum, 2021, one of the biggest soccer analytics conferences attended by professionals from over 80 clubs and federations worldwide.
- 2020  **Joint Statistical Meetings paper award**, *Section on Survey Research and Methods* of American Statistical Association.
- 2018  **Top five finalist**, US Soccer Hackathon 2018, Chicago, USA, July 14–15, 2018.
- 2012  **INSPIRE scholarship**, Department of Science and Technology, Govt. of India.
- 2011, 2012  **Awardee**, Regional Math Olympiad, India.
- 2011  **Top 5%**, UNESCO Science Olympiad.

Professional Activities and Service


Affiliations and Research Groups

- 2025–  Affiliated Member, Center for Biological Clocks Research (CBCR), Texas A&M University.
 Active member, Space Time Research Group, Department of Statistics, Texas A&M University.



University Service

-  Member, Undergraduate Program Committee, Department of Statistics, Texas A&M University.
2025  Judge, poster competition at the 8th Annual Symposium on Cancer Research: Basic Science to Bioinformatics, Rudder Tower, Texas A&M University (in collaboration with the Center for Statistical Bioinformatics).

Editorial and Peer Review Activities

- 2021–  Peer reviewer for: *Biometrics*, *Journal of Machine Learning Research*, *Journal of Computational and Graphical Statistics*, *Annals of Applied Statistics*, *Econometrics and Statistics*, *Mental Health and Physical Activity*, *TEST*, *Environmetrics*, *Sensors*, *Journal of Interpersonal Violence*.

Technical Skills

- Programming  R, Python, HTML, \LaTeX
Languages  English, Bengali, Hindi.

References

Dr. Vadim Zipunnikov, *PhD advisor*

Associate Professor
Johns Hopkins Bloomberg School of Public Health,
Dept. of Biostatistics
Email: vzipunni@jhu.edu

Dr. Kathleen Merikangas, *Postdoc advisor*

Chief of the Genetic Epidemiology Research
Branch
National Institute of Mental Health,
NIH
Email: merikank@mail.nih.gov

Dr. Abhirup Datta, *PhD advisor*

Associate Professor
Johns Hopkins Bloomberg School of Public Health,
Dept. of Biostatistics
Email: abhidatta@jhu.edu

Dr. Sudipto Banerjee, *Collaborator*

Professor
University of California Los Angeles,
Fielding School of Public Health,
Dept. of Biostatistics
Email: sudipto@ucla.edu