Ian M. Taylor

STATISTICIAN   IAN.M.TAYLOR1@GMAIL.COM • WWW.IANMTAYLOR.NET

Education

PhD, Statistics, Colorado State University, Fort Collins, CO.   2023
Bayesian Models and Streaming Samplers for Complex Data with Application to Network Regression and Record Linkage
PhD Advisors: Andee Kaplan and Bailey K. Fosdick

MS, Statistics, Colorado State University, Fort Collins, CO.   2019

BS, Mathematical Sciences, Clemson University, Clemson, SC.   2013
Concentration: Computer Science, *Magna Cum Laude*

Honors and Awards

- Science in 3 Honorable Mention, NREL   2024
- Stanford Open Datathon Finalist: data modeling track, Stanford U.   2021
- Graybill Award for Excellence in Linear Models, Colorado State U.   2019
- Phi Beta Kappa, Clemson U.   2013
- Clemson University Honors College: General Honors, Clemson U.   2013
- National Society of Collegiate Scholars, Clemson U.   2010

Professional Experience

Postdoctoral Researcher, National Renewable Energy Laboratory, Golden, CO.   2024–Present

- Developed Bayesian Neural Network-based surrogate models for multi-modal simulations, improving out-of-sample prediction accuracy with calibrated uncertainty through cross-modal learning.
- Applied novel simulation-based inference methods to parameter calibration and uncertainty quantification in electrical grid simulations, supporting development of digital twin models of grids with high penetration of inverter-based resources.
- Designed and launched a collaborative "Statistical Help Desk" providing data analysis consultation to NREL researchers; responded to 5 inquiries and hosted 2 tutorial sessions in the first month, increasing lab-wide awareness of best practices in statistical modeling.
- Partnered with researchers in machine learning and energy systems to integrate probabilistic modeling into ongoing projects, increasing the statistical rigor and interpretability of results.

Graduate Statistical Consultant, Graybill Statistics Laboratory at CSU, Fort Collins, CO.   2022

- Advised researchers across diverse fields including veterinary medicine and microbiology to select appropriate statistical models and experimental designs, yielding improved analytical rigor and interpretability of results.
- Contributed to methods sections of peer-reviewed publications, ensuring reproducible analyses and accurate interpretation of findings; one study later published in *Community Mental Health Journal*.
- Organized and delivered training sessions on Rmarkdown and git for external faculty, increasing participants' ability to create reproducible statistical analyses in collaborative publications.

Graduate Research Assistant, Colorado State University, Fort Collins, CO.   2020–2022

- Designed and implemented a Bayesian record linkage framework for streaming data, achieving comparable linkage accuracy to state-of-the-art offline methods at a fraction of the computational cost.
- Created novel MCMC algorithm for efficient streaming Bayesian updates leveraging parallel computing resources, enabling real-time inference on continually growing datasets.
- Collaborated with Laboratory for Analytic Sciences project sponsors to align delivered outcomes with project funding goals.
- Presented findings at JSM and CMStatistics conferences through posters, contributed, and invited talks.

CarMax Auto Finance, Kennesaw, GA.

Senior Analyst   2016–2017
- Managed $70M reserve forecast for Extended Service Plan (ESP) returns while improving actuarial models; refined models with team, creating $2M reserve surplus.
- Researched and implemented a Markov network clustering algorithm as record linkage solution for ESP claims data. Consolidated 350,000 records into groups for each distinct repair facility.
- Analyzed ESP claims data by repair facility, supporting the launch of a repair partnership program for small-format stores.
- Supervised a summer analyst intern, guiding research into improvements to risk segmentation by vehicle make and model.

Strategy Analyst   2013–2016
- Updated behavioral scorecard to predict account delinquency, reducing 1 to 30-day collections costs by 15%.
- Automated collection of data from post-call customer surveys; created customer service evaluation metric used across the department to increase customer satisfaction.
- Developed TCPA compliance strategy for telephone collections departments, minimizing risk with minimal business process impacts.

