BORA JIN

+1 919 448 4611 bora.jin@duke.edu

Department of Statistical Science, Duke University, Durham, NC 27708

RESEARCH INTERESTS

Environmental health, Spatial statistics, Multivariate data, Hierarchical models, Latent variables, Bayesian methods

EDUCATION

Ph.D. Candidate in Statistical Science **2018 – 2022 (expected)**
Duke University Durham, NC, USA
Advisors: Amy H. Herring, David Dunson

Master of Applied Statistics **2015 – 2017**
Korea University, 4.25 Grade Point Average (4.5 Scale) Seoul, SOUTH KOREA

Bachelor of Economics in Statistics **2011 – 2015**
Korea University, 4.3 Grade Point Average (4.5 Scale) Seoul, SOUTH KOREA

RESEARCH EXPERIENCE

Bag of DAGs: Flexible & Scalable Modelling of Spatiotemporal Dependence **Present**
with Peruzzi, M., Johndrow, J.E., and Dunson, D.

- Propose a computationally efficient approach to construct a well-defined spatial Gaussian process (GP) with the nonstationary covariance using multiple yet simple directed acyclic graphs (DAGs), which leads to computational efficiency, flexibility, and interpretability in point-referenced geostatistical models.
- Develop Bag of DAGs-Gaussian process (BDAG-GP), each DAG of which is chosen to represent a different possible dependence structure, to induce nonstationarity.
- Analyze spatiotemporal variability of fine particulate matter (PM2.5) in California in which a DAG represents a prevailing wind direction causing some associated covariance in the pollutants.

Scalable Gaussian Processes on Physically Constrained Domains **Present**
with Herring, A.H. and Dunson, D.

- Motivated by applications in point-referenced geostatistics that have measurements collected and meaningful only within a constrained domain.
- Develop the Barrier Overlap-Removal Acyclic directed graph GP (BORA-GP), a scalable GP method that incorporates the constrained domain via sparsity-inducing DAGs.
- Enable characterization of dependence in constrained domains by removing an edge in a DAG if a linear path between two points overlaps physical barriers.
- Analyze levels of chlorophyll a along the east coast of the United States.

Bayesian Matrix Completion for Hypothesis Testing **2019 – 2020**
with Dunson, D., Rager, J.E., Reif, D., Engel, S.M., Herring, A.H.

- Adapt Bayesian heteroscedastic nonparametric regression to a multiple hypothesis testing framework.
- Impose a generalized latent factor model to form a non-exchangeable prior for testing.
- Develop a matrix completion method for a latent matrix.
- Tackle sparsity of the ToxCast data using hierarchical framework.
- Enable prediction for non-tested chemical's activity.
- Broaden the definition of activity including heteroscedasticity.

Master's thesis on Bayesian Methods **March 2015 – February 2017**
Korea University Seoul, SOUTH KOREA

- Applied Bayesian inference for a seemingly unrelated regression model and examined novel statistical methods on an extended instrumental variables model with random effects using the MCMC method.
- Employed the extension to a nonparametric model using cosine basis functions and the Dirichlet process location-scale mixture for a great deal of flexibility on the proposed model.
- Conducted a real data application explaining Annex I Parties variations in compliance to the Kyoto Protocol.

Prediction of Carbon Emissions in Industrial Setting **June 2016 – August 2016**
Korean Environment Ministry Seoul, SOUTH KOREA

- Trained toward a generalist dealing with global environmental issues in both national and international settings through the 8th International Environmental Expert Training Program.
- Applied statistical prediction analysis in studying afforestation practices within industrial complex areas.
- Developed strategic programs to implement carbon reduction targets of industrial complexes.

Clustering of Cancer Patients' Symptoms **March 2016 – August 2016**
College of Nursing at Chungnam National University Seoul, SOUTH KOREA

- Conducted clustering analysis to see if symptoms are divided into particular clusters as expected in a theory.
- Measured the effect of symptom clusters on the quality of life through physical functions.
- Provided expertise on all aspects from the interpretation of statistical results to the presentation of statistics and graphics.

Emotionality of Language in Online Platforms **October 2014 – March 2015**
MezzoMedia & SungKyunKwan University Seoul, SOUTH KOREA

- Contributed to the development of a Korean morphological analyser.
- Embedded factor analysis and principal component analysis to determine appropriate weights of frequently used morphemes based on their verbal and social context.
- Designed and interpreted quantitative research examining the emotionality of language in online platforms.

WORK EXPERIENCE

ENVIRONMENTAL STATISTICS

Internship in Chemicals and Waste **February 2017 – August 2017**
UN Environment Geneva, SWITZERLAND

- Managed national reports that Parties are obliged to submit under Basel and Stockholm Conventions and analysed national reporting data so as to identify regional patterns and temporal trends.
- Designed a query system to facilitate proper visualizations and the use of collected data from national reports.
- Participated in the formulation of indicators with regards to Sustainable Development Goals.

