

EDUCATION

- **Iowa State University** Ames, IA
Doctor of Philosophy in Statistics; GPA: 3.93/4.00 Aug. 2017 – Expected May. 2022
- **University of Science and Technology of China** Hefei, China
Bachelor of Science in Statistics; Aug. 2013 – July. 2017

EXPERIENCE

- **Biostatistics and Research Decision Sciences, Merck & Co.** Rahway, NJ
Biostatistics Graduate Intern May. 2021 – Aug. 2021
 - Applied **Bayesian Shrinkage method** to conduct **subgroup analysis** on a clinical data and analyzed reduction rate of the targeted treatment to placebo at four different regions at both aggregate level and patient level.
 - Under patient level, built up a **Extend Dixon-Simon Model** using **hierarchical autoregression** with **R** to include correlation of records at different time.
 - Achieved more consistent reduction rates at four regions and narrower confidence intervals with significant margin of error reduction compared with no shrinkage method.
- **Center for Survey Statistics and Methodology, Iowa State University** Ames, IA
Research Assistant May. 2018 – Present
 - Research Interests: Survey Statistics, High-dimensional data analysis, Generalized Linear Mixed-effects Model, Graphical Model, Machine Learning.
- **Department of Statistics, Iowa State University** Ames, IA
Graduate Teaching Assistant Jul. 2017 – May. 2018
 - Assisted with STAT 231 (Probability and Statistical Inference for Engineers), STAT 322 (Probabilistic Methods for Electrical Engineers) and STAT 342 (Introduction to the Theory of Probability and Statistics II).

PROJECTS

- **High-dimensional Mixed Graphical Network** Jan. 2020 – Present
 - Provided a **high-dimensional mixed graphical model** containing both continuous and categorical variables.
 - Established the **consistency of graph reconstruction** under complex survey sample designs.
 - Developed **design-based BIC** for **neighbour selection** with **group lasso** to recover the true neighbourhood.
 - Optimized **survey pseudo composite likelihood** with **coordinate gradient descent** to estimate edge parameters.
- **Road Change Detection** August. 2020 – Present
 - Applied convolutional neural network model, **AD-LinkNet**, to detect roads on satellite images based on **PyTorch** and achieved **0.8** dice score with models trained on **30,000** images.
 - Utilized **transfer learning** to reduce training time and **hyperparameter optimization** to turn hyperparameters.
 - Applied **RoadTracer** algorithm to extract road graph and developed a **Hidden Markov Random Field** model to detect road map change based on **AD-LinkNet** outputs.
- **Shiny App Development For National Resource Inventory** Jun. 2020 – Aug. 2020
 - Designed a **Shiny App** using **R** and **Javascript**.
 - Realized **web scrabing** satellite image, searching point level land cover type, calculating each land cover area and showing the curve of growth trend under county/state level.
- **Mixture Responses for Small Area Estimation** Jan. 2019 – Dec. 2019
 - Created a **multivariate mixed-effects model** containing mixed types of response variables for **informative sampled data**.
 - Adopted **Monte Carlo EM algorithm** with **importance sampling** to estimate model parameters
 - Applied **empirical best predictor** to estimate small area parameters and **bootstrap** approach to estimate MSE.
 - Reduced MSE by **32 ~ 93%** compared with univariate independent models if the response variables are conditionally correlated.

PUBLICATIONS

- **Hao Sun**, Berg, E. and Zhu, Z. (2021). “Bivariate Small Area Estimation for Binary and Gaussian Variables Based on a Conditionally Specified Model”. *Biometrics*. DOI: 10.1111/biom.13552.

WORKING PAPERS

- **Hao Sun**, Berg, E. and Zhu, Z. (2022+). “High-dimension Mixed Graphical Model Under Complex Survey Design”. manuscript in preparation for *Biometrika*.

CONTRIBUTED TALKS & POSTERS

- **Hao Sun**, Hua, J., Li, Q. and Kaur, A. “Treatment Effects Across Subgroups Based on Shrinkage Estimation”. In: Joint Statistical Meetings, Virtual, August 2021.
- **Hao Sun**, Berg, E. and Zhu, Z. “Edge Selection for Graphical Models with Mixed Types under Informative Sampling”. In: Statistics in the Era of Evidence Based Inference (IISA 2021 Conference), Virtual, May 2021.
- **Hao Sun**, Berg, E. and Zhu, Z. “Joint Small Area Estimation for Categorical and Continuous Response Variables Based on a Conditionally Specified Model”. In: Joint Statistical Meetings, Virtual, August 2020.
- **Hao Sun**, Berg, E. and Zhu, Z. “Small Area Estimation for an Informative Sample Design”. In: Joint Statistical Meetings, Virtual, August 2020.
- **Hao Sun**, Dutta, S. “A Penalized H-Likelihood Method for Gaussian Spatial Additive Model on Regular Lattice”. In: Joint Statistical Meetings, Denver, Colorado, August 2019.

ACHIEVEMENTS AND AWARDS

- **Joint Statistical Meeting Student Travel Award, 2020**: Supported for Joint Statistical Meeting 2020 in Philadelphia by American Statistical Association for excellent research work.
- **Holly and Beth Fryer Scholarship, 2019**: Superior academic achievement by 2nd year graduate student.
- **Kemphorne Award in Statistics, 2018**: Presented annually to recognize outstanding performance in linear models.
- **Outstanding Student Scholarship, 2016**: Presented to top 10% students.

PROGRAMMING AND SKILLS

- **Programming**: R, Python, Matlab, C, SAS, SQL, Jupyter Notebooks, Linux
- **R Packages**: tidyverse, ggplot2, plotly, Leaflet, shiny, Rcpp, rstan, caret, glmnet