Raphaël Morsomme

raphael.morsomme@duke.edu
https://rmorsomme.github.io
(858) 226-7691

Education

Research interests: statistics, causal inference, Bayesian inference, clinical trials, statistical programming, MCMC, conformal prediction, data augmentation, multistate models, contagious disease modeling, cancer modeling.

2019-present: **Ph.D. candidate in *Statistical Science***,
Department of Statistical Science, Duke University.
Advisor: Prof. Jason Xu.
Thesis: Efficient Sampling Algorithms for High-Dimensional Latent-Data Models.

2019-2023: **M.S. degree in *Statistical Science***,
Department of Statistical Science, Duke University.

2014-2018: **Double B.S. degree in *Liberal Arts and Sciences***,
University College Maastricht, The Netherlands; University College Freiburg, Germany.
Honors program, *Summa cum Laude*.
Advisors: Dr. Patrick Lindsey; Prof. Dirk Neumann.
Theses: Embryonic and Mitochondrial Modeling in the Context of *In-Vitro* Fertilization; Forecasting Financial Instability from Soft Content Anomalies.

Research and Professional Experience

05/23-current: Student Trainee (Admin Support) GS-0399-07,
Office of Biostatistics Research, Division of Clinical Research, National Institute of Allergy and Infectious Diseases, National Institute of Health.
Part-time (25 hours per week)

- Design a Monte-Carlo EM algorithm for fitting multistate semi-Markov models to panel data.
- Develop a Julia package implementing the Monte-Carlo expectation-maximization algorithm.
- Study an event-driven design for clinical trials with time-to-event data.

05/23-08/23: Statistical consultant,
MetLife Investment Management.
Part-time (15 hours per week)

- Develop scalable dynamic Bayesian system for long-term forecasting of high-dimensional macroeconomic time series.
- Design a novel framework for the dynamic selection of macroeconomic time series.
- Implement the forecasting system in python.

08/19-current: Graduate research assistant,
Department of Statistical Science, Duke University.
Full-time

- Implement a uniformly ergodic block sampler for stochastic epidemic models in continuous time in R, with an application to the 2013-2015 outbreak of Ebola in Western Africa.
- Design a novel family of non-parametric temporal stochastic epidemic models and implement a scalable data-augmentation MCMC sampler in R, with applications to the 1995 and the 2018 outbreaks of Ebola in Congo.

- Design a semi-Markov multistate model for cancer natural history to estimate the overdiagnosis rate and implement a scalable data-augmentation MCMC sampler in C++.

09/18-07/19: Data scientist,
University College Maastricht.
Full-time.

- Develop a course recommender system for Liberal Arts students.
- Construct a topic model of course content in a Liberal Arts college.
- Apply conformal prediction for student course grades.

Manuscripts

Published

Morsomme, R., & Xu, J. (2024). Uniformly Ergodic Data-Augmented MCMC for Fitting the General Stochastic Epidemic Model to Incidence Data. arXiv preprint arXiv:2201.09722. (under Review)

Huang, J., **Morsomme, R.**, Dunson, D., & Xu, J. (2024). Detecting changes in the transmission rate of a stochastic epidemic model. Statistics in Medicine.

Morsomme, R., & Smirnov, E. (2020). Valid Prediction Intervals for Course Grades with Conformal Prediction. In 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 936-941). IEEE.

Morsomme, R., & Smirnov, E. (2019). Conformal Prediction for Students' Grades in a Course Recommender System. *Conformal and Probabilistic Prediction and Applications* (pp. 196-213).

Morsomme, R., & Alferez, S. V. (2019). Content-based Course Recommender System for Liberal Arts Education. In *Proceedings of The 12th International Conference on Educational Data Mining (EDM 2019)* (Vol. 748, p. 753).

In Progress

Morsomme, R., Holloway, S., Ryser, M., & Xu, J. (2023). A Bayesian Approach for Fitting Semi-Markov Mixture Models of Cancer Latency to Individual-level Screens.

Morsomme, R., Chen, M., & Xu, J. (2023). Scalable Non-Parametric Temporal Stochastic Epidemic Models.

Morsomme, R., Liang, J., Follman, D., & Fintzi, J. (2023). A Data Assimilation Framework for Assessing Treatment Efficacy with Multistate Semi-Markov Models.

