Brock Stewart, PhD

brock@brockstewart.com • BrockStewart.com • HyLown.com • PowerAndSampleSize.com • Publications

Seeking a Senior Data Scientist position.

EXPERIENCE

Senior Mathematical Statistician **2018 – present**

Mathematical Statistician **2010 – 2018**

US Centers for Disease Control & Prevention (CDC) **Atlanta, GA**

Deliverables

- **30** published, peer-reviewed scientific manuscripts
- **12** internal manuscripts and technical reports
- **12** scientific presentations
- **10** technical presentations on, e.g., R, R Shiny, Rmarkdown, interfacing C and R, LaTeX, Bayesian statistics
- **7** software products using **R**, **Python**, **C**, **SQL**, **SAS**, **LaTeX**
- Statistical consulting and technical assistance on 100+ scientific projects as a non-author
- Routinely performed statistical analyses, gave critical review of analyses, and assisted others in analyses using methods including logistic and Poisson regression, survival analysis, linear mixed models, exact inferential methods, nonparametric statistical tests, zero-inflated and hurdle models, segmented regression models, generalized liner models, and generalized estimating equations.

Professional Activities

- **Subject matter expert** in statistics & data mining for internal/external workgroups and mid-level leadership.
- **Chaired** a 52-member statistical workgroup consisting of internal and external partners.
- **Statistical Advisory Group**: Active 2015–2018; Chaired the annual Statistics Day 2015, 2016.
- **R User Group**: Founding member (2013), Chair (2014), active leader (2015–present), 300+ members.
- **R course**: taught a 4-day course to CDC staff with three other statisticians.
- **Invited Reviewer** for **6** scientific journals.

Selected Project Activities

- I developed a **web app** written in **Python** and **R** to compute sample size & CI width for multi-stage, stratified cluster surveys; outputs tables, graphs, summary statements suitable for use in protocols/manuscripts. (GitHub)
- I **designed a data storage model**, implemented with **SQL**, that reduced storage size **90%** and query runtimes **50%**.
- I **invented A/B Testing algorithms** written in **C** (optional interfaces in **R** & a stand-alone DLL) for fast computation of Type I error rate under Binomial & Poisson designs; also computes exact-test critical value, statistical power, & lists decision-making data points. **Runtime < 1 – 3 seconds** compared to previous software running overnight-to-days. (poster)
- I used **empirical Bayesian data mining** on a database of ~2M records which gains ~3k records/month to perform a recurring search for rare-event associations between ~150 candidates and any of ~70k potential outcomes. I used SQL for data preparation and used numerical optimization functions in R to maximize the marginal likelihoods. Example uses: HPV, HepB, Flu, Hib
- I used **Bayesian Model Averaging** for prediction, variable selection/importance and to assess model uncertainty. (slides)
- In a peer-reviewed scientific manuscript, co-authors and I demonstrated an interesting contrast between a frequentist and a **Bayesian** method for sequential analysis of surveillance data; namely, that in the setting of vaccine adverse event sequential testing, the frequentist approach had more favorable false-positive rates, while the Bayesian approach consistently provided quicker time-to-signal and lower false-negative rates. We wrote **R** code to perform the simulations and analyses for this project.

Awards and Accolades

- **Two-time winner** of **BEST THEORETICAL PAPER**, CDC's Statistical Science Award for Statistical Excellence:
  Continuous Sequential Boundaries for Vaccine Safety Surveillance Stat in Med (33):19
  A Bayesian Approach to Sequential Analysis in Post-licensure Vaccine Safety Surveillance Pharm Stat (19):3
- **HONOR AWARD**, National Center for Emerging and Zoonotic Infectious Diseases (NCEZID), for:
  Excellence in rapidly assessing the safety of Tdap vaccination in pregnant women.
- **CERTIFICATE OF APPRECIATION**, CDC Division of Tuberculosis Elimination, for:
  Exemplary collaboration & innovation in quantitative science to respond to the leveling of Tuberculosis incidence.
- **DIRECTOR'S SPECIAL ACT IN RECOGNITION**, NCEZID/Division of Healthcare Quality Promotion, for:
  Development of innovative methods to aid Rapid Cycle Analysis vaccine safety surveillance activities.

Postdoctoral Research Associate, Graduate Researcher **2002 – 2010**
University of Georgia **Athens, GA**

- **6** published scientific manuscripts; **12** scientific presentations.
- Published a paper examining the predictive abilities of **non-linear regression predictive models** used in forestry.
- Published a paper presenting **Bayesian predictive models** of tree height.
- Developed a new class of forest growth estimators and examined their statistical efficiency on simulated forests I generated using Markov Chain Monte Carlo (MCMC) methods with code I wrote in **R** and **C**.
- Developed a novel system using cameras for ground-based tree measurements, accounting for perspective distortion, using software I wrote in Visual BASIC. I performed data management in SAS.
- I used non-linear mixed-effects models to predict treatment response of forest growth using SAS.
- Wrote SAS MACROs to prepare summary tables & graphical output from large linked databases for a web app.

EDUCATION

University of Georgia **2001, 2004, 2007, 2007**

2007 **PhD** Forestry	Dissertation: *Sampling efficiencies of new and existing variable radius plot growth estimators*. Awarded the Jerome L. Clutter Memorial **Fellowship** for outstanding student research. Inducted into the Xi Sigma Pi International Forestry **Honor Society**.
2007 **MS** Statistics	Coursework including Probability Theory *I & II*, Mathematical Statistics *I & II*, Non-linear Regression, Time Series Analysis, Numerical Computing.
2004 **MS** Forestry	Thesis: *Using a camera as an angle gauge in angle-count sampling*.
2001 **BS** Forestry	With a focus on quantitative decision making.

PERSONAL INTERESTS

- PowerAndSampleSize.com **8k** unique visitors/month. Cited in **1,100+** published scientific manuscripts/books.
- **Git & Linux**: I setup **2 personal, private git servers** on: Debian\Linux at home; my webhost's server (RedHat).
- HyLown.com I setup a home HTTP server, GitWeb, a free VPN (no 3rd party), and network-wide ad blocking.
- Ethereal to Make Docs A framework using **Python** to convert a collection of Markdown files into a website
- Faith, Family, Farming, Fabricating, and Fishing