David A. Buch

Contact Information	Department of Statistical Science Duke University Durham, NC 27705	*Phone:* (304) 860-5895 *Email:* david.buch@duke.edu *Website:* davidbuch.github.io

Quantitative and Technical Skills

- Relevant coursework *Time Series and Dynamic Models, Bayesian Forecasting, Multivariate Statistics, Linear Models, Probability and Measure Theory, Machine Learning, Bayesian Nonparametrics, Numerical Analysis, Linear Algebra*
- Extensive experience with R, C/C++, and git
- Experience with Python, Matlab, Linux/Unix command line tools, and cluster computing
- Experience in collaborative software development, project management, and teaching

Education

Duke University, Durham, NC
Ph.D. Statistical Science (GPA 3.94/4) **2019-**

West Virginia University, Morgantown, WV
M.S., Applied Mathematics (GPA 3.9/4) **2018-2019**
B.S., Mathematics, Physics - Honors College (GPA 3.9/4) **2014-2018**

Selected Honors and Awards

- *Citadel Data Open Championship 2021*: 1st Place Team ($100,000 Prize); *Citadel East Coast Regional Data Open 2021*: 3rd Place Team ($2,500 Prize); *USC, Big Data Health Science Competition 2021*: 5th Place Team
- Duke Statistics: Honorable Mention TA of the Year 2020-2021
- West Virginia University Outstanding Senior 2018
- United States Presidential Scholar 2014, National Merit Scholar 2014

Research Activities

A Semi-Parametric Mixed-Effects SIR Model **In preparation**
D.A. Buch, J. Johndrow, and D. B. Dunson.
Lead researcher, charged with task of using epidemic incidence data to learn about the impact of spatiotemporal covariates on the rate of spread of an infectious disease (e.g. using daily deaths to understand how air quality affects transmission rate). Skills: **Gaussian Process models with non-gaussian likelihood, mechanistic/ODE models, adaptive Metropolis algorithms**. Language: **R/C++, Matlab**.

A Taxonomy of Statistical Sins: Catalogued and annotated failure as a tool for success in Data Science **In preparation**
D.A. Buch, D. B. Dunson, and A.H. Herring.
Lead researcher, became personally interested in systematizing "bad" applied statistics. Reviewed dozens of student projects and synthesized other reviewer feedback to develop a taxonomy documenting and explaining common analytic errors.

Transforming energy access to real demand: evaluating the impact of information provision on rural electricity consumption **In preparation**
Z. Han, **D.A. Buch**, and T. R. Fetter.
Following role as Summer 2020 **project manager** for the Duke Data+ program, mentored economics masters degree student Zhihao Han, who used archival data to

investigate the impact of a New Deal era program on residential electricity consumption in the rural United States.

Predicting Transcription Factor Binding Affinity **Ongoing research**
D.A. Buch, R. Gordan, and D. B. Dunson.
Lead researcher, charged with developing an **interpretable model that accurately predicts binding affinity** of certain transmission factors to various DNA sequences. Skills: **high-dimensional regression, categorical data analysis**. Language: **R**.

Tie Strength, Optimal Connections, and Distance in Social Networks **2019**
D.A. Buch. WVU Master's Thesis Article. https://doi.org/10.33915/etd.7400
Masters research assistant, charged with analyzing social network data for a STEM retention program. Developed and implemented descriptive statistics for network connectedness. Skills: **shortest path algorithms, the graph laplacian form, and resistor array simplification**. Language: **R/C++**.

Improving performance of SEOBNRv3 by $\sim 300x$ **2018**
T.D. Knowles, C. Devine, **D.A. Buch**, S.A. Bilgili, T.R. Adams, Z.B. Etienne, and S.T. McWilliams. Classical and Quantum Gravity. 35: 15
Undergraduate research assistant, **optimized software written by members of the international LIGO scientific collaborative**. Efficient computations were essential to an MCMC algorithm used to draw inferences from gravitational wave data. Language: **C**.

Collective repacking reveals that the structures of protein cores are uniquely specified by steric repulsive interactions **2017**
J.C. Gaines, A. Virrueta, **D.A. Buch**, S.J. Fleishman, C.S. O'Hern, and L. Regan. Protein Eng. Des. Sel. 30: 387
Summer research intern at Yale University, reviewed the literature of protein folding software and developed pipeline to compare the performance of various fold prediction strategies. Language: **Python**.

Software Experience	*smeSIR* (R/C++) **Beta testing**

Software Experience

smeSIR (R/C++) **Beta testing**
Individual project, user friendly implementation of the method developed in "A Semi-Parametric Mixed-Effects SIR Model." R package available on github.

Phenom Microbial Growth Model (Python Web App) **Beta testing**
Served as **project manager** for team of Duke undergraduate students, Summer 2021.

U.K. Football Analysis (R Shiny App) **2021**
Individual project, created as part of larger team effort in the 2021 Citadel East Coast Regional Datathon. Web app published on RShiny server.

Gretel (R/C++) **2019**
Individual project, implementation of the methods developed in "Tie Strength, Optimal Connections, and Distance in Social Networks." R package available on CRAN.

LALSimulation SEOBNRv3 PDE Solver (C) **2016-2018**
International collaboration, as discussed in "Improving performance of SEOBNRv3 by $\sim 300x$". Scientific software used by LIGO.