Morsomme, R., Liang, J., & Fintzi, J. (2023). MultistateModels: Fitting Semi-Markov Multistate Models to Panel Data.

Brittain, E., **Morsomme, R.**, & Proschan, M. (2023). BEDD: Binary-Event-Driven Design for Clinical Trials with Time-to-Event Data.

Invited Talks

2024: A Data Assimilation Framework for Assessing Treatment Efficacy with Multistate Semi-Markov Models.
Joint Statistical Meeting, American Statistical Association.

2023: Exact inference for stochastic epidemic models via uniformly ergodic block sampling.
6th International Conference on Econometrics and Statistics.

2019: Conformal prediction for students' grades in a course recommender system.
8th Symposium on Conformal and Probabilistic Prediction with Applications.

2019: Content-Based Course Recommender System for Liberal Arts Education.
Educational Data Mining 2019.

Awards and Grants

2024: Travel award, ISBA World Meeting,
The International Society for Bayesian Analysis.

2023: Travel award, Summer Institute in Statistics and Modeling in Infectious Diseases,
University of Washington.

2022: Outstanding Mentor of Undergraduate Research Award,
Department of Statistical Science, Duke University.

2022: Summer Course Development Grant,
Duke University.

2022: Full scholarship, Summer Institute in Statistics and Modeling in Infectious Diseases,
University of Washington.

2021: Young Investigator Award,
The ASA Section on Statistics in Epidemiology.

Outreach

2024: Guest lectures for the Ph.D. course STA863 Advanced Statistical Computing,
Department of Statistical Science, Duke University.

2024: Data organizer for the 53 institutions participating in DataFest 2024,
American Statistical Association.

2023: Peer reviewer,
Statistics in Medicine.

2023: Peer reviewer,
IEEE Journal of Biomedical and Health Informatics.

2023: Coordinator for the 2023 DataFest event at Duke University,
American Statistical Association.

2018-2022: Semi-annual workshop: Introduction to *R*,
University College Maastricht.

2021: Judge for the DataFest 2021,
American Statistical Association.

2016: Organizer of the Global Order Project conference: Mobility & Identity in a Globalizing World,
University College Freiburg.

Teaching and Mentoring

2021-current: Certificate in College Teaching,
Duke University.

Instructor of Record

2022: STA101 Data Analysis and Statistical Inference,
Department of Statistical Science, Duke University.

2021: STA101 Data Analysis and Statistical Inference,
Department of Statistical Science, Duke University.

Teaching Assistant

2024: STA863 Advanced Statistical Computing (Ph.D.),
Department of Statistical Science, Duke University.

2023: STA561 Probabilistic Machine Learning (masters),
Department of Statistical Science, Duke University.

2022: STA310 Generalized Linear Models (undergraduate),
Department of Statistical Science, Duke University.

2021: STA723 Case Studies in Bayesian Statistics (Ph.D.),
Department of Statistical Science, Duke University.

2020: STA540 Case Studies in Statistical and Data Science (masters),
Department of Statistical Science, Duke University.

2019: STA440 Case Studies in the Practice of Statistics (undergraduate),
Department of Statistical Science, Duke University.

2017: Introduction to Statistics and Data Analysis (undergraduate),
University College Freiburg, Freiburg University.

Tutoring and Mentoring

2024: Research mentor for the Intro to Undergrad Research in Statistical Science Workshop,
Department of Statistical Science, Duke University.

2023-current: Academic mentor of M. Chen,
Masters in Statistical Science, Duke University.

2021-current: Academic tutor,
SPIRE Fellows Program, Duke University.

2023-current: Research mentor for the Thesis Writer's Mentoring Workshop,
Department of Statistical Science, Duke University.

2020-current: Private tutoring.

2021-2023: Academic mentor of J. Huang,
Major in Statistical Science and Computer Science, Duke University.

2020-2021: Research mentor,
Lumiere Research Scholar Program.

Programming Skills

Proficiency in **R**, **Julia**, **MATLAB**, **LaTeX**, **Git**, **Quarto**, **STAN, Shiny**.

Working knowledge of C++, Python, SAS, SQL, Tableau, Weka.