Collin A. Politsch

Kavli Institute for Cosmology
University of Cambridge
Madingley Road, Cambridge, CB3 0HA
United Kingdom

October 6, 2022
collin.politsch@ast.cam.ac.uk
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, planetary transits

Academic Positions

University of Cambridge — Cambridge, U.K.
Institute of Astronomy and Kavli Institute for Cosmology — Sep 2022 – present
Postdoctoral Research Associate

Carnegie Mellon University — Pittsburgh, PA
Machine Learning Department — July 2020 – Aug 2021
Postdoctoral Fellow

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 A.C. Croft

Carnegie Mellon University — Pittsburgh, PA
M.Sc. in Machine Learning — 2017

Thesis: *Exploring the Intergalactic Medium*

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

1. **Three-dimensional cosmography of the high redshift Universe using intergalactic absorption**
   C. A. Politsch, J. Cisewski-Kehe, R. A. C. Croft, L. Wasserman
   In preparation. Pre-submission inquiry approved by *Nature*.

2. **Trend Filtering – I. A Modern Statistical Tool for Astronomical Spectroscopy and Time-Domain Astronomy**

C. A. Politsch, J. Cisewski-Kehe, R. A. C. Croft, L. Wasserman
Monthly Notices of the Royal Astronomical Society, Volume 492, Issue 3, March 2020, Pages 4005-4018.
Publisher. arXiv. trendfiltering R package.

* **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**
C. A. Politsch, J. Cisewski-Kehe, R. A. C. Croft, L. Wasserman
Monthly Notices of the Royal Astronomical Society, Volume 492, Issue 3, March 2020, Pages 4019-4032.
Publisher. arXiv. trendfiltering R package.

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

4. **Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the United States**
E. Y. Cramer, E. L. Ray, V. K. Lopez, et al.
Proceedings of the National Academy of Sciences, Volume 119, Issue 15, April 2022.
Publisher. medRxiv. COVID-19 Forecast Hub.

5. **An Open Repository of Real-Time COVID-19 Indicators**
A. Reinhart, L. Brooks, M. Jahja, A. Rumack, J. Tang, et al.
Proceedings of the National Academy of Sciences, Volume 118, Issue 51, December 2021.
Publisher. medRxiv. Supplement. Data Access.

6. **Augmenting Adjusted Plus-Minus in Soccer with FIFA Ratings**
F. Matano, L. F. Richardson, T. Pospisil, C. A. Politsch, J. Qin
To appear in *Journal of Quantitative Analysis in Sports*.
arXiv. Sports Analytics Website.

Other Publications

1. **Mapping the Large-scale Universe through Intergalactic Silhouettes**
C. A. Politsch and R. A. C. Croft
CHANCE, Volume 32, Issue 3, Sep. 2019, Pages 14-19.
Publisher.

Awards and Honors

- 2020-'21 Umesh K. Gavaskar Memorial Award for Best Ph.D. Dissertation in Statistics and Data Science at Carnegie Mellon University.
- Finalist for best paper in the 2020 ASA Astrostatistics Student Paper Competition, sponsored by the Astrostatistics Interest Group.
- 2nd Place: The Data Open 2018 at CMU, presented by Citadel and Correlation One.
  - 300+ applications, ∼125 selected to compete for $25,000 in prizes
- 2nd Place: 2017 NBA Basketball Analytics Hackathon, New York, NY, hosted by the NBA.
  - 900+ applications, ∼200 selected to compete for ∼$20,000 equivalent in tickets, etc.
- 2nd Place: The Data Open 2017 at CMU, presented by Citadel and Correlation One.

  - 550+ applications, ∼125 selected to compete for $25,000 in prizes

- 3rd Place: 2017 CMU BrainHub NeuroHackathon, sponsored by Google.
  - 51 CMU graduate students selected to compete for free tuition and travel stipends

Selected Talks

Invited

- *Three-dimensional cosmography of the high redshift Universe using intergalactic absorption*
  - *Into the Impossible With Brian Keating*, Aug. 2021. YouTube.
  - University of Chicago, Machine Learning in Complex Phenomena seminar series, 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.
  - The "Physics of the Future" NSF AI Planning Institute at Carnegie Mellon & the Statistical Methods for the Physical Sciences (STAMPS) Research Group at Carnegie Mellon. Oct. 2020.
  - The Flatiron Institute, Center for Computational Astrophysics & 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*
  - Joint Statistical Meetings. Best Astrostatistics Student Paper Award Session. Aug. 2020.
  - The "Data-Driven Discovery in Physics" NSF AI Planning Institute at Carnegie Mellon. Oct. 2019.

- *From Mapping the Universe to Forecasting the Pandemic*
  - OnSolve Nexus 2021: Managing Uncertainty for Organizational Resiliency (Paid speaker). April 2021.

- *A Multi-Resolution 3D Map of the Intergalactic Medium via the Lyman-α Forest*
  - Uber Technologies, Inc., San Francisco, CA. Aug. 2018.
  - Statistical and Applied Mathematical Sciences Institute (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*
  - Joint Statistical Meetings. Session: *Statistical Challenges in Cosmology*. Aug. 2021.

