

Collin A. Politsch

Carnegie Mellon University
Machine Learning Department
5000 Forbes Avenue
Pittsburgh, PA, United States 15213

July 19, 2021
capolitsch@cmu.edu
<https://collinpolitsch.com>

Research Interests

Machine Learning/Statistics: Massive spatial datasets, spatial modeling, distributed spatial models, time series analysis, signal processing, forecasting, data mining, nonparametric statistics, uncertainty quantification, high-dimensional statistics, statistical machine learning

Astrophysics: Astrostatistics and astroinformatics, cosmostatistics, nonparametric and data-driven astrophysics, Lyman- α forest, intergalactic medium, statistical cosmography, large-scale structure of the Universe, time-domain astronomy, planetary transits, spectroscopic classification and redshift estimation

Academic Positions

Carnegie Mellon University Pittsburgh, PA
Machine Learning Department 2020 - 2021
School of Computer Science
Postdoctoral Fellow, Supervisor: Ryan Tibshirani

Education

Carnegie Mellon University Pittsburgh, PA
Joint Ph.D. in Statistics and Machine Learning 2020
Dissertation: *Statistical Astrophysics: From Extrasolar Planets to the Large-scale Structure of the Universe*
Award: Umesh K. Gavaskar Memorial Award for Best Ph.D. Dissertation
Advisors: Larry Wasserman, Jessi Cisewski-Kehe, Rupert Croft

Carnegie Mellon University Pittsburgh, PA
M.Sc. in Machine Learning 2017
Thesis: *Exploring the Intergalactic Medium*
Advisors: Larry Wasserman, Jessi Cisewski-Kehe, Rupert Croft

University of Kansas Lawrence, KS
B.Sc. in Mathematics (With Honors), Minor in Latin 2014
Honors Thesis: *On Discrete-Time Linear Quadratic Control*
Advisor: Tyrone E. Duncan

Peer-Reviewed Publications

- Three-dimensional cosmography of the high redshift Universe using intergalactic absorption**
[Collin A. Politsch](#), Jessi Cisewski-Kehe, Rupert A.C. Croft, Larry Wasserman
In preparation. Pre-submission inquiry approved by *Nature*.
 - Trend Filtering – I. A Modern Statistical Tool for Astronomical Spectroscopy and Time-Domain Astronomy**
[Collin A. Politsch](#), Jessi Cisewski-Kehe, Rupert A.C. Croft, Larry Wasserman
Monthly Notices of the Royal Astronomical Society, Volume 492, Issue 3, March 2020, Pages 4005-4018.
[arXiv:1908.07151](#); [GitHub](#).
- * **Finalist for best paper in the 2020 ASA Astrostatistics Student Paper Competition**, sponsored by the Astrostatistics Interest Group.

3. **Trend Filtering – II. Denoising Astronomical Signals with Varying Degrees of Smoothness**
Collin A. Politsch, Jessi Cisewski-Kehe, Rupert A.C. Croft, Larry Wasserman
Monthly Notices of the Royal Astronomical Society, Volume 492, Issue 3, March 2020, Pages 4019-4032.
[arXiv:2001.03552](#); [GitHub](#).

* **Finalist for best paper in the 2020 ASA Astrostatistics Student Paper Competition**,
 sponsored by the Astrostatistics Interest Group.

