

Xucheng (Fred) Huang

xucheng.huang@emory.edu | (404) 398-3186 | Atlanta, GA

PROFESSIONAL SUMMARY

Ph.D. candidate in Population Health Data Science with an **M.S. in Biostatistics** and **5+ years** of combined research and analytics experience, including **3 years of full-time industry consulting experience**. Specializes in **Bayesian inference**, **functional data analysis**, and end-to-end **ML pipelines** on large-scale complex real-world data spanning 60M+ sensor networks to longitudinal biosignals. Combines deep methodological expertise with strong business acumen to deliver high-impact, data-driven insights to both technical and executive audiences.

TECHNICAL SKILLS

Programming & Data Tools: Python, R, SAS, Stata, GitHub

Data Science & Modeling: Time-series analysis, functional data analysis, high-dimensional data analysis, Bayesian modeling, causal inference, longitudinal/multilevel modeling, spatial analysis, machine learning

Visualization & Analytical Platforms: Tableau, Power BI, Salesforce, Qualtrics

EDUCATION

Emory University, Rollins School of Public Health

Atlanta, GA

Ph.D., Environmental Health Sciences — Population Health Data Science

Aug. 2024–Expected May 2028

M.S., Biostatistics | GPA: 4.0/4.0 | Rank 1/26 | Penina & Michael Haber Award | Delta Omega

Aug. 2022–May 2024

APPLIED DATA SCIENCE & RESEARCH EXPERIENCE

Data Analyst Graduate Assistant | Emory University

Feb. 2023–Present

Atlanta, GA

- Bayesian Spatiotemporal Modeling:** Developed a 2-stage spatiotemporal pipeline integrating quasi-Poisson time-series with a **Bayesian spatial hierarchical model (CAR prior)** across **61.8M+ ED records, 279 counties, 13 years**; demonstrated that conventional diagnosis-coded surveillance underestimates true influenza burden by **25–38%**; published as first author.
- High-Frequency Longitudinal Modeling:** Engineered bedtime-aligned functional features from **300K+ biosensor records (30-sec intervals)**; applied multilevel scalar-on-function regression to decompose between-person and within-person variation, revealing that skin temperature trajectory shape **3.5–4 hours** before sleep onset predicts sleep efficiency while scalar averages carry no signal.
- Multimodal Attention Decoding:** Open-source R pipeline ([Code](#)) synchronizing **fMRI (96K+ features/scan) with eye-tracking (340K+ gaze events/participant)** across heterogeneous sampling rates; applied **Bayesian spatial GLM** on cortical surfaces, revealing dissociable neural signatures of visual attention engagement and disengagement from real-time gaze behavior.
- Spatial Distribution Modeling:** Captured full daily temperature distributions from **1-km gridded data (19 years, 3,496 days)** and linked distribution shape directly to ED visit counts via scalar-on-quantile-function regression ([Code](#)); uncovered that mean-based models underestimate peak-day heat burden by **76%** by collapsing spatial heterogeneity into a single scalar.
- IoT-Scale Environmental Inference:** Processed **58M+ minute-level records** from **13,487 PurpleAir sensors** on Emory HPC; built automated QC pipeline, reducing processing time from **1 day to 50 min**; applied random-effects meta-analysis and unsupervised clustering across **250K+ days**, uncovering 4 distinct indoor smoke infiltration patterns and 53% higher outdoor PM_{2.5} on smoke days.
- Counterfactual Disease Detection:** Modeled **7.5M+ respiratory emergency department visits** across 15 years via **quasi-Poisson time-series models**; incorporated spline-based interactions, counterfactual prediction, and random-effects meta-analysis to surface disease signals invisible to standard coding systems, detecting **6x more flu-attributable visits** than rule-based surveillance; produced **two first-author publications**.

PROFESSIONAL EXPERIENCE

Data Analyst Graduate Assistant | Emory Global Engagement

Aug. 2025–Present

Atlanta, GA

- Data integration:** Led collection, cleaning, integration, and validation of **four years** of global-engagement data across **11 university schools and units**; analyzed longitudinal trends to inform institution-wide planning.
- Dashboards and reporting:** Built interactive Tableau dashboards and global-engagement maps covering **80+ countries and territories**; produced a statistical report used by university leadership at a conference in Germany.

Business Consultant | Noldus (Beijing) Information Technology Co., Ltd.

Mar. 2019–Apr. 2022

Beijing, China

- Client analytics:** Designed and led usability studies combining **EEG, ECG, EDA, eye-tracking**, and emotion-recognition data; translated behavioral and physiological findings into product recommendations for commercial clients.
- Project delivery:** Executed 10+ subprojects and completed 300+ hours of stakeholder interviews.
- Business development:** Developed tailored client proposals that helped secure **two multimillion-dollar annual contracts**.

SELECTED PUBLICATIONS

Huang, X. F., et al. “A Time-Series Approach for Estimating Emergency Department Visits Attributable to Seasonal Influenza: Results From Six U.S. Cities, 2005–2006 to 2016–2017 Seasons.” *American Journal of Epidemiology*, 2026; 195(1):40–48. [DOI](#)

Huang, X. F., et al. “County-Level Influenza-Attributable Emergency Department Visits and Their Spatial Correlates in the United States: Cross-Sectional Observational Study.” *JMIR Public Health and Surveillance*, 2025; 11:e82879. [DOI](#)