Magdalena Siwek, PhD

Research Scientist

New York, NY · mssiwek92@gmail.com · mssiwek.github.io · github.com/mssiwek · linkedin.com/in/mssiwek

Summary

I build statistical and machine learning surrogate models for data from expensive simulations. I am currently writing **fates**, a context-aware, transformer-based model in JAX with flexible, uncertainty-aware loss functions that forecasts time series with uncertainty. Before that I built **calypso**, a stochastic emulator that predicts time series with estimates of both aleatoric and epistemic uncertainty, released on PyPI with a live web app. I completed my PhD in astrophysics at Harvard in 2024, and I am now a Junior Fellow of the Simons Society of Fellows at NYU and the Flatiron Institute.

Skills

Machine learning: transformers, autoregressive sequence models, time-series forecasting, probabilistic and generative modeling, surrogate modeling, uncertainty quantification, PCA.

Programming: Python (JAX, PyTorch, NumPy, SciPy, Pandas), SQL, C/C++, Cython, Bash, Git, LaTeX.

ML engineering: end-to-end ML pipelines, GPU training on HPC clusters, MPI, CI/CD and release automation (GitHub, PyPI, Zenodo), Streamlit apps, reproducible environments (pixi), unit testing (pytest).

AI-assisted development: agentic coding with Claude Code, coordinating several agents on parallel projects, writing design specs that agents work from, reviewing and correcting AI-generated code and text.

Data visualization: Matplotlib, Seaborn, Plotly; interpolating and visualizing unstructured (Voronoi) mesh data; publication figures, interactive dashboards and animations.

Communication: research talks to specialist and general audiences, teaching (Harvard Bok Center Distinction, twice), mentoring.

Projects

fates: Forecasting Autoregressive Transformer for Ensemble Sampling *2026 – present, in progress*

Project page: github.com/mssiwek/fates-showcase

- Writing a transformer in JAX that forecasts accretion-rate time series from hydrodynamic simulations. It has 0.79M parameters, 4 attention blocks and causal attention, with positions encoded by rotary position embeddings (RoPE).
- Developed a Gaussian output head, allowing the model to predict a mean and variance at every step for stochastic sampling.
- Chose a noise-aware sampling cadence from the data, which cut the time per training step from 31 s to 1.7 s.
- Analyzed the attention maps of the forecast model to verify the salience of previous points in the time series.

calypso: stochastic time-series emulator *2025 – present*

GitHub | live demo | paper | `pip install calypso-emulator`

- Designed a generative model for high-dimensional time series (PCA, then multivariate Gaussian fits interpolated across a 2-D parameter space) that returns calibrated synthetic realizations anywhere in parameter space.
- Derived the likelihood analytically, so the model can be used directly for Bayesian inference.
- Shipped v1.0 to PyPI and Zenodo (citable DOIs for code, trained model and data) with a live Streamlit app; releases are automated from a git tag through GitHub, Zenodo, PyPI and the app redeploy.
- Wrote and ran the simulations that produced the training data: about 100 MPI-parallel C/C++ hydrodynamic simulations on HPC clusters, 3 million CPU-hours in total, published in two first-author papers in 2023.

Experience

Junior Fellow, Simons Society of Fellows *Sept 2024 – present*

New York University and Flatiron Institute (Center for Computational Astrophysics), New York, NY

- Built and released calypso, and am now building fates (see Projects).
- Wrote an LSTM emulator for the accretion time series in PyTorch, and later replaced it with the PCA-based model that became calypso.
- Set up the GPU training and large-scale data processing workflows on HPC clusters.

Postdoctoral Researcher *May 2024 – Sept 2024*

Harvard University, Cambridge, MA

- Built statistical population models of massive black hole binaries from my simulation results, and measured how the small-scale physics changes population-level predictions.

- For this work, wrote a Cython-accelerated dynamics class for holodeck, the NANOGrav collaboration's population code.

Doctoral Researcher, Department of Astronomy *Sept 2018 – May 2024*
Harvard University, Cambridge, MA

- Ran a 3 million CPU-hour simulation campaign on HPC clusters across a large parameter space; this became the training data for calypso and fates. Results published in five first-author papers.
- Added new physics modules to arepo, a production C/C++ simulation code, with new MPI-aware data structures and communication patterns.
- Co-developed holodeck, the NANOGrav collaboration's Python code for massive black hole binary populations (several dozen developers), and extended its KDE resampling to correlated parameters, work I continued as a postdoc.
- Wrote Python tools for loading, post-processing and analyzing large distributed simulation outputs.

Teaching Fellow *2021 – 2022*
Harvard University, Cambridge, MA

- Taught cosmology to Harvard undergraduates, and was awarded the Bok Center Certificate of Distinction in Teaching in both years.

Education

PhD, Computational and Theoretical Astrophysics, Harvard University *2024*
Dissertation: developed a scale-free model of how binaries evolve in circumbinary disks, built from large MPI-parallel simulation suites and now widely used in astronomy. Keto Prize for best thesis in theoretical astrophysics.

MSci and BSc, Physics and Astronomy (First Class Honours), University of Glasgow *2017, 2016*

Publications

Six first-author papers in *Monthly Notices of the Royal Astronomical Society*, plus NANOGrav collaboration papers (full list on Google Scholar).

- Siwek et al., *calypso: a Parameter-Conditioned Stochastic Surrogate Model for Circumbinary Accretion Time-Series*, submitted to MNRAS, 2026.
- Siwek et al., *Signatures of Circumbinary Disk Dynamics in Multi-Messenger Population Studies of Massive Black Hole Binaries*, MNRAS, 2024.
- Siwek et al., *Orbital Evolution of Binaries in Circumbinary Disks*, MNRAS, 2023.
- Siwek et al., *Preferential Accretion and Circumbinary Disk Precession in Eccentric Binary Systems*, MNRAS, 2023.
- Siwek et al., *Differential Accretion and the Gravitational Wave Background*, MNRAS, 2020.
- Siwek et al., *Optical and X-ray Luminosities of Expanding Nebulae*, MNRAS, 2017.

Awards

Simons Society of Fellows Junior Fellowship (2024) | Institute for Advanced Study membership (2024, declined) | Harvard Keto Prize, best thesis in theoretical astrophysics (2024) | Bok Center Certificate of Distinction in Teaching (2021, 2022) | Harvard Scholarship Foundation Germany Merit Award (2018) | University of Glasgow prizes (2015–2017).