Gemma E. Moran

CONTACT INFORMATION

475 Hill Center, 110 Frelinghuysen Road, Piscataway, NJ 08854
Primary email: gm845@stat.rutgers.edu
Secondary email: gem.e.moran@gmail.com
Website: https://www.gemma-moran.com

EMPLOYMENT

Statistics Department, Rutgers University — Piscataway, NJ, USA

Assistant Professor (tenure-track) — September 2023 - present

Data Science Institute, Columbia University — New York, NY, USA

Postdoctoral Research Scientist — August 2019 - July 2023
Supervisor: David Blei

EDUCATION

The Wharton School, University of Pennsylvania — Philadelphia, PA, USA

Ph.D. Statistics — August 2014 - May 2019
Thesis Title: "Bayesian Approaches for Modeling Variation"
Thesis Advisors: Edward George and Veronika Ročková

The University of Sydney — Sydney, NSW, Australia

B.Sc. Advanced Mathematics (First Class Honours) — March 2010 - November 2013
Majors: Mathematics, Statistics
Thesis Advisor: John Ormerod

HONORS AND AWARDS

- Visiting Scholar, The University of Sydney Mathematical Research Institute — 2026
- Lindley Prize, International Society for Bayesian Analysis — 2024
- Rising Star in Data Science, University of Chicago — 2022
- Poster Award, International Society for Bayesian Analysis — 2022
- Donald S. Murray Prize for excellence in teaching, Statistics Department, Wharton — 2018
- Honorable Mention for Best Talk at the Bayesian Young Statisticians Meeting — 2018
- Wharton Doctoral Program Fellowship — 2014
- The George Allen Scholarship for Mathematical Statistics, the University of Sydney — 2013

GRANTS

- NIH Grant 1R21HL181649 — 2025-2027
  - PI: Anandi Krishnan, Co-Investigator: Gemma Moran
- Rutgers Cyberinfrastructure and AI for Science and Society (CASS) Award — 2024
  - PI: Anandi Krishnan, Co-PI: Gemma Moran

Selected Peer-Reviewed Publications

* indicates equal contributions; † student advisee

[1] G. Chatzoudis, Z. Li, **G. E. Moran**, H. Wang, and D. N. Metaxas. "Beyond Interpretability: When, Why, and How Sparse Autoencoders Enable Label-Free Visual Steering". In: *Neural Information Processing Systems*. 2026.

[2] G. Chatzoudis, K. D. Polyzos, Z. Li, D. Gu, **G. E. Moran**, H. Wang, and D. N. Metaxas. "Can Cross-Layer Transcoders Replace Vision Transformer Activations? An Interpretable Perspective on Vision". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. June 2026, pp. 4016–4025.

[3] J. Linhart, **G. E. Moran**, and D. Blei. ""All Models Are Right": The Blind Spots of Posterior Predictive Checks in Simulation-Based Inference". In: *Sim2Science: ML with Imperfect Scientific Models*. 2026.

[4] **G. E. Moran** and B. Aragam. "Toward Interpretable Deep Generative Models via Causal Representation Learning". In: *Journal of the American Statistical Association* 121.553 (2026), pp. 259–275.

[5] K. Donhauser, K. Ulicna, **G. E. Moran**, A. Ravuri, K. Kenyon-Dean, C. Eastwood, and J. Hartford. "Towards scientific discovery with dictionary learning: Extracting biological concepts from microscopy foundation models". In: *International Conference on Machine Learning*. 2025.

[6] S. F. Friedman, **G. E. Moran**, M. Rakic, and A. Phillipakis. "Genetic Architectures of Medical Images Revealed by Registration of Multiple Modalities". In: *Bioinformatics and Biology Insights* 18 (2024).

[7] **G. E. Moran**, D. M. Blei, and R. Ranganath. "Holdout predictive checks for Bayesian model criticism". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 86.1 (2024), pp. 194–214.

[8] **G. E. Moran**, J. P. Cunningham, and D. M. Blei. "The posterior predictive null". In: *Bayesian Analysis* 18.4 (2023), pp. 1071–1097. **Lindley Prize Winner**.

[9] R. Bai*, **G. E. Moran***, J. L. Antonelli*, Y. Chen, and M. R. Boland. "Spike-and-slab group lassos for grouped regression and sparse generalized additive models". In: *Journal of the American Statistical Association* 117.537 (2022), pp. 184–197.

