Collin Politsch

Born in Shawnee, Kansas, USA
Email: collin.politsch@ast.cam.ac.uk
Website: https://collinpolitsch.com
Last updated: March 19, 2023

Kavli Institute for Cosmology Cambridge
c/o Institute of Astronomy
Madingley Road
Cambridge CB3 0HA, U.K.

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, supernova cosmology

Academic Positions and Affiliations

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

Flatiron Institute, Guest Researcher — New York, NY, USA
Center for Computational Astrophysics — July 2021 – May 2022

Carnegie Mellon University, Postdoctoral Fellow — Pittsburgh, PA, USA
Machine Learning Department and the Delphi Group — July 2020 – Aug 2021

Education

Carnegie Mellon University — Pittsburgh, PA, USA
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, USA
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, USA
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
   Pre-submission inquiry approved by *Nature*. Preparing to submit in full.

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.
   [Publisher] [arXiv] [Software]

   * **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.
   [Publisher] [arXiv] [Software]

   * **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] [Data Access]

5. **The United States COVID-19 Forecast Hub dataset**
   E. Y. Cramer, Y. Huang, Y. Wang, et al.
   Scientific Data, Volume 9, Issue 462, August 2022.
   [Publisher] [Data Access]

6. **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] [Data Access]

7. **The Young Supernova Experiment Data Release 1 (YSE DR1): Light Curves and Photometric Classification of 1978 Supernovae**
   P. D. Aleo, K. Malanchev, S. Sharief, D. O. Jones, et al.
   To appear in *The Astrophysical Journal Supplement Series*. [arXiv] [Data Access]

8. **Mapping the Large-scale Universe through Intergalactic Silhouettes** [*Lightly-refereed]
   C. A. Politsch and R. A. C. Croft
   CHANCE, Volume 32, Issue 3, September 2019. [Publisher]

9. **Augmenting Adjusted Plus-Minus in Soccer with FIFA Ratings**
   F. Matano, L. F. Richardson, T. Pospisil, C. A. Politsch, J. Qin
   Journal of Quantitative Analysis in Sports, Volume 19, Issue 1, March 2023.
   [Publisher] [arXiv] [Data Access]

Awards and Honors

- 2020-'21 Umesh K. Gavaskar Memorial Award for Best Ph.D. Dissertation in Statistics and Data Science at Carnegie Mellon University.
- 2021 Statistical Partnerships Among Academe, Industry, and Government (SPAIG) Award [Link]
  - Awarded to the Delphi group (in conjunction with our government and corporate partners: the U.S. Centers for Disease Control and Prevention, Google, Facebook, Amazon, Change Healthcare, Optum, and Quidel Inc.) by the American Statistical Association "for commitment to the theory and practice of epidemic tracking and forecasting through building and modeling unique public health data streams," for our partnership on COVIDcast.
- 2021 Allen Newell Award for Research Excellence [Link]
  - Awarded to the Delphi Group by the Carnegie Mellon University School of Computer Science, "for advancing the theory and practice of epidemic tracking and forecasting, and enabling national collaborative scientific response."
- Finalist for best paper in the 2020 ASA Astrostatistics Student Paper Competition, sponsored by the Astrostatistics Interest Group. [Link]
- 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 game 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 Carnegie Mellon University 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, University of Cambridge, Department of Applied Mathematics and Theoretical Physics. 10/2022

(Invited) Machine Learning for Astronomy, Into the Impossible With Brian Keating. 08/2021

Three-dimensional cosmography of the high redshift Universe using intergalactic absorption, Joint Statistical Meetings. *Statistical Challenges in Cosmology* session. 08/2021

(Invited) From Mapping the Universe to Forecasting the Pandemic, OnSolve Nexus 2021: Managing Uncertainty for Organizational Resiliency. 04/2021

(Invited) Three-dimensional cosmography of the high redshift Universe using intergalactic absorption, University of Chicago, Machine Learning in Complex Phenomena seminar series. 02/2021

(Invited) Three-dimensional cosmography of the high redshift Universe using intergalactic absorption, University of Maryland, Department of Mathematics. 11/2020

(Invited) Three-dimensional cosmography of the high redshift Universe using intergalactic absorption, Duke University, Department of Statistical Science. 11/2020

(Invited) Three-dimensional cosmography of the high redshift Universe using intergalactic absorption, "Physics of the Future" NSF AI Planning Institute at Carnegie Mellon University. 10/2020

(Invited) Three-dimensional cosmography of the high redshift Universe using intergalactic absorption, The Flatiron Institute, Center for Computational Astrophysics & NYU. 10/2020

(Invited) Three-dimensional cosmography of the high redshift Universe using intergalactic absorption, Los Alamos National Laboratory. 10/2020

(Award) Trend Filtering: A Modern Statistical Tool for Time-Domain Astronomy and Astronomical Spectroscopy, Joint Statistical Meetings, Best Astrostatistics Student Paper Award Session. 08/2020

(Invited) Trend Filtering: A Modern Statistical Tool for Time-Domain Astronomy and Astronomical Spectroscopy, "Data-Driven Discovery in Physics" NSF AI Planning Institute at Carnegie Mellon University. 10/2019

(Invited) A Multi-Resolution 3D Map of the Intergalactic Medium via the Lyman-α Forest, Uber Technologies Data Science. 08/2018

(Invited) A Multi-Resolution 3D Map of the Intergalactic Medium via the Lyman-α Forest, Joint Statistical Meetings. 07/2017