4. **An Open Repository of Real-Time COVID-19 Indicators**

Alex Reinhart, Logan Brooks, Maria Jahja, Aaron Rumack, Jingjing Tang, Wael Al Saeed, Taylor Arnold, Amartya Basu, Jacob Bien, Ángel A. Cabrera, Andrew Chin, Eu Jing Chua, Brian Clark, Nat DeFries, Jodi Forlizzi, Samuel Gratz, Alden Green, George Haff, Robin Han, Addison J. Hu, Sangwon Hyun, Ananya Joshi, Jimi Kim, Andrew Kuznetsov, Wichada La Motte-Kerr, Yeon Jin Lee, Kenneth Lee, Zachary C. Lipton, Michael X. Liu, Lester Mackey, Kathryn Mazaitis, Daniel J. McDonald, Balasubramanian Narasimhan, Natalia L. Oliveira, Pratik Patil, Adam Pereri, Collin A. Politsch, Samyak Rajanala, Dawn Rucker, Nigam H. Shah, Vishnu Shankar, James Sharpnack, Dmitry Shemetov, Noah Simon, Vishakha Srivastava, Shuyi Tan, Robert Tibshirani, Elena Tuzhilina, Ana Karina Van Nortwick, Valérie Ventura, Larry Wasserman, Jeremy C. Weiss, Kristin Williams, Roni Rosenfeld, and Ryan J. Tibshirani
 Submitted to *Proceedings of the National Academy of Sciences*.

5. **Augmenting Adjusted Plus-Minus in Soccer with FIFA Ratings**

Francesca Matano, Lee F. Richardson, Taylor Pospisil, Collin A. Politsch, Jining Qin
 Submitted to *Journal of Quantitative Analysis in Sports*. [arXiv:1810.08032](#); [intraocular.net](#).

6. **Mapping the Large-scale Universe through Intergalactic Silhouettes**

Collin A. Politsch and Rupert A.C. Croft
CHANCE, Volume 32, Issue 3, Sep. 2019, Pages 14-19.

Competitions

Media: *CMU Statistics and Data Science Graduate Students Keep Winning Big*

- 2nd Place:** The Data Open, presented by Citadel and Correlation One 2018
 – 300+ applications, ~125 selected to compete for \$25,000 in prizes (and job offers)
- 2nd Place:** NBA Basketball Analytics Hackathon, New York, NY, hosted by the NBA 2017
 – 900+ applications, ~200 selected to compete for ~\$20,000 equivalent in tickets, etc.
 (and job offers)
- 2nd Place:** The Data Open, presented by Citadel and Correlation One 2017
 – 550+ applications, ~125 selected to compete for \$25,000 in prizes (and job offers)
- 3rd Place:** CMU BrainHub NeuroHackathon, sponsored by Google 2017
 – 51 CMU graduate students selected to compete for free tuition and travel stipends

Selected Talks and Posters

Invited

- *Three-dimensional cosmography of the high redshift Universe using intergalactic absorption*
 – *Into the Impossible With Brian Keating* (Journal Club Guest Lecture), To appear on his *Dr Brian Keating* YouTube channel, July 2021.
 – University of Chicago, Machine Learning in Complex Phenomena seminar, Chicago, IL. Feb. 2021.
 – University of Maryland, Department of Mathematics, College Park, MD. Nov. 2020.
 – Duke University, Department of Statistical Science, Durham, NC. Nov. 2020.
 – NSF AI Planning Institute: Physics of the Future (jointly hosted by STAMPS@CMU). Oct. 2020.
 – Flatiron Institute Center for Computational Astrophysics (jointly hosted by NYU). Oct. 2020.
 – Los Alamos National Laboratory, Los Alamos, NM. Oct. 2020.

- *Trend Filtering: A Modern Statistical Tool for Time-Domain Astronomy and Astronomical Spectroscopy*
 - NSF AI Planning Institute for Data-Driven Discovery in Physics @CMU, Pittsburgh, PA. Oct. 2019.
 - ASA Joint Statistical Meetings. Astrostatistics Interest Group: Best Student Paper Finalist. Aug. 2020.
- *From Mapping the Universe to Forecasting the Pandemic*
 - OnSolve Nexus 2021: Managing Uncertainty for Organizational Resiliency. April 2021.
- *A Multi-Resolution 3D Map of the Intergalactic Medium via the Lyman- α Forest*
 - Uber Technologies, Inc., San Francisco, CA. Aug. 2018.
 - ASA Joint Statistical Meetings, Baltimore, MD. July 2017. (Poster)
 - SAMSI, Cosmology Working Group Seminar Series, Durham, NC. Nov. 2016.

