Mengling Qiao

Associate Research Scientist, Columbia University

Tel: +1 332-258-2102 | Email: m.qiao@columbia.edu | LinkedIn | Google Scholar

CORE COMPETENCIES

- Time-Series Analysis & Forecasting • Applied ML/DL • Spatial-Temporal Modeling & Analysis • Data Science • NLP • Applied Generative AI • Data-Driven Research

EDUCATION

Wuhan University | Direct-Entry Ph.D. in *Geographic Information Science* | CN   *Sept. 2015 - Dec. 2020*

- o **Awards**: Recipient of 4 graduate studies scholarships×4 (10%), The Scholarship for Outstanding Freshmen (5%)

Nanjing Normal University | B.Sc. in *Geographic Information Science* | CN   *Sept. 2011 - Jun. 2015*

- o **Awards**: Recipient of Talent Scholarship (1%), Outstanding Student Scholarship (2%), National Encouragement Scholarship (1%)

APPOINTMENTS

Columbia University in the City of New York   *New York, U.S.*

Associate & Postdoctoral Research Scientist | *Earth and Environmental Engineering*   *Jan. 2023 - present*
Guest Lecturer | *Civil Engineering & Engineering Mechanics* | *Data Science Institute*   *Jan. 2023 - Dec. 2024*

- **Secured & Executed Two Multi-Institutional Projects**: Co-led *$600K* projects on data-driven disaster response. Defined technical strategy and coordinated execution in multivariate time-series modeling, forecasting, and data engineering, driving project direction and ensuring all major milestones were met.
- **Comprehensive Knowledge Integration:** Synthesized 900+ academic papers and 400+ reports using AI tools to guide decisions.
- **Proposal Development & Technical Leadership**: Coordinated 3+ teams to prepare an NSF $500K proposal on spatiotemporal ML/DL, graph modeling, and transfer learning for smart drainage simulation. (ongoing).

The Hong Kong University of Science and Technology   *Hong Kong, SAR*

Postdoctoral Fellowship | *Computer Science & Engineering*   *Jul. 2022 - Jan. 2023*

- **Rapid Proposal Development:** Led a team of 5 PhD students to develop an AR/VR research proposal, completed successfully within one week
- **Visual Analytics and Pipeline Optimization:** Supported decision-support systems and visualization projects; accelerated delivery by 4 months through improved pipeline design.
- **Awards:** Recipient of the Research Talent Hub of Hong Kong (*HKD 431,264*)

The Chinese University of Hong Kong   *Hong Kong, SAR*

Postdoctoral Fellowship | *Geography and Resource Management*   *Jun. 2021 - Jul. 2022*

- **Executed Project on Risk Modeling**: Led a $450K project; developed indices and extended models for spatial-temporal risk quantification and prediction; delivered actionable insights for agencies using predictive analytics.
- **Outreach**: Delivered invited talks at multiple universities to disseminate findings.

PROJECTS

Resilience Metrics for High-Frequency Travel Behavior ⊙ GitHub   *National Science Foundation (2023–2024)*

- Developed time-series anomaly detection, cross-correlation models, and network-based resilience metrics by integrating multimodal, high-frequency transportation data (3+ million time-series records spanning subway, taxi, rideshare, bike); Built automated Python workflows, reducing manual processing by 40%.
- Quantified impacts: Ida accounted for *40%* of anomalies, the subway most affected (*23%*); Citi Bike fastest to respond, FHV/taxi fastest to recover.

Time-Series Modeling of Indicators for Public Safety   *United States-Japan Foundation (2023–2024)*

- Modeled 15-year city-level time series of climate, mobility, and census indicators using Bayesian models and multiple ML/DL to quantify city safety. Achieved $R^2 = 0.85$ citywide, improving prediction accuracy by ≈3–5x over weaker baselines.
- Employed advanced feature analysis (Python Pipeline with SHAP model) to rank predictors and improve accuracy.

