Collin A. Politsch

Born in Shawnee, Kansas, USA
Email: collin.politsch@ast.cam.ac.uk
Website: https://collinpolitsch.com
Last updated: December 22, 2022

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
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] [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, Pages 4019-4032.
   [Publisher] [arXiv] [Software]
   * **Finalist for best paper in the 2020 ASA Astrostatistics Student Paper Competition**, sponsored by the Astrostatistics Interest Group.

4. **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 *ApJS*. [arXiv]

5. **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]

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. **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] [Data Access]

Lightly-Reviewed 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.

- 2021 Statistical Partnerships Among Academe, Industry, and Government (SPAIG) Award [Link]
  - Awarded to the Delphi group, the United States CDC, Google, Facebook, Amazon, Change Healthcare, Optum, and Quidel Inc. "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.
- 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 CMU BrainHub NeuroHackathon, sponsored by Google.
  - 51 CMU graduate students selected to compete for free tuition and travel stipends

Selected Talks

Invited

- *Machine Learning for Astronomy*
  - *Into the Impossible With Brian Keating*, Aug. 2021. [YouTube]
- *Three-dimensional cosmography of the high redshift Universe using intergalactic absorption*
  - University of Cambridge, Department of Applied Mathematics and Theoretical Physics, Cambridge, U.K. Oct. 2022.
  - 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. [YouTube]
  - 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. [YouTube]
  - 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.
  - Joint Statistical Meetings, Baltimore, MD. July 2017. (Poster)
  - Statistical and Applied Mathematical Sciences Institute, 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, Astronomy Transition Workshop, Durham, NC. May 2017.
  - American Statistical Association Banquet, Pittsburgh, PA. April 2017. (Poster)

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

- *Exploring the Intergalactic Medium*
  - Carnegie Mellon University, Department of Statistics & Data Science, Pittsburgh, PA. April 2017.
  - Statistical and Applied Mathematical Sciences Institute, Astronomy Opening Workshop, Durham, NC. Aug. 2016. (Poster)

- *Statistical Methods for Estimating Regression Quantiles*
  - Carnegie Mellon University, Machine Learning Department, Pittsburgh, PA. Dec. 2015. (Poster)

Software

R packages

- trendfiltering: *The state-of-the-art method for denoising 1D signals* [Link]
- SALTdenoiseR: *Statistical software for the SALT Observatory* [Link]
- trendfilteringSupp: *Optimal one-dimensional data analysis with trend filtering* [Link]
- aardvark: *COVID-19 forecasters from Carnegie Mellon's Delphi research group* [Link]

Teaching

University of Cambridge
Teaching Committee 10/2022 - present

- Postdoc Student Representative on the Institute of Astronomy's Teaching Committee

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

Equality, Diversity, and Inclusion

- Postdoc representative on the Equality, Diversity & Inclusion (EDI) Committee
  Institute of Astronomy, University of Cambridge. Oct. 2022 – present.
- Postdoc representative on the Institute of Physics (IOP) Project Juno Self-Assessment Team (SAT)
  Institute of Astronomy, University of Cambridge. Oct. 2022 – present.
- Chair of the Work-Life Balance focus group
  Institute of Astronomy, University of Cambridge. Nov. 2022 – present.

Academic Service

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