- *A Multi-Resolution 3D Map of the Intergalactic Medium via the Lyman-α Forest*
  - Statistical and Applied Mathematical Sciences Institute (SAMSI), Astronomy Transition Workshop, Durham, NC. May 2017.

- *Multi-resolution Regression, Divide and Conquer Risk Estimation, and the Large-scale Universe*
  - Department of Statistics & Data Science, Carnegie Mellon University, Pittsburgh, PA. May 2017.
  - Statistical and Applied Mathematical Sciences Institute (SAMSI), Durham, NC. April 2017.

- *Exploring the Intergalactic Medium*
  - Department of Statistics & Data Science, Carnegie Mellon University, Pittsburgh, PA. April 2017.

Software

- R package trendfiltering: *The state-of-the-art method for denoising 1D signals*
  Available at https://capolitsch.github.io/trendfiltering
- R package SALTdenoiseR: *Statistical software for the SALT Observatory*
  Available at https://capolitsch.github.io/SALTdenoiseR
- 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 research group*
  Available at https://github.com/cmu-delphi/covid-19-forecast

Teaching

Carnegie Mellon University

Instructor

- *Introduction to Programming in R.* Summer 2015. B.Sc. course

Guest Lecturer

- 36-401: *Modern Regression* B.Sc. course
  - Lecture: "Introduction to Programming in R and R Markdown." Aug. 2017.
- 36-217: *Probability Theory and Random Processes* B.Sc. course
  - Lecture: "Introduction to Markov Chains." Nov. 2015.

Head Graduate Teaching Assistant 06/2015 – 12/2018

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

Graduate Teaching Assistant 01/2015 – 05/2015

- 36-402/608: *Advanced Methods for Data Analysis* B.Sc./M.Sc. course

Advising

Undergraduate students

- Benjamin LeRoy (UC Berkeley), Summer Undergraduate Research Experience in Statistics, Dept. of Statistics & Data Science, Carnegie Mellon University, Summer 2015. "Dynamical Mass Measurements of Galaxy Clusters." (Received Ph.D. in Statistics & Data Science from CMU in 2021).

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
Description:

I was a core member of the CMU-based Delphi Research Group and Lead of the forecasting development and evaluation team. Our research was devoted to developing statistical models for forecasting COVID-19 incidence in the United States in order to help inform a data-driven national response to the COVID-19 pandemic.

Graduate Research Assistant, McWilliams Center for Cosmology 01/2019 – 06/2020

Project: Intensity Mapping the Universe
Funding: NASA Grant #NNX17AK56G
PI: Rupert A.C. Croft

Graduate Research Assistant, Department of Statistics & Data Science 01/2015 – 08/2016

Project: Nonparametric Procedures that Exploit Structured Data and Models
Funding: NSF Grant #1521786
PI/co-PI(s): Ann Lee, Chad Schafer, Shirley Ho

Project: Statistics and Machine Learning for Scientific Inference
Funding: NSF Grant #1043903
PI: Larry Wasserman

Uber Technologies, Inc. San Francisco, CA
Data Scientist Intern 06/2018 – 08/2018

Team: UberEverything Data Science
Project: A Holistic Approach to Uber Eats Home Feed Ranking Optimization
Description:

I 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, Chile
La Serena School for Data Science: Applied Tools for Astronomy 08/2015

Project: Cosmology with the Cosmic Microwave Background Through Cross Correlations
Funding: NSF Grant #1637359, MAS, CONICYT
Mentors: Jeffrey McMahon, Chris Miller

North Carolina State University Raleigh, NC
Undergraduate Research Assistant 05/2013 – 07/2013

Project: Portfolio Optimization with Conditional Value-at-Risk (CVaR)
Funding: NSF Grant #1461148, NSA

PI: Tao Pang

University of Kansas Lawrence, KS
Undergraduate Research Assistant 01/2013 – 05/2014

Project: Optimal Control of Stochastic Systems Driven by Fractional Brownian Motions
Funding: U.S. Army Research Contract #W911NF-10-1-0248
PI/co-PI(s): Tyrone E. Duncan, Bozenna Pasik-Duncan

Project: Optimal and Adaptive Control of Stochastic Systems
Funding: Air Force Office of Scientific Research Grant #FA9550-09-1-0554
PI/co-PI(s): Tyrone E. Duncan, Bozenna Pasik-Duncan

Project: Control of Stochastic Systems
Funding: NSF Grant #1108884
PI/co-PI(s): Tyrone E. Duncan, Bozenna Pasik-Duncan

Professional Service

Referee	*Journal of Cosmology and Astroparticle Physics (JCAP)* *NASA Experimental Program to Stimulate Competitive Research* *Astronomy and Computing (A&C)* *CHANCE Magazine*
Program Chair	Serving as Program Chair-Elect for the ASA Astrostatistics Interest Group during the 2022-'23 academic year and Program Chair during the 2023-'24 year, culminating with the organization of the full Astrostatistics program at the Joint Statistical Meetings 2024.
Session Organizer	*Statistical Challenges in Cosmology*, Joint Statistical Meetings 2021, Seattle, WA.
Session Chair	*Computing, Graphics, and Programming Statistics*, Joint Statistical Meetings 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*
- Cristiano Ronaldo effect: Are Man United, Portugal benefitting from today's version of the superstar? (Joint work with Francesca Matano, Lee Richardson, et al.)

- 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*