

William Michael Landau

Education

- PhD Statistics, [Iowa State U](#), 2016
- MS Statistics, [Iowa State U](#) 2013
- BS Math, [U of Chicago](#) 2011

Contact

- will.landau@gmail.com
- wlandau.github.io
- linkedin.com/in/wlandau
- github.com/wlandau

Publications

- Niemi, J., Mittman, E., **Landau, W.**, and Nettleton, D. (2015), "Empirical Bayes Analysis of RNA-seq Data for Detection of Gene Expression Heterosis," *Journal of Agricultural, Biological, and Environmental Statistics*, 20, 1-15. Available at link.springer.com.
- **Landau, W.** and Liu, P. (2013), "Dispersion Estimation and Its Effect on Test Performance in RNA-seq Data Analysis: A Simulation-Based Comparison of Methods," *PLOS One*, 8. Available at journals.plos.org.
- Ratliff, B., Womack, C., Tang, X., **Landau, W.**, Butler, L., and Szpunar, D. (2010), "Modeling the Rovibrationally Excited C₂H₄OH Radicals from the Photodissociation of 2-Bromoethanol at 193 nm," *Journal of Physical Chemistry*, 114, 4934-4945. Available at ncbi.nlm.nih.gov.

Articles under review

- **Landau, W.**, Niemi, J., and Nettleton, D., "Fully Bayesian analysis of RNA-seq counts for the detection of gene expression heterosis", submitted to the *Journal of the American Statistical Association* on June 24, 2016.
- **Landau, W.** and Niemi, J., "A fully Bayesian strategy for high-dimensional hierarchical modeling using massively parallel computing," submitted to the *Journal of Computational and Graphical Statistics* on March 8, 2016. Preprint available at arxiv.org.

Open Source Software

- [drake](#), a solution in R for reproducibility and high-performance computing. Currently [under review](#) by [rOpenSci](#).
- [downsize](#), an R package to toggle between the test and production versions of large workflows.
- Infrastructure-related R package utilities [epl](#), [grapes](#), and [wildcard](#).

Awards

- 2017 Lilly Innovator Award. During my first year at Lilly, I created and led a successful campaign to transform the company's open source software policies and procedures.
- [Student Paper Award](#), American Statistical Association Section on Statistical Computing, Jan 2016. Awarded for an early draft of the preprint at arxiv.org/abs/1606.06659.
- [Vince Sposito Statistical Computing Award](#), Iowa State University, Aug 2013.
- [GlaxoSmithKline Industrial Scholarship](#), Iowa State University, Sep 2011.
- Alumni Scholarship, Iowa State University, Aug 2011.

Skills

- Reproducible research, hierarchical models, Bayesian methods, Markov chain Monte Carlo, statistical computing, high-dimensional data analysis, genomics data analysis, exploratory analysis, visualization, linear and nonlinear models, data mining, machine learning, predictive modeling, multivariate analysis.
- High-performance computing, R, R package development, general-purpose graphics processing unit (GPU) computing, CUDA, shell scripting, LaTeX, HTML, CSS.
- Past experience with C/C++, MPI, OpenMP, Python, JavaScript, AWK, Fortran.

Research scientist

- **October 2016 - Present**
- [Eli Lilly and Company](#)
- Advanced Analytics Immunology Hub
- Used tree-based algorithms to help tailor new treatments of autoimmune diseases to patients.
- Used modeling and simulation to help teams optimize clinical trial pipelines.
- Created interactive R-powered web applications to deliver cogent real-time insights to clinical teams.
- Created [drake](#), an R package for reproducibility and high-performance computing. Currently [under review](#) by [rOpenSci](#).

Research assistant

- **May 2013 - Aug 2016**
- RNA-sequencing Working Group, Department of Statistics, [Iowa State University](#).
- Funded by NIH grant R01GM109458 with Drs. Dan Nettleton and Jarad Niemi.
- Developed a new fully Bayesian analysis method for high-dimensional genomic datasets using hierarchical models.
- Implemented massively parallelized Markov chain Monte Carlo.

- Created the [fbseq](#) R package to distribute the analysis method.
- Implemented and distributed parallel computing backends for [CUDA GPUs \(fbseqCUDA\)](#) and [OpenMP \(fbseqOpenMP\)](#).
- Created the [remakeGenerator](#), [parallelRemake](#), and [downsize](#) packages to manage, ameliorate, expedite, and accelerate computationally heavy reproducible workflows that are under heavy development.

Seminar instructor

- **Aug - Dec, 2012 and 2013.**
- Department of Statistics, [Iowa State University](#).
- [GPU computing](#) seminar series at [wlandau.github.io/gpu](#).
- Educated faculty and graduate students on massively parallel computing with [general-purpose graphics processing units](#).
- Constructed, curated, and distributed slides, video, and example code at [wlandau.github.io/gpu](#) and on [YouTube](#).

Course instructor

- **Jan - May, 2012 and 2013.**
- Department of Statistics, [Iowa State University](#).
- STAT 305: Engineering Statistics ([wlandau.github.io/stat305](#)).

Grader

- **Aug - Dec, 2011.**
- Department of Statistics, [Iowa State University](#).
- STAT 231: Engineering Probability.
- STAT 105: Introduction to Engineering Statistics.

Leadership at Iowa State University

- Founder and leader, Cloud Computing Working Group, Sep - Dec 2015.
- Member, Computation Advisory Committee, Sep 2015 - May 2016.
- Volunteer instructor, Office of Precollegiate Programs for Talented and Gifted ([OPPTAG](#)), Mar 13, 2014.
- Fellow, [Preparing Future Faculty](#), Aug 2013 - May 2014.
- Assistant Coach, Boxing Club, Aug 2013 - Dec 2013.

References

- [Jarad Niemi](#), PhD advisor and major professor, niemi@iastate.edu.
- [Dan Nettleton](#), lead principal investigator of the RNA-sequencing Working Group (Iowa State Department of Statistics), dnett@iastate.edu.
- [Peng Liu](#), MS advisor and major professor, pliu@iastate.edu.
- Additional references available on request.

Hobbies

- Sailing, windsurfing, climbing, CrossFit, Brazillian Jiu Jitsu.