[10] **G. E. Moran**, D. Sridhar, Y. Wang, and D. M. Blei. "Identifiable Deep Generative Models via Sparse Decoding". In: *Transactions on Machine Learning Research*. 2022.

[11] **G. E. Moran**, V. Ročková, and E. I. George. "Spike-and-slab lasso biclustering". In: *The Annals of Applied Statistics* 15.1 (2021), pp. 148–173.

[12] **G. E. Moran**, V. Ročková, and E. I. George. "Variance prior forms for high-dimensional Bayesian variable selection". In: *Bayesian Analysis* 14.4 (2019), pp. 1091–1119.

[13] V. Ročková, **G. Moran**, and E. George. "Determinantal regularization for ensemble variable selection". In: *Artificial Intelligence and Statistics* (2016).

Manuscripts

* indicates equal contributions; † student advisee

[14] M. Cui, K. Eng, J. T. Greene, Z. Ke, A. Sodagartojgi, Z. Xia, **G. E. Moran**, and M. L. Stein. "Stochastic weather generators for high-frequency wind vector time series". In: *arXiv preprint arXiv:2606.09941* (2026).

[15] S. Dutta† and **G. E. Moran**. "Aggregated Posterior Predictive Checks for Generative Modeling". In: *arXiv preprint arXiv:2609.20999* (2026).

[16] Z. Ke†, M. Cui†, **G. Moran**, and J. Cabrera. "Distributional Balancing with Machine Learning for Clinical Trial Augmentation Using Real-World Data". In: *arXiv preprint arXiv:2609.23524* (2026).

[17] **G. E. Moran** and A. Krishnan. "Nonlinear multi-study sparse factor analysis". In: *arXiv preprint arXiv:2601.18128* (2026).

Presentations

Invited Seminars

- "Bayesian Model Criticism: From Holdout Checks to Model Comparison"
  - Statistics Seminar, University of Kentucky, March 2026.
- "Identifiable deep generative models for multi-study data"
  - Statistics Seminar, University of South Carolina, November 2025.
  - Statistics Seminar, Cornell University, November 2025.
  - Statistics Seminar, Ohio State University, March 2025.
- "Identifiable variational autoencoders via sparse decoding."
  - Biostatistics/Biomedical Informatics/Big Data (B3D) Seminar, Harvard, April 2023.
  - Data Science Seminar, University of Chicago, February 2023.
  - Econometrics and Statistics Seminar, Chicago Booth, February 2023.
  - Biotatistics Seminar, Columbia University, February 2023.
  - Statistics Seminar, Rutgers, February 2023.
  - Statistics Seminar, University of Connecticut - Storrs, February 2023.
  - Data Science and Operations Seminar, USC Marshall, February 2023.
  - Statistics Seminar, Boston University, January 2023.
  - Statistics Seminar, Duke University, January 2023.
  - Statistics Seminar, University of Michigan, January 2023.
  - Statistics Seminar, University of Wisconsin - Madison, January 2023.
  - Machine Learning Seminar, Microsoft Research New England, December 2021.
  - Math and Data Seminar, New York University, October 2021.

Conferences

- "Nonlinear Multi-Study Factor Analysis." (Invited Talk) Workshop on Identifiable Representation Learning. Banff International Research Station. Banff, Canada. October 2026.
- "Nonlinear Multi-Study Factor Analysis." (Invited Talk) Workshop on Collaborative Research at the Intersection of Statistics and Engineering. University of Pennsylvania, Philadelphia, PA, USA. August 2026.
- "Aggregated Posterior Predictive Checks" (Invited Talk) Joint Statistical Meetings, Boston, MA, USA. August 2026.
- "Nonlinear Multi-Study Factor Analysis." (Invited Talk) International Society for Bayesian Analysis World Meeting. Nagoya, Japan. July 2026.
- "From causal representation learning to multi-study genomics" (Invited Talk) Causality and Machine Learning Workshop. Isaac Newton Institute, Cambridge University, UK. June 2026.
- "Nonlinear Multi-Study Factor Analysis." (Invited Talk) Maryland Statistical Symposium, University of Maryland, College Park, MD, USA. December 2025.
- "Nonlinear Multi-Study Factor Analysis." (Invited Talk) International Day of Women in Statistics & Data Science Conference, Online. October 2025.

