

Yan-Ting (Vicky) Liao

yliao1@hotmail.com ♦ 480-2890869 ♦ www.linkedin.com/in/vicky-liao/ ♦ github.com/vickyting0910 ♦ Richardson, Texas

Ph.D. Candidate in Geospatial Information Sciences(GIS), Specialty in Spatial Statistics

CORE COMPETENCY

- **Programming:** Python (Tensorflow, PyTorch, PySpark, Scikit-Learn, Numpy, Scipy, Statsmodels, PySAL, pandas, NLTK, re, Pystan, etc.), R (Spatstat, Gstat, etc.), SQL, Matlab, SPSS, Excel, AWS (Lambda, etc.), WinBUGS, etc.
- **Statistical Modeling:** Generalized Linear Model, Generalized Linear Mixed Model (GLMM), Multiple Imputation, Exploratory Spatial Data Analysis (ESDA), Spatial Analysis, Time-Series Analysis, Dynamic Time Warping (DTW), Bayesian Analysis, Principal Component Analysis (PCA), Factor Analysis (FA), Cross-Decomposition Methods (e.g., Canonical Correlation Analysis (CCA)), Stepwise Regression/LASSO/Ridge/Elastic-Net, segmentation, A/B testing.
- **Machine Learning Algorithms:** Decision Tree, Random Forest, Neural Network Models, Support Vector Machines.
- **Database and Deployment:** PostgreSQL, Amazon Web Service (Dynamodb, S3, Redshift), Azure (Datalake Store)
- **Data Visualization Tools:** D3.js, three.js, Tableau, ggplot, matplotlib, CorelDRAW, Ulead PhotoImpact

WORK EXPERIENCE

Graduate Teaching Assistant - The University of Texas at Dallas (UT Dallas) September 2015 – Present

- Led 2 team members for 3D visualization and won high appreciation by the National Geospatial-Intelligence Agency.
- Convinced three agencies to obtain fine-resolution, multi-sourced datasets (around 1TB big data) for experiments.
- Pointed out insufficiencies of the most influential cause for imputation-based measurement error in Econometrics.
- Developed experiments to pioneeringly assess impacts of using spatial imputations (e.g., geostatistics) in regression.
- Experimented how spatial sampling strategies (e.g., stratified random sampling) influence imputations and inference.
- Propose imputations initially to deal with missing data, expected better than the most widely used multiple imputation.

National Water Center Research Fellow - National Oceanic & Atmospheric Administration June 2016 – July 2016

- Formulated a new moisture index and applied filters for detecting flooding extents with 5% and 10% improvements.

Graduate Research Associate for NIJ Project - University of Oklahoma April 2014 – September 2014

- Refined geocoding methods by text matching, 20% more match rates, and 80% less time, over ESRI geocoder.
- Detected crime patterns by DTW (10TB GPS tracking data) and wrote SQL to extract data from a PostgreSQL database.

SELF-LEARNING PROJECTS

Developing Spatial-Temporal Matching Algorithms to Predict Crime Incidents June 2018 – August 2018

- Predicted 11 crimes (705,874 incidents) by parsing Zillow, Yellow Page, Yelp, and Google Maps by 484 features.
- Developed matching algorithms, incorporating cross-decomposition methods, increasing adjusted R^2 by 0.5.
- Implemented feature matching algorithms, integrating GLMM with elastic net, to reduce the influences of missing data.
- Designed matching algorithms, extended from Multivariate Adaptive Regression Splines, for lagged incidents.

Comparisons of Model Specifications to Enhance Bayesian Prediction October 2017 – November 2017

- Adapted Bernardinelli (1995) Spatiotemporal Bayesian model for evaluating specifications on predictions.

Multivariate Spatial Crime Analysis March 2017 – April 2017

- Extended PCA with variance inflation factor to improve modeling performance and handling multicollinearity.
- Implemented the Eigenvector Spatial Filtering in Python to handle spatial autocorrelation in residuals.

EDUCATION

Ph.D. in GIS (Spatial Statistics) UT Dallas, Richardson, TX August 2015 – May 2020 (expected)

Master in Geography, Arizona State University, Tempe, Arizona August 2011 – December 2013

- Developed segmentation by random walker for plants, superior to the most widely used software, eCognition.

Master in Geography, National Taiwan Normal University, Taipei, Taiwan September 2008–June 2010

- Managed databases for spatiotemporal data from over 400 references across 50 years for predictions.
- Pioneered predictions of long-term landscape changes by PCA, FA, and ESDA.