YANG LONG

🏛 Department of Statistics, George Mason University

✉ **ylong5@gmu.edu**   🔗 **ylong-chn.github.io**

EDUCATION

PhD in Statistical Science, George Mason University, *Fairfax, VA*   Jan 2021 – May 2026 (Exp.)
- *Advisors:* Dr. David Kepplinger and Dr. Lily Wang
- *Dissertation Title:* Trustworthy AI Through Robust Functional Data Analysis and Statistical Inference for Imaging Data

MS in Statistics, CUNY–Baruch College, *New York, NY*   Jan 2017 – May 2019

MS in Finance, University of Rochester, *Rochester, NY*   Jul 2014 – Dec 2015

BEcon in Finance, Zhongnan University of Economics and Law, *Wuhan, China*   Sep 2010 – Jun 2014

RESEARCH INTERESTS

- Trustworthy AI
- Multispectral imaging analysis
- Robust nonparametric statistics
- Non-convex optimization
- Functional data analysis
- Time series analysis

HONORS & AWARDS

- **Dissertation Completion Grant,** George Mason University   2026
- **R. Clifton Bailey Travel Award,** Department of Statistics, George Mason University   2025
- **Graduate Student Travel Fund,** George Mason University   2024, 2025
- **NSF Travel Award,** ICSA Applied Statistics Symposium 2025   2025
- **Student Paper Runner-Up (Theory and Methods),** 2025 Statistical Methods in Imaging Conference   2025
- **WSS Outstanding Graduate Student Award,** Washington Statistical Society   2025
- **Academic Excellence Scholarship in MS Statistics (one recipient annually),** Baruch College   2019
- **Merit-Based Scholarship,** Simon Business School, University of Rochester   2014

PUBLICATIONS

Publications (Published & Accepted)

P1. **Y. Long**, G. Cao, D. Kepplinger, and L. Wang, "**Robust mean signal estimation and inference for imaging data**," *Statistica Sinica*, Accepted, 2025.

P2. Z. Li, S. Bruce, C.J. Wutzke, and **Y. Long**, "**Conditional adaptive Bayesian spectral analysis of replicated multivariate time series**," *Statistics in Medicine*, vol. 40, pp. 1989–2005, 2021.

P3. D. Feldman, S. Gross, and **Y. Long**, "**Gender competitiveness and predictability, and prize money in Grand Slam tennis tournaments**," *Quarterly Journal of Finance*, vol. 10, no. 2, 2020.

Working Papers (Submitted & In Preparation)

W1. G. Wang, S. Yu, **Y. Long**, Z. Gu, and L. Wang, "Distributed synthetic surrogate functional regression (D-SSFR): a scalable and robust framework for estimating spatially varying covariate effects in AI-augmented neuroimaging," *To be submitted to Annals of Applied Statistics*, 2026+.

W2. **Y. Long**, Z. Gu, and L. Wang, "Synthetic surrogate image-on-scalar regression for robust brain imaging analysis," *In preparation*.

W3. **Y. Long**, Z. Gu, G. Wang, and L. Wang, "Robust and scalable distributed learning for surface-based imaging regression with applications to neuroimaging," *In preparation*.

W4. **Y. Long**, D. Hanley, and D. Kepplinger, "Illuminant spectrum estimation and inference to study animal coloration from multispectral camera images," *In preparation*.

W5. **Y. Long**, C.E. Lee, Z. Li, "Conditionally uncorrelated components for replicated tensor time series via martingale difference divergence matrix," *In preparation*.

PRESENTATIONS AND ACTIVITIES

Invited Talks

T1. **EcoSta 2025,** *Tokyo, Japan*   Aug 2025
Robust and Scalable Distributed Learning for Surface-Based Imaging Regression with Applications to Neuroimaging

T2. **ICSA Applied Statistics Symposium 2025,** *Storrs, CT*   Jun 2025
Robust and Scalable Distributed Learning for Surface-Based Imaging Regression with Applications to Neuroimaging

T3. **2025 Statistical Methods in Imaging Conference,** *Houston, TX*   May 2025
Robust Mean Signal Estimation and Inference for Imaging Data

T4. **StatConnect 2025,** *Fairfax, VA*   Mar 2025
Robust Mean Signal Estimation and Inference for Imaging Data

T5. **ICORS meets DSSV 2024,** *Fairfax, VA*   Jul 2024
Robust Learning and Inference for Mean Functions in Functional Data Analysis of Imaging Data

Contributed Talks

T6. **2025 Joint Statistical Meetings,** *Nashville, TN*   Aug 2025
Robust and Scalable Distributed Learning for Surface-Based Imaging Regression with Applications to Neuroimaging

T7. **2024 Joint Statistical Meetings,** *Portland, OR*   Aug 2024
Illuminant Spectrum Estimation to Study Animal Coloration from Multispectral Camera Images

Contributed Posters

T8. **The Past, Present and Future of Statistics in the Era of AI,** *Washington, D.C.*   May 2025
Robust Mean Signal Estimation and Inference for Imaging Data

T9. **The Conference on Evolving Statistical Data Science,** *Fairfax, VA*   Mar 2023
Accelerated Algorithms for Elastic Net S-Estimators

RESEARCH PROJECTS

Syn-ISR: Synthetic surrogate image-on-scalar regression for robust neuroimaging analysis [W2]
- Developed a misspecification-robust surrogate-augmented image-on-scalar regression framework for settings with missing imaging outcomes and fully observed scalar covariates
- Introduced a joint modeling strategy for observed and surrogate images with correlated error components, improving robustness when the surrogate generator is imperfect
- Built a bootstrap-based simultaneous confidence corridor procedure that improves domain-wide inference compared to observed-only analyses and remains stable under surrogate-model misspecification