A Multi-Resolution 3D Map of the Intergalactic Medium via the Lyman-α Forest, Statistical and Applied Mathematical Sciences Institute. 05/2017

Multi-resolution Regression, Divide and Conquer Risk Estimation, and the Large-scale Universe, Carnegie Mellon University, Department of Statistics and Data Science & Machine Learning Department. 05/2017

Multi-resolution Regression, Divide and Conquer Risk Estimation, and the Large-scale Universe, Statistical and Applied Mathematical Sciences Institute. 04/2017

(Invited) A Multi-Resolution 3D Map of the Intergalactic Medium via the Lyman-α Forest, Carnegie Mellon University, McWilliams Center for Cosmology. 03/2017

(Invited) A Multi-Resolution 3D Map of the Intergalactic Medium via the Lyman-α Forest, Statistical and Applied Mathematical Sciences Institute, Cosmology Working Group. 11/2016

(Poster) Exploring the Intergalactic Medium, Statistical and Applied Mathematical Sciences Institute, Astronomy Opening Workshop. 08/2016

Software

R packages

- trendfiltering: *The state-of-the-art method for denoising 1D signals* [Link]
- SALTdenoiseR: *Statistical software for the SALT Observatory* [Link]
- aardvark: *COVID-19 forecasters from Carnegie Mellon University's Delphi Group* [Link]

Teaching

University of Cambridge
Instructor, Department of Applied Mathematics and Theoretical Physics 02/2023 – present

- *Mathematical Foundations of Data-Intensive Science*

Teaching Committee, Institute of Astronomy 10/2022 – present

- Postdoc representative on the IoA Teaching Committee

Carnegie Mellon University
Instructor, Department of Statistics and Data Science

- *Introduction to Programming in R* (B.Sc. course) 05/2015 – 07/2015

Guest Lecturer, Department of Statistics and Data Science

- 36-401: *Modern Regression* (B.Sc. course) 08/2017
- 36-217: *Probability Theory and Random Processes* (B.Sc. course) 11/2015

Head Graduate Teaching Assistant 06/2015 – 12/2018
Department of Statistics and Data Science & Machine Learning Department

- 10-702: *Statistical Machine Learning* (Ph.D. course)
- 10-705: *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, Department of Statistics and Data Science 01/2015 – 05/2015

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

Scientific Collaborations

Delphi Delphi Research Group 08/2020 – 08/2021

I was a core member of the CMU-based Delphi Group and Team Lead of our forecasting development and evaluation initiative. The team I personally led devoted our work to developing statistical models for forecasting COVID-19 incidence in the United States in order to support and advise the COVID-19 pandemic responses of the U.S. Centers for Disease Control and Prevention (CDC) and the White House. The Delphi Group received numerous awards for our collective work. Particularly relevant to my team's work was the *2021 Statistical Partnerships Among Academe, Industry, and Government Award*[1] by the American Statistical Association and the *2021 Allen Newell Award for Research Excellence*[2] by the Carnegie Mellon University School of Computer Science.

LSST *Large Synoptic Survey Telescope (Vera C. Rubin Observatory)* 01/2023 – present

WEAVE *William Herschel Telescope Enhanced Area Velocity Explorer* 01/2023 – present

YSE *Young Supernova Experiment* 10/2022 – present

Industry Experience

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

Team: UberEverything Data Science

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

Role: 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.

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).

Equality, Diversity, and Inclusion

- Postdoc representative on the department EDI Committee 10/2022 – present
  Institute of Astronomy, University of Cambridge.
- Co-Chair of the department EDI Work-Life Balance focus group 10/2022 – present
  Institute of Astronomy, University of Cambridge.
- Postdoc representative on the Institute of Physics Project Juno Self-Assessment Team 10/2022 – present
  Institute of Astronomy, University of Cambridge.

Academic Service

- Postdoc representative on the department Teaching Committee 10/2022 – present
  Institute of Astronomy, University of Cambridge.
- Member of the department Postdoc Committee 01/2023 – present
  Institute of Astronomy, University of Cambridge.

Referee *Astronomy and Computing (A&C)*
Journal of Cosmology and Astroparticle Physics (JCAP)
NASA Experimental Program to Stimulate Competitive Research (EPSCoR)
CHANCE Magazine

Judging Panel ASA Astrostatistics Student Paper Competition 2023
Tartan Data Science Cup 2017, Carnegie Mellon University.

Program Chair Program Chair-Elect for the ASA Astrostatistics Interest Group during the 2022-'23 academic year and Program Chair during the 2023-'24 year. Responsible for organizing 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.

Outreach

Using Statistics to Explore the Universe, Hillel Academy of Pittsburgh, AP Statistics. 03/2017

Research Funding

University of Cambridge Cambridge, U.K.
Postdoctoral Research Associate 09/2022 – present
Kavli Institute for Cosmology and the Institute of Astronomy

Project: Next-Generation Data-Driven Probabilistic Modelling of Type Ia Supernova SEDs in the Optical to Near-Infrared for Robust Cosmological Inference – BayeSN
Funding: ERC Grant #101002652
PI: Kaisey Mandel

Carnegie Mellon University Pittsburgh, PA, USA
Postdoctoral Fellow, Machine Learning Department 07/2020 – 08/2021

Project: COVIDcast
Funding: Unrestricted Gift from Google.org [Link]
PI: Ryan J. Tibshirani

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

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

North Carolina State University Raleigh, NC, USA
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, USA
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 Memberships

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

References

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

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

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