- "Identifiable deep generative models for multi-study data." (Invited Talk) Joint Statistical Meetings, Nashville, TN, USA. August 2025.
- "Holdout Predictive Checks for Bayesian Model Checking." (Invited Talk) Objective Bayes Methodology Conference, Athens, Greece. June 2025.
- "Identifiable deep generative models for multi-study data." (Invited Talk) Frontiers of Bayesian Inference and Data Science (BIRS-CMO), Oaxaca, Mexico. September 2024.
- "Identifiable deep generative models for multi-study data." (Invited Talk) Joint Statistical Meetings, Portland, OR, USA. August 2024.
- "Holdout Predictive Checks for Bayesian Model Checking." (Invited Talk) Joint Conference on Statistics and Data Science, Kunming, China. July 2024.
- "Identifiable deep generative models for multi-study data." International Conference on Frontiers of Data Science, Hangzhou, China. July 2024.
- "Spike-and-Slab Lasso Biclustering." (Invited Talk) New England Statistical Symposium, Storrs, CT, USA. May 2024.
- "Identifiable Deep Generative Models via Sparse Decoding." (Invited Talk) First Workshop on Causal Representation Learning at the Conference on Uncertainty in Artificial Intelligence (UAI), Eindhoven, Netherlands. August 2022.
- "The Posterior Predictive Null." (Poster) International Society for Bayesian Analysis World Meeting, Montreal, Canada. June 2022. **Poster Award (top 10% of posters)**
- "Identifiable deep generative models via sparse decoding." (Invited Talk) International Chinese Statistical Association Applied Statistics Symposium, Gainesville, FL, USA. June 2022.
- "Identifiable variational autoencoders via sparse decoding." (Invited Talk) CMStatistics, London, UK. December 2021.
- "Spike-and-Slab Lasso Biclustering." Joint Statistical Meetings, Vancouver, BC, Canada. August 2018.
- "Spike-and-Slab Lasso Biclustering." Bayesian Young Statisticians Meeting, University of Warwick, UK. July 2018. **Honorable Mention for Best Talk**
- "On variance estimation for Bayesian variable selection." (Poster) International Society for Bayesian Analysis World Meeting, Edinburgh, Scotland. June 2018.
- "Independence Variance Priors for Penalized Likelihood Variable Selection." Joint Statistical Meetings, Baltimore, MA. July 2017.

Software

- `SSLB` (R Package) **Moran, G. E.** [gemoran/SSLB]
- `SSLASSO` (R Package) Ročková, V. and **Moran, G. E.** [CRAN.R-project.org/package=SSLASSO]
- `EMVS` (R Package) Ročková, V. and **Moran, G. E.** [CRAN.R-project.org/package=EMVS]

Professional Activities

- Co-Organizer, NeurIPS Workshop on Interpretability as a Science, 2026.
- Scientific Committee, International Society for Bayesian Analysis World Meeting 2026.

- Area Chair, Uncertainty in Artificial Intelligence 2025, 2026.
- Workshop Co-Chair, International Joint Conference on Artificial Intelligence 2025.
- Organizer, Invited Session at International Conference on Frontiers of Data Science (Hangzhou, China) 2024.
- External Reviewer for *Journal of the American Statistical Association, Journal of the Royal Statistical Society (Series B), Biometrika, Journal of Machine Learning Research, Transactions of Machine Learning Research, Statistical Science, Biometrics, Journal of Computational and Graphical Statistics, Bayesian Analysis, AISTATS (Selected as a "Top Reviewer" in 2022), NeurIPS, ICML, CLeaR, Statistics and Computing, International Statistical Review, Statistical Methods in Medical Research, Briefings in Bioinformatics*

Teaching

Instructor

Rutgers University

- STAT486: Applied Statistical Learning (Spring 2026)
- MSDS597: Data Wrangling with Python (Spring 2025)
  - *Major revision of course - developed modules on visualization, git and GitHub, text data, and big data.* [website]
- STAT588: Financial Data Mining (Spring 2024)
- STAT690: Special Topics Course on Generative Models (Spring 2024)
  - *New course - developed material on modern generative models and causal representation learning*
- MSDS534: Statistical Learning for Data Science (Fall 2023, 2024, 2025)
  - *Complete revision of course - co-developed new syllabus and material on deep learning, conformal inference, permutation testing and causal inference.* [website]