Robust FDA: Robust mean signal estimation and simultaneous inference for imaging on irregular domains [P1]
- Developed an M-type bivariate penalized spline over triangulation (M-BPST) estimator for reliable mean-image estimation under spatially localized outliers and heavy-tailed noise
- Proved consistency, optimal convergence rates, and asymptotic normality of the proposed estimator
- Implemented a weighted-bootstrap robust SCC procedure and benchmarked against classical smoothers and robust deep neural network-based alternatives for accuracy and computational efficiency

R-DISR-SS: Robust distributed image-on-scalar regression with scalable inference on cortical manifolds [W3]
- Designed a scalable distributed image-on-scalar regression framework for cortical surface images by domain decomposition of their spherical projection to enable structure-preserving learning at scale
- Developed robust estimation pipelines combining M-type penalized regression with triangulated spherical spline smoothing to stabilize learning under outliers, heavy-tailed errors, and high-noise imaging conditions
- Established convergence rates and asymptotic normality for distributed regression estimators, and implemented weighted-bootstrap simultaneous confidence corridors for domain-wide uncertainty quantification

Spectral illumination estimation and inference from multispectral images [W4]
- Proposed a statistical framework for recovering the illumination spectrum from multispectral camera images with off-the-shelf cameras to support animal-vision reconstruction in dynamic lighting and video settings
- Developed penalized constrained likelihood estimation for basis coefficients and the illumination covariance function, enabling simultaneous confidence bands for the illumination spectrum

CUC-MDDM: Covariate-assisted dimension reduction for replicated tensor-valued time series [W5]
- Proposed a covariate-assisted dimension reduction framework where subject-level covariates modulate serial dependence in replicated tensor time series
- Introduced conditionally uncorrelated components (CUCs) and constructed a cumulative mode-wise martingale difference divergence matrix (CUC-MDDM) using lagged cross-moment features aggregated across lags
- Developed eigen-based subspace recovery with ratio-based rank selection and sparse eigen-computation

TEACHING EXPERIENCE

Instructor
- **STAT 346: Probability for Engineers,** George Mason University   Fall 2025

Teaching Assistant
- **STAT 250: Introductory Statistics I,** George Mason University   Spring 2021
- **STA 2000: Business Statistics I,** Baruch College   Fall 2017, Spring 2018

- **STA 3154: Business Statistics II,** Baruch College   Fall 2017, Fall 2018
- **STA 9719: Foundations of Statistical Inference (Graduate),** Baruch College   Spring 2018

PROFESSIONAL EXPERIENCE

Summer Associate (Data Scientist), Navy Federal Credit Union, *Vienna, VA*   May 2024 – Aug 2024
- Enhanced credit card probability of default models for the CECL quantitative modeling team

Quantitative Analyst Intern, Truist Bank, *Charlotte, NC*   Jun 2023 – Aug 2023
- Developed a machine learning framework with SAS and Python for suspicious transaction monitoring

Graduate Assistant, Statistical Consulting Laboratory, Baruch College, *New York, NY*   Aug 2017 – May 2019
- Advised business school faculty and graduate students on data visualization and statistical software

Quantitative Research Associate, Terrapin Asset Management, LLC, *New York, NY*   Oct 2015 – Dec 2016
- Performed empirical data analysis on hedge fund activism to validate and enhance a new hedge fund strategy

SERVICE TO THE PROFESSION

Conference Session Chair
- **2025 Joint Statistical Meetings,** *Nashville, TN*   Aug 2025
  Functional Regression for Complex Data: Accommodating Ordinal, Truncated, and Zero-Inflated Structures

Journal Referee
- Journal of Applied Statistics: Enviro Stats (1 article)
- Journal of Computational and Graphical Statistics (1 article)
- Journal of Nonparametric Statistics (2 articles)
- Statistical Analysis and Data Mining (2 articles)
- TEST (1 article)

Conference Volunteer
- **ICSA Applied Statistics Symposium 2026,** *Fairfax, VA*   Jun 2026
- **StatConnect 2025,** *Fairfax, VA*   Mar 2025
- **ICORS meets DSSV 2024,** *Fairfax, VA*   Jul 2024
- **IMS Meeting of New Researchers in Statistics and Probability,** *Fairfax, VA*   Aug 2022
- **SC21 (ACM/IEEE Supercomputing Conference),** *St. Louis, MO*   Nov 2021

UNIVERSITY SERVICES

George Mason University
- **President/Vice President,** Statistics Graduate Student Association (SGSA)   Jun 2021 – Present
- **PhD Representative,** Graduate and Professional Student Association (GAPSA)   2021 – 2023, 2025 – 2026

REFERENCES

David Kepplinger

Assistant Professor

Department of Statistics

George Mason University

Fairfax, VA 22030

dkepplin@gmu.edu

Lily Wang

Professor

Department of Statistics

George Mason University

Fairfax, VA 22030

lwang41@gmu.edu

Anand Vidyashankar

Professor

Department of Statistics

George Mason University

Fairfax, VA 22030

avidyash@gmu.edu

Zeda Li

Associate Professor

Paul H. Chook Department of Information Systems and Statistics

Baruch College, CUNY

New York, NY 10010

zeda.li@baruch.cuny.edu