Contributed Conference Proceedings & Seminars

- *Three-dimensional cosmography of the high redshift Universe using intergalactic absorption*
 - ASA Joint Statistical Meetings. Session: *Statistical Challenges in Cosmology*. Aug. 2021.
- *A Multi-Resolution 3D Map of the Intergalactic Medium via the Lyman- α Forest*
 - SAMSI Astronomy Transition Workshop, Durham, NC. May 2017.
 - ASA Pittsburgh Chapter Banquet, Pittsburgh, PA. April 2017. (Poster)
- *Multi-resolution Regression, Divide and Conquer Risk Estimation, and the Large-scale Universe*
 - Carnegie Mellon University, Pittsburgh, PA. May 2017.
 - Statistical and Applied Mathematical Sciences Institute, Durham, NC. April 2017.
- *Exploring the Intergalactic Medium*
 - Carnegie Mellon University, Pittsburgh, PA. April 2017.
 - SAMSI Astronomy Opening Workshop, Durham, NC. Aug. 2016. (Poster)
- *Statistical Methods for Estimating Regression Quantiles*
 - CMU 10-725 Convex Optimization Symposium, Pittsburgh, PA. Dec. 2015. (Poster)

Experience

Carnegie Mellon University

Postdoctoral Fellow

08/2020 - 08/2021

Lab: The Delphi Research Group

PI/co-PI(s): Roni Rosenfeld, Ryan J. Tibshirani

Personal role: Lead of COVID-19 forecasting development and evaluation team

Official duties: Lead a team of faculty and Google software engineers and data scientists assembled to develop statistical models for forecasting the COVID-19 pandemic, as part of Delphi's COVID-19 response. Take the leading role in developing new forecasting methodology, ensure the group adheres to good programming and development practices, implement and test models, build supporting infrastructure, write documentation and notebooks and blog posts detailing the use cases for our tools, and our findings.

Graduate Research Assistant, McWilliams Center for Cosmology

01/2019 - 06/2020

Project: Intensity Mapping the Universe

PI: Rupert A.C. Croft

Funding: NASA ([Grant NNX17AK56G](#))

Graduate Research Assistant, Department of Statistics & Data Science

01/2015 - 08/2016

Project: Nonparametric Procedures that Exploit Structured Data and Models

PI/co-PI(s): Ann Lee, Chad Schafer, Shirley Ho

Funding: National Science Foundation ([Award #1521786](#))

Project: Exploring the Intergalactic Medium

Advisors: Larry Wasserman, Jessi Cisewski-Kehe, Rupert A.C. Croft

Funding: National Science Foundation ([Award #1043903](#))

Uber Technologies, Inc.

Data Scientist Intern

San Francisco, CA

06/2018 - 08/2018

Team: UberEverything Data Science

Project: A Holistic Approach to Uber Eats Home Feed Ranking Optimization

Description: Completed an end-to-end project which culminated in a new personalized ranking and recommendation algorithm for the Uber Eats iOS/Android home feed that showed significant improvement over the current ranking algorithm in both offline evaluation and online A/B testing, and was subsequently launched.

Association of Universities for Research in Astronomy Observatory

La Serena School for Data Science: Applied Tools for Astronomy

La Serena, Chile

08/2015

Project: Cosmology with the Cosmic Microwave Background Through Cross Correlations

Mentors: Jeffrey McMahon, Chris Miller

Funding: NSF ([Award #1637359](#)), MAS, CONICYT

North Carolina State University

Undergraduate Research Assistant

Raleigh, NC

05/2013 - 07/2013

Project: Portfolio Optimization with Conditional Value-at-Risk (CVaR)

PI: Tao Pang

Funding: NSF ([Award #1461148](#)), NSA

University of Kansas

Undergraduate Research Assistant

Lawrence, KS

01/2013 - 05/2014

Project: Optimal Control of Stochastic Systems Driven by Fractional Brownian Motions

PI/co-PI(s): Tyrone E. Duncan, Bozenna Pasik-Duncan

Funding: U.S. Army Research Office ([Contract W911NF-10-1-0248](#))