Teaching & Mentoring

Courses Taught
- "Introduction to Statistical Methods" (300-level, Statistics), Colorado State U.   F18, S23
- "Introduction to Biostatistics" (300-level, Statistics), Colorado State U.   S19 ,F19
- "Statistics with Business Applications Recitation" (200-level, Statistics), Colorado State U.   S18
- "General Statistics Recitation" (200-level, Statistics), Colorado State U.   F17

Courses Designed
- "Introduction to Statistical Methods," (300-level, Statistics), Colorado State U.   2019
  - 10 sections per semester offered
  - 600 total students per semester enrolled
- "Introduction to Biostatistics," (300-level, Statistics), Colorado State U.   2019
  - 3 sections per semester offered
  - 150 total students per semester enrolled

Workshops Led
- "Coding and Cookies: Version Control Using Git," Colorado State U.   S21, S22
- "Coding and Cookies: Reproducible Reports using Rmarkdown," Colorado State U.   F21

Guest Lectures

  3. "Statistical Machine Learning" (400-level, Data Science), Colorado State U.   F20

  2. "Advanced Theory of Statistics" (700-level, Statistics), Colorado State U.   S20

  1. "Bayesian Data Analysis" (400-level, Statistics), Colorado State U.   S19

Professional Activities and Affiliations

Professional Affiliations

- Member, American Statistical Association   2025–Present

Conference Organizational Roles

- Session Chair, "Computing in Large and Complex Data Analysis," JSM, Washington, DC   2022
- Conference Co-Organizer, "IMS New Researchers Conference," IMS NRC, Fort Collins, CO   2019

Journals Reviewed

- *Journal of Computational and Graphical Statistics*

Review Committees

- Workforce Development for Teachers and Scientists Abstract Competition, Office of Science   2025

Technical Expertise

Programming Languages

- R (8 years)
- Python (6 years)
- PyTorch
- SQL (5 years)
- Rmarkdown (6 years)
- Quarto
- LaTeX (8 years)
- Stan
- Bash (12 years)
- Julia

Publications

Software Authored

- bstrl: Bayesian Streaming Record Linkage
  - GitHub: https://github.com/ianmtaylor1/bstrl
  - CRAN: https://cran.r-project.org/package=bstrl
- gammacount: Distributions related to renewal processes with gamma-distributed interarrival times
  - GitHub: https://github.com/ianmtaylor1/gammacount

Peer-Reviewed Journals

  3. Taylor, I., Kaplan, A., & Betancourt, B. (2025). Generative Filtering for Recursive Bayesian Inference with Streaming Data. *Journal of Computational and Graphical Statistics*, 1–13. https://doi.org/10.1080/10618600.2024.2442000

  2. Taylor, I., Kaplan, A., & Betancourt, B. (2024). Fast Bayesian Record Linkage for Streaming Data Contexts. *Journal of Computational and Graphical Statistics*, *33*(3), 833–844. https://doi.org/10.1080/10618600.2023.2283571

1. Hughes, S., Rondeau, M., Shannon, S., Sharp, J., Ivins, G., Lee, J., Taylor, I., & Bendixsen, B. (2021). A Holistic Self-learning Approach for Young Adult Depression and Anxiety Compared to Medication-Based Treatment-As-Usual. *Community Mental Health Journal*, *57*(2), 392–402. https://doi.org/10.1007/s10597-020-00666-9

Work-In-Progress

3. "Statistical Parameter Calibration with Quantified Uncertainty in Next-Gen Electrical Grid Simulations." (in preparation)

2. "Bayesian Optimization with Multi-Modal Surrogate Models." (in preparation)

1. **"**Multi-modal Bayesian Neural Network Surrogate Models with Conjugate Estimation." (submitted)

Professional Presentations

Invited

2. "Generalizable Machine Learning Tools for Optimization," Science in 3, Golden, CO.	2024
1. "Fast Bayesian Record Linkage for Streaming Data Contexts," CMStatistics, Virtual.	2021

Contributed

3. "A Non-degrading Streaming Sampler for Recursive Bayesian Inference," Joint Statistical Meetings, Washington, DC.	2022
2. "Streaming Record Linkage for Online Data Deduplication," Joint Statistical Meetings, Virtual.	2021
1. "Restricted Regression in Networks," ISBA, Virtual.	2021

Posters

7. "Multi-modal Surrogate Models for Bayesian Optimization," Conference on Data Analysis, Santa Fe, NM.	2025
6. "Fast Bayesian Record Linkage for Streaming Data Contexts," Colorado State U, Fort Collins, CO.	2022
5. "Restricted Regression in Networks," ISBA, Virtual.	2021
4. "Sampling Methods for Streaming Record Linkage Models," Colorado State U, Fort Collins, CO.	2021
3. "Record Linkage: Basics and Streaming RL," Colorado State U, Virtual.	2020
2. "The Impact of Prior Choice on Latent Variable Network Models," Joint Statistical Meetings, Denver, CO.	2019
1. "Impact of Prior Choice on Network Models with Random Effects," Colorado State U, Fort Collins, CO	2019