Josie Park

An experienced Climate Data Scientist and Mathematician with a PhD in mathematics and ocean modelling

Cambridge, UK · josephine.park@natwest.com
https://josiepark.github.io/

Personal Profile

Experienced Climate Data Scientist with a PhD in Ocean Modelling from Imperial College London and a strong track record of delivering data-driven solutions to complex climate risk challenges. Currently Senior Climate Data Scientist at NatWest Group, where I lead the development of the mortgage climate impact model.
I have led multidisciplinary teams of data scientists, engineers, and researchers in high-pressure environments, fostering collaboration and driving projects to successful delivery.
I bring a deep passion for climate science and a commitment to harnessing data for positive environmental and societal impact. I am now seeking to step into a Lead Climate Data Scientist role, where I can help shape the strategic direction of climate modelling efforts within NatWest Group, mentor junior scientists, and drive innovation at the intersection of climate, data, and finance.

Technical Skills

Python (including pandas, scikit-learn, nltk, tensorflow)
R (including tidyverse, ggplot, rstan)
Fortran
Matlab
Version Control
Cloud Computing (e.g. AWS and Azure)
Linux/Unix and Shell Scripting
High Performance Computing
SQL
CI/CD

Work Experience

March 2023 - Present

Senior Climate Data Scientist — NatWest Group, London

Stream Lead of the Mortgage Climate Impact Model

- Collaborated with key stakeholders in credit risk, group economics and the scenario analysis working group to design the Mortgage Climate Impact Model.
- Initiated the first refactor of the model to conform to OOP principles and adopt software engineering best practices - documentation, modularisation, reproducibility, unit-testing.
- Co-ordinates development and productionisation with MLOPs and Data Engineering.
- Attends networking events and works to create links with other internal teams in NWG - such as retail and NatWest Markets.
- Currently mentoring a data analyst at NWG who wants to move into data science.
- Aims to up-skill the wider CDT team and drive a standardised approach to software development and create common ways of working.

September 2021 - March 2023

Principal Data Scientist (G7) — UKHSA, (formerly JBC), London

Lead Data Scientist of the Epidemiological Ensemble Modelling Team

- Led the team to build long-term capability within the UKHSA to publish national consensus epidemiological estimates and develop an internal modelling group.
- Led discussions and developed working relationships with senior stakeholders, from government, civil service and academia, to finalise the consensus statistic.
- Ran, understood, interpreted and helped develop several epidemiological models, written in a mix of Python, R and Stan, using a range of methods, such as differential equations, Bayesian inference and agent-based modelling.
- Led the standardisation of model packaging and deployment, including Docker-ising the models, using github actions and ensuring models are automatically validated before publication.
- Co-authored two associated papers.

Lead Developer of Test & Trace effectiveness model

- Developed a Markov-Chain Model at pace in response to urgent request from a Senior Stakeholder.
- Wrote the associated technical paper that was published on the gov.uk website.

Data Scientist in the Research Engineering Team

- Scoping out external commercial and academic products to support Covid-19 analysis and modelling,
- Establishing and communicating methods of best practices,
- Carrying out several pieces of QA on key analysis, such as hospital occupancy forecasts.

Josie Park

An experienced Climate Data Scientist and Mathematician with a PhD in mathematics and ocean modelling

Cambridge, UK · josephine.park@natwest.com
https://josiepark.github.io/

November 2019 - September 2021 — **Data Scientist** — Avanade (previously Altius), London

Building re-usable data science assets. Led the development of the time series capability of an object-oriented framework in python to speed up the process of productionising Data Science models.
NLP project processing prescriptions. Used Azure's Read API to automatically process handwriting. An accuracy of 98% with a coverage of 60% was achieved.
NLP project analysing and summarising consultation response data. Used TF-IDF and LDA to build a topic model and detect common keywords. Built a sentiment analysis model and combined this with the topic model to detect topics individuals were positive or negative about. Worked alongside a business analyst who presented our findings in a PowerBI dashboard.
Speech recognition and NLP project to create a tool to automatically process restaurant orders. Developed a Flask App, and associated SQL database storing customer orders as generated using a speech recognition tool.

Summer 2018 — **Summer Research Project** — Met Office, Exeter

Coded and tested a finite volume numerical method and different interpolation methods used for modelling the transport equation and isolated causes of any inaccuracies due to the new grid discretisation developed for the future Met Office unified model, and suggested possible solutions which were then collected into a report.

2017 - 2019 — **Graduate Teaching Assistant** — Imperial College London

Summer 2012 — **Statistician** — Wheelright, Begbroke Science Park

Education

2016 - 2020 — **Ph.D candidate in Applied Mathematics** — Imperial College London, Mathematics of Planet Earth CDT

Thesis Title: Lagrangian Transport in an Idealised Meandering Oceanic Jet.
Josie ran and further developed an idealised ocean model simulating meandering oceanic jets and built an accurate, fully tested transport model which advects Lagrangian particles, employing novel numerical methods developed specifically for this program. Motivated by statistical analysis of the data output, new parameterisation methods were explored, using stochastic Markov models and dimensionality reduction. The numerical models were written in Fortran and were documented and shared among colleagues for use on High Performance Compute clusters. Data analysis was performed in Matlab and Python.

2015 - 2016 — **Masters in Research, Mathematics, Distinction** — Imperial College London, Mathematics of Planet Earth CDT

Courses included:
Data and Uncertainty, Geophysical Fluid Dynamics, Computational Stochastic Processes, Numerical Methods, Dynamical Systems and Partial Differential Equations.

2011-2015 — **M.Sci. Mathematics, 1st Class** — University College London, Department of Mathematics

Thesis Title: Vortex dynamics in a uniform flow around two cylinders.

2009-2011 — **A Levels** — Cherwell Secondary School, Oxford

A* - Maths, A* - Further Maths, A - Physics, B - French, A - Chemistry (AS Level)

Extracurricular Activities

2017 - 2019 — **Scientific Computing Support Network Founder** — Imperial College, London

2013-15 — **President (14-15), Treasurer (13-14) - Won the award for most improved sports club at UCL both years.** — University College London, Swimming Club