Yanni Papandreou

Website: YanniPapandreou.github.io
GitHub: github.com/YanniPapandreou

Email: ypapandreou7@gmail.com
Mobile:+44 (0)783 326 8559
Location: London, UK

Professional Experience

Moody's RMS - Climate Hazard Modeller November 2024 - Present

- Applying the latest academic research to understand and quantify the hazard posed by tropical cyclones.
- Developing R packages, R scripts and Bash scripts for running/analysing tropical cyclone simulations, including developing a bespoke R Shiny dashboard with Quarto to explore simulation results.

Bayforest Technologies - Quantitative Researcher September 2023 - November 2024

- Researching, back-testing, and developing new signals for Bayforest's propriety financial models.
- Experience with data pipelines (using Airflow and AWS): from ingestion to processing and storing.
- Experience writing custom parsers to extract information from the web.

Arabesque AI - Research Intern Aug - Nov 2021

- Investigated approaches for Market Regime Detection utilising Change-point Detection (CPD) algorithms.
- Adapted a kernel CPD algorithm which maps time-series into a high-dimensional feature space in order to detect arbitrary changes in distribution for the original time-series for Market Regime Detection.
- Developed a Python package that combined some of these methods for use by Arabesque.
- Presented my findings to the AI team at the end of the internship and advised further research into Market Regime Detection.

GTA and Tutoring 2018 - 2023

- Graduate Teaching Assistant helping out with tutorials/marking for Maths undergrads and MSc Stats students.
- Online maths tutor with MyTutor UK and Keystone Tutors.

Education

Imperial College London
PhD Mathematics 2019 - 2023
Section: Statistics
Supervisors: Dr Andrew Duncan & Dr Jon Cockayne
Thesis : Bayesian uncertainty quantification for PDE models: new perspectives on inference and dimension reduction
Research interests: Kernel-based methods for inference of complex models, Gaussian processes, Machine learning, Bayesian modelling of differential equations, Probabilistic modelling

Imperial College London
MSc Statistics 2018 - 2019
Grade: **Distinction** (85.7%) Thesis: Kernel-Based Inference Methods for Ordinary Differential Equations (awarded a **distinction**: 84.5%)

University of Cambridge
BA Mathematics 2015 - 2018
Grade: **High 1st Class Honours** (76%)

Technical skills

Programming Languages

- **Core:** Python, R, Nix
- **Intermediate:** Rust, Lua, Fortran, Julia, C++, MATLAB
- **Python Frameworks/Packages:** JAX, Numpyro, Numpy, Pandas, Scipy, PyTorch, Matplotlib, Jupyter, Papermill, Sphinx, Airflow

Tools

- **OS:** Linux (have used several distributions including NixOS, Ubuntu, and Manjaro)
- **Other:** SQL, Quarto, R Shiny, Docker, Airflow, Neovim, Git, Make, Latex

Research and Projects

Publication : Theoretical Guarantees for the Statistical Finite Element Method

- Published in the **SIAM/ASA Journal on Uncertainty Quantification (JUQ)**.
- Presented a new theoretical analysis of the Statistical Finite Element Method (StatFEM), demonstrating that it has similar convergence properties to the finite element method on which it is based.
- Demonstrated a bound on the 2-Wasserstein distance between the ideal prior and posterior and the StatFEM approximation thereof, and showed that this distance converges at the same mesh-dependent rate as the finite element solution converges to the true solution.
- **Coded up simulations in Python** demonstrating the error bounds. This involved using the popular open-source package **FEniCS** to apply the FEM method to various PDEs as well as working with Gaussian Processes using a combination of **Numpy**, **Scipy**, and **Numba**. The Python package **joblib** was utilised to parallelise some of the computations required. The code is publicly available on **GitHub** and can be run as a **Docker** container.

PhD Project : GrassGP

- PDE models often involve a large number of parameters increasing computational burden. Thus, the study of methods to perform dimension reduction are of vital importance.
- Introduced and motivated the concept of a **localised active subspace** for dimension reduction, where the dimension reducing subspace varies smoothly as a function of the input parameters.
- Presented a fully Bayesian model based on **Gaussian Processes**, called **GrassGP**, which can be used to estimate and interpolate localised active subspaces by viewing them as a field of subspace-valued response-variables.
- Provided a **code implementation** of this model in Python using **JAX** and the probabilistic programming package **NumPyro**. The code is publicly available on **GitHub** and my **PhD thesis**, which can be found **online here** , contains a detailed description of this project.

MSc Thesis : Kernel-based Inference Methods for Ordinary Differential Equations

- Investigated the use of Maximum Mean Discrepancy for parameter inference in generative models based on ODEs.
- Studied an adjoint method for gradient descent in high-dimensional parameter spaces.
- Developed **Jupyter Notebooks** for demonstration purposes. **Numpy** and **Scipy** were used to solve the ODEs involved. **Vanilla and Natural Gradient Descent** were utilised for parameter fitting. **PyTorch** was utilised for some experiments involving **neural networks** and for its autodiff capabilities. The code is publicly available on **GitHub**

MSc Coursework

- Sampling and simulation methods such as inverse transform, rejection sampling, and MCMC methods including Metropolis-Hastings and Gibbs samplers, and particle filters (achieved distinction grades of 82.9% and 80.5% on Computational Statistics and Advanced Simulation Methods courseworks respectively). All code for this coursework was implemented in **R**.
- Machine Learning algorithms including: fitting models using Gaussian Processes, binary classification using methods such as logistic regression and generative linear classifiers and PCA (achieved a distinction grade of 83.5% on Machine Learning Coursework). All code for this coursework was implemented in **R**.
- Time series modelling. All code for the coursework was implemented in **MATLAB**.
- Big Data - Map-reduce, Hadoop, Spark (with **Scala**)

Open Source Contributions

- Engage in the **Neovim** plugin community via adding features through pull requests.
- Added a feature to the popular probabilistic programming package **Numpyro** via a successful pull request. My contribution helped enable rendering of deterministic sites for Bayesian Models.
- I am a **Nixpkgs maintainer**; I packaged an open-source Python package in **Nixpkgs** via a successful pull request.

Awards

- **Winton Capital Prize** awarded for best MSc Statistics student (top of class) at Imperial College London.
- **Georges Lemaître Prize** awarded for achieving the highest mark in the Maths Tripos at St Edmund's College, University of Cambridge.
- **Imperial College MSc Stats Challenge - 1st place**: Winning team in the annual challenge where we had to fit a model to noisy financial time-series data in order to optimize the Residual Sum of Squares. Our algorithm later went on to be in the top bracket of Auquan's spring challenge competition.
- **A Level Highest Subject Marks** (2013): Maths (Internationally), Further Maths, and Physics (in Cyprus).
- **A Levels**: Maths (A*), Further Maths (A*), Physics (A*), Chemistry (A*)