Project: Optimal and Adaptive Control of Stochastic Systems

PI/co-PI(s): Tyrone E. Duncan, Bozenna Pasik-Duncan

Funding: Air Force Office of Scientific Research ([Grant FA9550-09-1-0554](#))

Project: Control of Stochastic Systems

PI/co-PI(s): Tyrone E. Duncan, Bozenna Pasik-Duncan

Funding: National Science Foundation ([Award #1108884](#))

Teaching and Advising

Carnegie Mellon University

Graduate Teaching Assistant

01/2015 - 12/2018

- 10/36-702: *Statistical Machine Learning* Head TA, PhD course
- 10/36-705: *Intermediate Statistics* Head TA, PhD course
- 36-618: *Experimental Design & Time Series* Head TA, MS course
- 36-467/667: *Special Topics: Data over Space & Time* Head TA, MS course
- 36-401/607: *Modern Regression* Head TA, BS/MS course
- 36-402/608: *Advanced Methods for Data Analysis* BS/MS course
- 36-225: *Introduction to Probability Theory* Head TA, BS course
- 36-226: *Introduction to Statistical Inference* Head TA, BS course
- 36-217: *Probability Theory and Random Processes* Head TA, BS course

Lecturer

- **Summer Lecture Series**, Carnegie Mellon University, Summer Undergraduate Research Experience in Statistics, Pittsburgh, PA. *Introduction to Statistics in R*. June - July 2015.

- **Guest Lecture**, Carnegie Mellon University, STAT 217 (Probability Theory and Random Processes), Pittsburgh, PA. *Introduction to Markov Chains*. Nov. 2015.
- **Guest Lecture**, Carnegie Mellon University, STAT 401 (Modern Regression), Pittsburgh, PA. *Introduction to Programming in R and R Markdown*. Aug. 2017.

Research Advisor

05/2015 - 08/2015

Undergraduate student: Benjamin Leroy (UC Berkeley; now CMU PhD student)

Project: Dynamical Mass Measurements of Galaxy Clusters

Funding: National Science Foundation ([Award #1043903](#))

Software

R package `trendfilteringSupp`: *Optimal one-dimensional data analysis with trend filtering*. Available at <https://github.com/capolitsch/trendfilteringSupp>.

R package `aardvark`: *COVID-19 forecasters from Carnegie Mellon's Delphi Lab*. Available at <https://github.com/cmu-delphi/covid-19-forecast>.

Professional Service

Referee *Journal of the Royal Statistical Society: Series B*
Journal of Cosmology and Astroparticle Physics (JCAP)
NASA Experimental Program to Stimulate Competitive Research (EPSCoR)
Astronomy and Computing (A&C)
CHANCE Magazine

Session Organizer *Statistical Challenges in Cosmology*, JSM 2021, Seattle, WA.

Session Chair *Computing, Graphics, and Programming Statistics*, JSM 2017, Baltimore, MD.

Judging Panel *Tartan Data Science Cup 2017*, Carnegie Mellon University.

Outreach Talks *Astrostatistics*, Hillel Academy of Pittsburgh, AP Statistics class, 2017.

Professional Memberships

AAS *American Astronomical Society*
ASA *American Statistical Association*
COIN *Cosmostatistics Initiative*
IAA *International Astrostatistics Association*
IAIA *International Astroinformatics Association*

In the News

- [CMU Statistics and Data Science Graduate Students Keep Winning Big](#)
- [Mr. Indispensable, from Lionel Messi to Virgil Van Dijk: Which player can your team not live without?](#) (Joint work with Francesca Matano, Lee Richardson, et al.)
- [NBA Hackathon 2017 Recap](#)

References

- Larry Wasserman
Ph.D. Advisor
Department of Statistics & Data Science
Machine Learning Department
Carnegie Mellon University

- Jessi J. Cisewski-Kehe
Ph.D. Co-advisor
Department of Statistics
University of Wisconsin-Madison

- Rupert A.C. Croft
Ph.D. Co-advisor
Department of Physics
McWilliams Center for Cosmology
Carnegie Mellon University