Spatial-Temporal Risk Quantification and Prediction   *The Chinese University of Hong Kong (2021–2022)*

- Designed PCA-weighted time-series indices integrating mobility and demographic features to assess public health risk across multiple spatial scales. Achieved ≈60–65% stronger correlation with ground truth compared to traditional index (correlation improved from 0.48 to 0.79).
- Improved prediction accuracy by extending GTWR methods to capture dynamic time-series patterns in disease spread and community response; Outperformed baseline methods: R^2 improved from 0.23 (OLS) to 0.93 (GTWR) in temporal models (≈300% gain) and from 0.36 to 0.8 in spatial models *(≈120% gain)*.

Social Media Posts and Street View Images Analytics   *Multiple Studies (2018–2025)*
- **Risk and Sentiment Modeling:** Built NLP (Chinese-BERT-wwm) and geospatial models to analyze *10+ million* social media posts; identified public responses to heatwaves (F1-score =0.97); geospatial models achieved enhanced model explanatory power with R^2 = 0.87, *1.5x* higher than standard regression.
- **Street View–Driven Insights:** Developed computational frameworks integrating semantic and geometric features from *166K+* street view imagery and *125K+* building footprints, and human emotions from *1.6+million* social media posts; applied Generalized Additive Mixed Model to capture nonlinear relationships of built environment and human well-being; Identified key thresholds, such as green view optimal at 45%.
- **Economic Segregation Analysis:** Built multilevel measures of economic segregation using large-scale human activity data from *15+ million* social media, *121K+* customer review comments, and *39K+* property transaction records. The segregation index developed with TF-IDF weighting reveals finer-grained and more realistic economic segregation patterns than traditional methods.

HIGHLIGHTS

- ***Team Leadership***: Led applications for 3 grant attempts (1 funded, 1 under review); As PI for 1 research project; Led a capstone project ranked 1st of 6 teams; Supervised 3 graduate-level training projects on NLP, ML/DL, and Graph Theory (delivered 32+ invited classes for 30+ graduates); mentoring 2 graduate students at the Data Science Institute.
- ***Cross-Functional Engagement***: Led a 4-hour hands-on workshop on "ML/DL for socio-environmental data analytics" at I-GUIDE Forum 2025.
- ***Project Delivery***: Acted as lead representative in a multi-PI research collaboration team, regularly presenting and coordinating across faculty and researchers; Managed multiple concurrent research timelines across universities and international collaborators, consistently delivering output on time.
- ***Multi-Source Data Perspective***: Leveraged human activity data, spatial big data, and emerging technology–generated streams (sensors, imagery, IoT, social platforms) to uncover hidden dynamics beyond traditional sources, enabling improved forecasting, resilience assessment, and decision-making.
- ***Research Impact***
  - Specialized in long-term and high-frequency time-series modeling across domains
  - Published 15+ peer-reviewed papers in top outlets and 2 book chapters, advancing methods in data-driven time-series analytics and resilient human-environment systems. *Full list: Google Scholar*
  - Delivered 10+ presentations at top-tier international conferences.
  - Served as Guest Editor for Applied Science Special Issue "Geospatial Data Processing, Mining and Application".
  - Reviewer of 12+ international journals, delivering 40+ review reports

SKILLS

- **Core Data Science & AI**
  - *ML & DL*: PyTorch, TensorFlow, Hugging Face Transformers, Scikit-Learn, NumPy, Pandas, CUDA, AWS
  - *Time-Series Analysis & Forecasting*: ARIMA/SARIMA, Prophet, LSTM/GRU, Bayesian, Anomaly Detection
  - *Spatial-Temporal Modeling*: Spatio-Temporal Regression, Spatial Panel Data Models.
  - *NLP*: Sentiment Analysis, Topic Modeling, LLMs (BERT, Hugging Face).
  - *Math & Stats*: Probability, Linear Algebra, Statistical Inference
  - *Graph & Network Science*: Graph-Based Modeling, Centrality Analysis, Modularity Detection, Random Walks.
- **Programming & Tools**
  - *Languages*: Python (10+ yrs), R, SQL, C#, Java
  - *Databases & Platforms*: Jupyter, MongoDB, Oracle
  - *GIS & Visualization*: ArcGIS, QGIS, Gephi, Origin.