

EDUCATION

Harvard T.H. Chan School of Public Health Boston, MA
Ph.D. in Biostatistics (Expected Graduation: 2027) 2022–Present
Advisor: Nima Hejazi

University of Massachusetts Dartmouth North Dartmouth, MA
B.S. in Data Science, GPA: 4.0/4.0 2018–2022

EXPERIENCE

Research Assistant, Computational Statistics & Data Science Lab North Dartmouth, MA
University of Massachusetts Dartmouth 2020–2022

- Developed R package for COVID-19 modeling in small areas
- Prepared journal and conference papers on topics in smart health and autonomous driving
- Used SAS to harmonize data from multiple dietary studies in Massachusetts

Project Manager, Research in Industrial Projects for Students Los Angeles, CA
Institute for Pure and Applied Mathematics, UCLA Summer 2021

- *Client:* The Aerospace Corporation
- Led team to develop object tracking simulation software and dashboard in Python
- Presented results weekly to client scientists and executives

Research Experience for Undergraduates: Ecological Modeling Remote
University of Wisconsin La Crosse Summer 2020

- Led team to develop novel forest cover classification model using R
- Presented results to U.S. Geological Survey and U.S. Army Corps of Engineers

Research Assistant, Public Policy Center North Dartmouth, MA
University of Massachusetts Dartmouth 2019–2020

- Analyzed socioeconomic data and created infographics for non-technical audiences
- Scraped US Patent and Trademark Office data using Python

PUBLICATIONS

- [1] **S. V. Balkus** and D. Yan, “Improving short text classification with augmented data using gpt-3”, *Natural Language Engineering*, pp. 1–30, Aug. 2023, ISSN: 1469-8110.
- [2] **S. V. Balkus**, H. Fang, and H. Wang, “Federated fuzzy clustering for decentralized longitudinal behavioral health data”, *submitted to IEEE Transactions on Big Data*, 2022.
- [3] **S. V. Balkus**, H. Wang, B. D. Cornet, C. Mahabal, H. Ngo, and H. Fang, “A survey of collaborative machine learning using 5G vehicular communications”, *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 1280–1303, 2022.

- [4] V. S. Gurugubelli, H. Fang, J. M. Shikany, **S. V. Balkus**, J. Rumbut, H. Ngo, H. Wang, J. J. Allison, and L. M. Steffen, “A review of harmonization methods for studying dietary patterns”, *Smart Health*, vol. 23, p. 100 263, Mar. 2022.
- [5] **S. V. Balkus**, H. Fang, J. Rumbut, A. Moormann, and E. Boyer, “A multi-level biosensor-based epidemic simulation model for COVID-19”, *IEEE Internet of Things Journal*, pp. 1–1, 2021.
- [6] **S. V. Balkus**, J. Rumbut, H. Wang, and H. Fang, “An adaptive and dynamic biosensor epidemic model for COVID-19”, in *2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI)*, IEEE, Aug. 2020.

PRESENTATIONS AND POSTERS

1. “Nonparametric Network Causal Inference for Continuous Exposures in Mobile Source Air Pollution,” *American Causal Inference Conference, May 2024*
2. “Assumption-Lean Causal Inference for Mobile Source Air Pollution,” *ASA Boston Chapter Student Research Symposium on Statistics and Data Science, April 2024*
3. “Improving Natural Language Classification with Augmented Data from GPT-3,” *University of Massachusetts Dartmouth, April 2022*
4. “Language Models That Teach Themselves: Augmenting Training Data for Topic Classification Using GPT-3,” *ASA Boston Chapter Student Research Symposium on Statistics and Data Science, April 2022*
5. N. Pai, **S. V. Balkus** and T. Zeng, “Multi-Hypothesis Tracking of Space Objects and Targets,” *AMS Joint Mathematics Meetings (JMM) Poster Session, April 2022*
6. N. Pai, **S. V. Balkus**, T. Zeng, and E. Sosa. “Multi-Hypothesis Tracking of Space Objects and Targets,” *Institute for Pure and Applied Mathematics, August 2021*
7. “Multi-Level Biosensor-based Epidemic Forecasting in Small Areas,” *ASA Joint Statistical Meetings, August 2021*
8. “Lunchtime Computing: Basics of AWS Sagemaker,” *Center for Science Computing and Visualization Research, University of Massachusetts Dartmouth, February 2021*
9. “A Classification System for Characterizing Diversity Across Floodplain Forests of the Upper Mississippi River System,” *University of Wisconsin La Crosse, August 2020*
10. “Lunchtime Computing: Getting Started with Git and GitHub,” *Center for Science Computing and Visualization Research, University of Massachusetts Dartmouth, February 2020*

TEACHING

- **Teaching Fellow** at Harvard T.H. Chan School of Public Health Fall 2023
Methods (BST 231)

SCHOLARSHIPS AND AWARDS

- Certificate of Distinction in Teaching, BST 231: Methods 2024
- National Science Foundation Graduate Research Fellowship 2022
- Academic Excellence Award: Honors College, *University of Massachusetts Dartmouth* 2022
- Academic Excellence Award: College of Engineering, *University of Massachusetts Dartmouth* 2022
- John H. Ohly Award for Outstanding Economics Minor, *University of Massachusetts Dartmouth* 2022
- Best Overall Analysis, *American Statistical Association DataFest* 2022

- Best Data Visualization, *American Statistical Association DataFest* 2021–2022
- Travel Award, *AMS Joint Mathematics Meetings* 2021
- Dean’s Scholarship, *College of Engineering, University of Massachusetts Dartmouth* 2021
- Chancellor’s List, *University of Massachusetts Dartmouth* 2018–2022
- University Commonwealth Scholarship, *University of Massachusetts Dartmouth* 2018

SKILLS

- **Programming:** Julia, Python, R
- **Computing:** Git, Linux
- **Writing:** L^AT_EX, Microsoft Office

SERVICE ACTIVITIES

- Co-Chair** 2024–2025
Harvard Biostatistics PhD Student Committee
- Organized events for Biostatistics PhD students
- Student Leader** 2023–2024
Harvard Biostatistics Peer Mentoring Program
- Matched incoming students to mentors and created instructional materials for mentoring
 - Served as a peer mentor for incoming students
 - Acquired funding for and organized Student Support Seminar series
- Graphic Designer and Blog Contributor** 2022–2024
Harvard Science in the News
- Designed infographics in Adobe Illustrator to explain complex scientific topics to a lay audience
 - Wrote blog posts communicating statistical ideas to a non-mathematical audience
- Session Chair** 2022
ASA Boston Chapter Student Research Symposium on Statistics and Data Science
- Chaired Session on Biostatistics
- President** 2020–2022
Big Data Club, University of Massachusetts Dartmouth
- Organized data science workshops
 - Created marketing campaigns to boost membership
 - Led annual DataFest teams to win multiple awards
 - Networked with local clients for consulting projects
- Student Panelist** 2019–2022
University of Massachusetts Dartmouth
- Presented opportunities in data science to prospective and incoming students
- Honors Council Representative** 2018–2020
University of Massachusetts Dartmouth
- Organized social activities for students in the university’s Honors Program