

Monica G. Bobra

Principal Data Scientist

San Francisco Bay Area | mbobra@alum.mit.edu | mbobra.github.io

Summary

I have over a decade of experience leading data science research. I develop novel machine learning algorithms and apply them to complex data to glean insights and inform public decision-making. I develop open data and open source scientific software for data-intensive research. I aim to foster interdisciplinary collaboration by bringing together disparate communities in scientific research, software development, public policy, storytelling, and ethics.

Education

University of New Hampshire, Durham NH

M.S. Physics
JANUARY 2010

Boston University, Boston MA

B.A. Astronomy
B.S. Communication
MAY 2004

Skills

Python scientific software stack

Machine learning algorithms for image data, metadata, and time series data as well as interpretability tools

Statistical modeling

Data visualization

Cloud computing (AWS, GCP, Snowflake)

Git (and CI/CD)

Awards

NASA Group Achievement Award — SunPy Development Team (2024)

American Astronomical Society Solar Physics Division Popular Media Award (2021)

NASA Group Achievement Award — Solar Dynamics Observatory Team (2017)

Robert H. Goddard Exceptional Achievement for Science Award (2016)

NASA Space Grant Fellowship (2008 - 2009)

NASA Group Achievement Award — Hinode Team (2007)

Experience

State of California, Office of Data and Innovation /

Principal Data Scientist

MAY 2023 - PRESENT, SACRAMENTO & SAN FRANCISCO BAY AREA, CA

Leading new initiatives with partner departments, such as the California Environmental Protection Agency, to develop and implement machine learning models designed to improve community safety, sustainability, and decision-making

Leading state-wide initiative on ethical and reproducible practices in data science by developing guidelines and advising departments

Supporting, together with other departments, the Governor's Executive Order on GenAI

Stanford University / Research Scientist

APRIL 2010 - JULY 2021, STANFORD CA

Published [several studies](#) and [presented talks](#) on solar flare prediction models using machine learning algorithms and image data taken by NASA satellites, that pioneered a new field of heliophysics research and garnered media attention

Authored a book titled [Machine Learning, Statistics, and Data Mining for Heliophysics](#)

Led the development of open source software as a founding member of the Python in Heliophysics community, Editor for the Journal of Open Source Software, and Vice-Chair of the Board for SunPy

Organized conferences to foster interdisciplinary collaboration, such as Machine Learning in Heliophysics (2019), Python in Astronomy (2020), and COSPAR Data Science Workshops (2021)

Led a culture of open scholarship at Stanford Data Science as a founding member of the Center for Open and REproducible Science

Wrote science policy to inform the direction of data science at a federal level as a member of the National Academy of Sciences Heliophysics Mid-Decadal Committee (2020)

Harvard-Smithsonian Center for Astrophysics /

Astrophysicist

OCTOBER 2005 - AUGUST 2007, CAMBRIDGE MA

Developed flight hardware tests, operations, and software for the JAXA/NASA Hinode satellite, and developed a numerical model of solar flares that accurately reproduces observations from the satellite