BIG DATA & FINANCE

Internship in the Division of Budget and Finance **January 2018 – July 2018**
International Atomic Energy Agency Vienna, AUSTRIA

- Managed data compliance of requests for all types of procurement and payment in the Agency.
- Analysed collected data as a member of Master Data Management Team.
- Assisted a clean-up project of the Agency's bank and branch pages through Oracle sql and MS Access.
- Published monthly infographics on activities of the whole division.

Internship in Banking and Finance, Market Surveillance Department **June 2014 – July 2014**
Korea Exchange (KRX) Seoul, SOUTH KOREA

- Focused upon the detection of unfair transactions, particularly with regards to high turnover volume accounts.
- Conducted data management including updating, arranging and organising big data.
- Participated in developing case-specific restrictions and market-wide regulatory practices.

TEACHING EXPERIENCE

Teaching Assistant **Present**
Duke University Durham, NC, USA
Case Studies in the Practice of Statistics (STA440)

Graduate Mentor **May 2021**
Duke University Online

- Intro to Undergraduate Research in Statistical Science Workshop
- Mentored a team of five undergraduates and facilitated their collaboration on identifying research questions, conducting relevant statistical analyses, writing a report in a common structure of statistics/medical journals, and presenting results for statistical audience.

Head Teaching Assistant **January 2021 – May 2021**
Duke University Online

- Theory and Methods of Statistical Learning and Inference (STA432)

Guest Lecture **April 28, 2021**
Harvard University Online

- Spatial Statistics (STAT141)

Teaching Assistant **January 2019 – May 2019**
Duke University Durham, NC, USA

- Statistics (STA250)

Teaching Assistant **September 2015 – December 2015**
Korea University Seoul, SOUTH KOREA

- Introduction to Probability Theory (STAT201)
- Topics in Mathematical Statistics (STAT412)

Mentoring in Mathematics **January 2013**
Samsung Seoul, SOUTH KOREA

KEY SKILLS

- R, MATLAB, Python, SAS, MySQL, MS Office, LATEX
- English (Professional Proficiency), Korean (Native)

HONOURS

Best Student/Postdoc Contributed Paper Award at the ISBA 2021	2021
NC ASA Student Travel Awards	2021
Global Korea Scholarship	2018 – 2020
First Prize in the Graduate Paper Session at the Korean Statistical Society's Annual Conference	2016
Yangcheon Foundation Scholarship for Academic Excellence	2016
So-Mang Presbyterian Church Scholarship for Academic Merit	2016
Second Prize in the Graduate Poster Session at the Korean Statistical Society's Annual Conference	2015
The Dean's Award for Academic Merit	2011 – 2014
Ministry of Gender Equality and Family Affairs Minister's Honour Award	2012
Seoul National University President's Prize	2011

PUBLICATIONS

Jin, B.*, Dunson, D., Rager, J.E., Reif, D., Engel, S.M., Herring, A.H. (2021+). **Bayesian Matrix Completion for Hypothesis Testing**. *submitted* ArXiv

WORKING PAPERS

Jin, B.*, Peruzzi, M., Johndrow, J.E., Dunson, D. **Bag of DAGs: Flexible & Scalable Modelling of Spatiotemporal Dependence**

Jin, B.*, Herring, A.H., Dunson, D. **Scalable Gaussian Processes on Physically Constrained Domains**

TALKS & POSTERS

Scalable Gaussian Processes on Physically Constrained Domains. *New Methods session at Bayesian Young Statisticians Meeting*, Online, Sep 2021.

Bag of DAGs: Flexible & Scalable Modelling of Spatiotemporal Dependence. *Contributed speed session at Joint Statistical Meetings*, Online, August 2021.

Bag of DAGs: Flexible & Scalable Modelling of Spatiotemporal Dependence. *Contributed session at World Meeting of the International Society for Bayesian Analysis*, Online, July 2021.

Scalable Gaussian Processes on Physically Constrained Domains. *Poster at The International Environmetrics Society – GRASPA Conference*, Online, June 2021.

Scalable Gaussian Processes on Physically Constrained Domains. *Lightning presentation at Symposium on Data Science and Statistics*, Online, June 2021.

Bayesian Inference to Multiple Equations in Seemingly Unrelated Regression Framework. *Graduate paper session at the Korean Statistical Society's Annual Conference*, November 2016.

Bayesian Inference to Multiple Equations in Seemingly Unrelated Regression Framework. *At Korea University – Hokkaido University Joint Conference in Statistics*, June 2016.

Bayesian Approaches to Instrumental Variable Models with Multiple Endogenous Regressors. *Graduate poster session at the Korean Statistical Society's Annual Conference*, November 2015.