

# MATEO DULCE RUBIO

mdulceru@andrew.cmu.edu

Personal website  $\diamond$  Google Scholar

## EDUCATION

---

<b>Carnegie Mellon University</b> Ph.D. in Statistics and Public Policy Advisor: <a href="#">Edward H. Kennedy</a> Thesis: “Robust Nonparametric Methods for Peacebuilding” Committee: <a href="#">Edward H. Kennedy</a> , <a href="#">Larry Wasserman</a> , <a href="#">Silvia Borzutzky</a> , <a href="#">Eli Ben-Michael</a>	Pittsburgh, PA 2020 - 2025
<b>Carnegie Mellon University</b> M.Sc. in Statistics	Pittsburgh, PA 2017 - 2022
<b>University of the Andes</b> M.Sc. in Economics, <i>Cum Laude</i>	Bogotá, Colombia 2017 - 2019
<b>University of the Andes</b> B.S. in Economics B.S. in Mathematics	Bogotá, Colombia 2012 - 2017 2012 - 2017

## AWARDS & HONORS

---

INFORMS Doing Good with Good OR Competition Finalist	2024
American Statistical Association JSM Student Paper Award (Social Statistics Section)	2024
K&L Gates Presidential Fellowship in Ethics and Computational Technologies	2024
Outstanding Teaching Assistant Award, Carnegie Mellon University	2023
<i>Cum Laude</i> M.Sc. in Economics, University of the Andes	2019

## RESEARCH INTERESTS

---

Theory	Nonparametric statistics, causal inference, population size estimation, responsible machine learning, geospatial data.
Application	Humanitarian mine action, criminal justice, health policy, sustainability, public policy.

## PUBLICATIONS & MANUSCRIPTS

---

### Working Papers

Identification of Hazard Clusters for Priority Landmine Clearance as a Quadratic Knapsack Problem.

**Mateo Dulce Rubio.**

**INFORMS Doing Good with Good OR Finalist.**

Advancements in Mine Action: Enhancing Remote Reporting and Analysis through Innovative Technologies.

Rory Collins, Lionel Fragniere, **Mateo Dulce Rubio.**

*Submitted.*

### Journal Articles

Population Size Estimation with Many Lists and Heterogeneity: A Conditional Log-Linear Model among the Unobserved.

**Mateo Dulce Rubio**, Edward Kennedy.

*arXiv:2407.03539. Submitted.*

**JSM Student Paper Award (Social Statistics Seccion).**

Effects of Adolescent Victimization on Offending: Flexible Methods for Missing Data & Unmeasured Confounding.

**Mateo Dulce Rubio**, Edward Kennedy, Valerio Baćak, Daniel Nagin.

*arXiv:2309.12595*. Under revision.

RELand: Risk Estimation of Landmines via Interpretable Invariant Risk Minimization.

**Mateo Dulce Rubio\***, Siqi Zeng\*, Qi Wang, Didier Alvarado, Francisco Moreno, Hoda Heidari, Fei Fang.

*ACM Journal on Computing and Sustainable Societies* 2 (2), 2024.

A Location Discrete Choice Model of Crime: Police Elasticity and Optimal Deployment.

Douglas Newball-Ramírez, Álvaro Riascos, Andrés Hoyos, **Mateo Dulce**.

*PLOS One*, 19(3), e0294020, 2024.

Modelling Underreported Spatio-Temporal Crime Events.

Álvaro J. Riascos Villegas, Jose Sebastian Nungo, Lucas Gómez Tobón, **Mateo Dulce Rubio**, Francisco Gómez.

*PLOS One*, 18(7), e0287776, 2023.

Spatio Temporal Sparsity in Homicide Prediction Models.

Álvaro J. Riascos Villegas, Juan S. Moreno Pabón, **Mateo Dulce**, Sebastián Quintero, Johan García, Hernán García.

*IEEE Access*, 10, 14359-14367, 2022.

Efficient Nearest Neighbors Methods for Support Vector Machines in High Dimensional Feature Spaces.

Diana C. Montañes, Adolfo J. Quiroz, **Mateo Dulce Rubio**, Álvaro J. Riascos Villegas.

*Optimization Letters* 15 (2), 391-404, 2020.

On the Relationship between the Inhomogeneous Wave and Helmholtz Equations in a Fractional Setting.

