Ioannis Zachos
a.k.a. **Yannis**

EU Settled Status
Cambridge, UK
Email

Website
LinkedIn
GitHub

I am passionate about devising Statistical and Probabilistic Machine Learning algorithms for industry-relevant Research & Development. I have over six years of Research and Commercial experience in Finance, Commerce, Transport, Environment, and Biology. My PhD research focuses on eliciting individual agent dynamics from aggregate data using Markov Bases, Markov Chain Monte Carlo, and Physics-informed ML.

EDUCATION

University of Cambridge — Cambridge, UK — 11/2020 – Present
PhD in Computational Statistics and Machine Learning.
Supervisors: Prof. Mark Girolami, Prof. Theodoros Damoulas.
- Thesis: *Probabilistic* Inference in Agent-Based Models: Advancements in Population Synthesis and Simulation.

University of Cambridge — Cambridge, UK — 10/2019 – 08/2020
MRes in Future Infrastructure & Built Environment (**Distinction**).
- Courses: Computational Statistics and Machine Learning (83%), Research Methods (79%).

University of Warwick — Coventry, UK — 10/2015 – 07/2018
BSc Data Science (**1st class honours**).
- Courses: Machine Learning (73%), Mathematical Statistics (79%), Linear Statistical Modelling (77%), Topics in Data Science (82%), Artificial Intelligence (72%).

Anatolia College — Salonika, GR — 09/2013 – 07/2015
International Baccalaureate (**39**/**45** - **top 7% globally**).
- Courses: Physics (7/7), Mathematics (6/7), Business Management (6/7), Extended essay on stock price forecasting using Statistics (35/36).

SKILLS & TOOLS

Coding Languages:	Python ($\approx 10^5$ lines), R, SQL, Java ($\approx 10^4$ lines), C, Matlab ($\approx 10^3$ lines each).
Libraries:	`numpy`, `pandas`, `PyTorch`, `TensorFlow`, `sklearn`, `xarray`, `PyMC3`.
Cloud, Databases & Big Data:	Amazon Web Services, Google Cloud Platform, PostgreSQL, MySQL, Geographic Information Systems, Git, Docker.
Machine Learning & Artificial Intelligence:	*Bayesian* Statistics, *Probabilistic* ML, Geospatial modelling, Deep Learning, Markov Chain Monte Carlo, Physics-informed ML, Multi-agent systems.

COMMERCIAL EXPERIENCE

Cervest Ltd (acquired by Mitiga Solutions): Statistical Scientist. — London, UK — 09/2018 – 07/2019
- Led the following research projects:
  - Change-point detection on climate-volatile geospatial data generating processes.
  - Sequential multinomial classification for assessing environmental resilience.
  - *Bayesian* models for *spatio-temporal* image and sensor data fusion.
- Designed and developed data acquisition infrastructures using Python.
- Engaged with clients and investors to facilitate science communication.

Eurobank Private Bank Luxembourg: Investment Advisory Intern. — Athens, GR — 06/2018 – 08/2018
- Derived optimal portfolios using efficient frontier theory with diversification and volatility constraints using R.
- Designed, developed, and deployed a web application for portfolio management.

iQom Ltd (acquired by Epsilon Net): Data Analyst Intern. — Salonika, GR — 08/2016 – 09/2016
- Performed exploratory data analysis of customer relationship management data using R and communicated results to management.
- Modelled call arrival times using homogeneous Poisson processes.

AWARDS & HONOURS

Full scholarship (tuition + stipend) co-sponsored by Arup and EPSRC. — Cambridge, UK — 10/2019 – 05/2024
- Awarded for MRes + PhD studies at the University of Cambridge.

Commendation Letters in 2/3 course components. — Cambridge, UK — 10/2019 – 09/2020
- Awarded for outstanding performance in MRes course.

Summer research project award (**1000 £**). — Coventry, UK — 06/2017 – 08/2017
- Awarded for outstanding performance in Mathematical Statistics exam.