**Mateo Dulce**, Alexander Getmanenko.

*Abstract and Applied Analysis*. Vol. 2019. *Hindawi*, 2019.

## Conference Articles

Statistical Inference Under Constrained Selection Bias.

Santiago Cortés-Gómez, **Mateo Dulce**, Carlos Patiño, Bryan Wilder.

*41st International Conference on Machine Learning (ICML)*, 2024. (to appear).

Self-exciting point processes with image features as covariates for robbery modeling.

**Mateo Dulce**, Paula Rodríguez, Juan Moreno, Álvaro Riascos, Jorge Camargo.

*Intelligent Computing: Proceedings of the 2021 Computing Conference, Volume 1*. Springer International Publishing, 2022.

Visual Representations to Evaluate the Heterogeneous Effects of Urban Parks Restoration on Crime.

Julián Chitiva Bocanegra, Douglas Newball Ramírez, Paula Rodríguez Díaz, Hamadys Benavides Gutiérrez, **Mateo Dulce**, Álvaro Riascos.

*Proceedings of the 4th ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS)*, 2021..

Interpreting a Conditional Generative Adversarial Network Model for Crime Prediction.

**Mateo Dulce**, Óscar Gómez, Juan Moreno, Christian Urcuqui, Álvaro Riascos.

*25th Iberoamerican Congress on Pattern Recognition (CIARP)*. Springer, Cham, 2021.

A Fair Allocation Algorithm for Predictive Police Patrolling.

Isabella Rodas, **Mateo Dulce**, Álvaro Riascos.

*8th International Conference on Artificial Intelligence & Applications (ARIA)*, 2021.

Dynamic Network Analysis of Spatio-Temporal Crime Incidents in Bogotá - Colombia.

Juan Moreno, Hernán García, Sebastián Quintero, Johan García, Alvaro Riascos, **Mateo Dulce**.

*8th International Conference on Behavioral and Social Computing (BESC)*, 2021.

Graph Restrictions for Signal Processing of Homicides Data.

Sebastián Quintero, Juan Moreno, **Mateo Dulce**, Álvaro Riascos and Gustavo Nonato.

*Proceedings of the International Conference on Applied Artificial Intelligence (ICAPAI), pp. 1-6, 2021.*

#### Zero-Inflated Embeddings to Analyze Homicide Occurrence Patterns.

Hamadys Benavides, Óscar Gómez, **Mateo Dulce**, Paula Rodríguez, Juan Moreno, Álvaro Riascos.  
*2nd International Conference on Computing and Data Science (CDS), 2021.*

#### A Manifold Learning Data Enrichment Methodology for Homicide Prediction.

Juan Moreno, **Mateo Dulce**, Álvaro Riascos, Yor Castaño, Paula Rodríguez.  
*7th International Conference on Behavioural and Social Computing (BESC), 2020.*

#### Accuracy and Fairness in a Conditional Generative Adversarial Model of Crime Prediction.

Christian Urcuqui, Juan Moreno, Carlos Montenegro, Álvaro Riascos, **Mateo Dulce**.  
*7th International Conference on Behavioural and Social Computing (BESC), 2020.*

**Best Paper Award Runner-up.**

#### Predicting Criminal Behavior with Lévy Flights using Real Data from Bogotá.

**Mateo Dulce.**

*Documento CEDE No. 2019-II.* Master's thesis. Contributed Talk at LXAI Workshop at Neurips, 2018.

#### Efficient Allocation of Law Enforcement Resources using Predictive Police Patrolling.

**Mateo Dulce**, Simón Ramírez-Amaya and Álvaro Riascos.

*NeurIPS 2018 Workshop on Machine Learning for the Developing World (ML4D): Achieving Sustainable Impact, 2018.*

## Book Chapters

#### Randomized Experiments.

Amanda Coston, **Mateo Dulce**, Edward Kennedy.

Book chapter in *AI for Social Impact*. Edited by Milind Tambe, Fei Fang, Bryan Wilder, 2022.

## WORK EXPERIENCE

---

### Apple

*Machine Learning Intern*

May 2024 - August 2024

*Seattle, WA*

- Use observational and guidelines informed human annotation data to evaluate foundation Large Language Models.

### RAND Corporation

*Summer Associate*

May 2022 - August 2022

*Pittsburgh, PA*

- Designed, developed, and implemented a pipeline for monitoring emerging trends in the interaction between society, technology, and the criminal justice system, using topic modeling (NLP) on text data and Markov switching models.

### Quantil | Applied Mathematics

*Data Mining Associate Director*

May 2017 - December 2021

*Bogotá, Colombia*

- Led the development and implementation of spatio-temporal crime prediction models in Bogotá, using Deep Neural Networks and Poisson point processes.
- Natural Language Processing for information retrieval, sentiment analysis, and classification of diverse text data sources.
- Applied ML and Markov chains methods in health data to develop prediction models for risk assessment.
- Conducted fairness auditing of matching algorithms for ridesharing and crime prediction models, ensuring equitable outcomes and promoting transparency and ethical considerations in algorithmic decision-making.

## TEACHING EXPERIENCE

---

### Teaching Assistant, Carnegie Mellon University

Critical Analysis of Policy Research	Fall 2021, 2022, 2023
Advanced Methods for Data Analysis ( <b>Outstanding TA Award</b> )	Spring 2022, 2023
Reasoning with Data	Summer 2023
Probability and Statistics	Fall 2022
Summer Undergraduate Research Apprenticeship	Summer 2022
Statistics for IT Managers	Fall 2021
Introduction to Statistical Inference	Summer 2021
Basic Mathematics for Management	Summer 2021
Data over Space and Time	Fall 2020

### Lecturer, University of the Andes

Analytics Project Management	Fall 2019
Machine Learning for Business intelligence	Fall 2019
Introduction to Data Analysis in Python	Fall 2019

## SERVICE

---

- Topic-Contributed Session Organizer on “Recent Advances in Capture-Recapture Methods for Population Size Estimation” at JSM 2024.
- Referee for ACM Journal on Computing and Sustainable Societies.
- Program Committee of ACM EAAMO Conference 2024.
- Co-founder of the [Data Analytics Center for Public Policy](#), a research initiative dedicated to fostering the advancement of mathematics, data science, and AI for the public sector in Colombia and Latin America (2021 - ).
- Program Committee of Machine Learning for Developing World (ML4D) NeurIPS Workshop (2020, 2021).

## SELECTED PRESENTATIONS

---

\* indicates presentation scheduled for future date

Statistical Inference Under Constrained Selection Bias <i>Contributed talk, ACM conference on Equity and Access in Algorithms, Mechanisms, and Optimization.</i>	2024*
Prediction of Landmine Contaminated Areas for Early Impact Analysis <i>Geneva International Centre for Humanitarian Demining, AI Applications for Mine Action.</i>	2024*
Identification of Hazard Clusters for Priority Landmine Clearance as a Quadratic Knapsack Problem <i>INFORMS Doing Good with Good OR Competition Final.</i>	2024*
Population size estimation with many lists and heterogeneity: A conditional log-linear model among the unobserved <i>Joint Statistical Meeting GSS/SRMS/SSS Student Paper Competition.</i>	2024
Population size estimation with many lists and heterogeneity: A conditional log-linear model among the unobserved <i>RAND Corporation Statistics Group.</i>	2024
RELand: Risk Estimation of Landmines via Interpretable Invariant Risk Minimization <i>Contributed talk, COMPASS Conference.</i>	2024
RELand: Risk Estimation of Landmines via Interpretable Invariant Risk Minimization <i>Contributed talk, EAAMO Doctoral Consortium.</i>	2023

Nonparametric Methods for Identification and Estimation of Total Population Size <i>Causal Inference Seminar, Carnegie Mellon University.</i>	2023
Effects of Adolescent Victimization on Offending: Flexible Methods for Missing Data & Unmeasured Confounding <i>Contributed talk, ICHPS Conference.</i>	2023
Statistical Analysis of Randomized Experiments <i>Guest Lecturer, AI for Social Impact, Harvard University.</i>	2022
Fair Algorithms for Learning in Allocation Problems <i>Fairness, Equity, Accountability and Ethics Seminar, Carnegie Mellon University.</i>	2022
Algorithmic Landmine Risk Prediction <i>Invited talk, K-12 Summit on Artificial Intelligence for Social Good.</i>	2022