Merit-based tuition scholarship (**10,000 €**). — Salonika, GR — 09/2013 – 06/2015
- Awarded for academic excellence and performance in Mathematics & English exams.

Honour in Mathematics. Salonika, GR
- Awarded by Hellenic Mathematics Society for performance in nationwide competition. 09/2013 – 05/2013

RESEARCH EXPERIENCE

Generating origin-destination matrices in neural spatial interaction models. Cambridge, UK
Lead author of paper under review (Python code). 08/2023 – 02/2024
- Introduced efficient framework for generating origin-destination matrices leveraging Neural Stochastic Differential Equations (optimisation) and Markov Bases (sampling).

Table inference for combinatorial origin-destination choices in agent-based population synthesis. Cambridge, UK
Lead author of paper published in Stat 2024 (Python code). 09/2021 – 07/2023
- Proposed Markov Chain Monte Carlo algorithm to explore the discrete combinatorial space of origin-destination matrices and their continuous physics-driven representation.

Model assessment of constitutive laws in traffic conservation laws. Cambridge, UK
First year project supervised by Prof. Mark Girolami (Python code). 10/2020 – 08/2021
- Estimated Bayes factors of constitutive laws embedded in traffic flow partial differential equations (PDEs) using thermodynamic integration.

Stochastic modelling of urban travel demand: A *Bayesian* inverse problem perspective. Cambridge, UK
MRes Thesis supervised by Prof. Mark Girolami (**79%**, Python code). 05/2020 – 08/2020
- Implemented Metropolis-Hastings, Hamiltonian Monte Carlo and Annealed Importance Sampling schemes to sample from a doubly intractable posterior distribution.

Bayesian hydrological modelling of road rainfall run-off. Cambridge, UK
Research project supervised by Prof. Mark Girolami (**80%**, Python code). 10/2019 – 01/2020
- Developed *probabilistic* hydrological model comparison and prediction framework using Sequential Monte Carlo.
- Collaborated with National Highways company executives to identify project scope and communicated results to them.

Bayesian online change-point detection for time series segmentation and forecasting in non-stationary point processes. Coventry, UK
Bachelor thesis supervised by Prof. Theodoros Damoulas (**79%**). 01/2018 – 05/2018
- Developed framework for change-point detection in discrete point processes.

Nuclei detection and segmentation from cell images. Coventry, UK
Machine learning project based on Kaggle competition (**84%**). 01/2018 – 04/2018
- Trained multi-layer perceptron and convolutional neural network and compared against heuristic techniques such as Watershed image segmentation.
- Performed data augmentation to achieve translation and rotation-invariance.

Summarising large binary sequences for RNA editing. Coventry, UK
Summer research project supervised by Prof. Anastasia Papavasiliou. 06/2017 – 08/2017
- Leveraged the theory of rough paths to compute signatures of binary representations of RNA sequences using R and Python.

LEADERSHIP ROLES

University of Cambridge Hellenic Society: Captain of Basketball team. Cambridge, UK
- Secured **600** € sponsorship from DeepSea, organised networking events and led the team. 10/2024 – 05/2024

Annual Future Infrastructure & Built Environment Conference: Lead organiser. Cambridge, UK
- Attracted 50 attendees of which **95%** rated their experience as positive and **90%** said they would recommend this conference to colleagues. 09/2022 – 11/2022

Judge Business School EnterpriseTECH: Team communicator. Cambridge, UK
- Developed unique value proposition for an air pollution prediction platform and pitched it to venture capitalists. 12/2019 – 03/2020

University of Warwick Statistics Department: Student Representative & Mentor. Coventry, UK
- Mentored students & liaised with staff to improve teaching quality and student support by collecting and discussing feedback. 10/2015 – 07/2018

University of Warwick and Deutsche Bank: Team lead in software engineering project. Coventry, UK
- Developed a real-time machine learning platform that detected anomalies in one million daily transactions of FTSE100 stocks and pitched our platform to company stakeholders. 01/2017 – 04/2017