ROBERT ARBON

https://orcid.org/0000-0001-6163-3029 | github.com/robertarbon | robert.arbon@gmail.com

SUMMARY

I am a computational chemist and data scientist with over five years' experience of statistical consulting. This includes hypothesis testing, machine learning, physics-based simulations, cheminformatics as well as consulting for physical, life sciences and humanities. I have contributed to open-source projects as well as building a custom ETL pipelines for analysis of large-scale simulation data. When building software I use a clean architecture approach, apply SOLID principles, and am fluent in Git, Github and have experience with CI/CD and AWS cloud infrastructure.

PROFESSIONAL EXPERIENCE

Aug. 2022 – Jul. 2023
Redesign Science, New York, USA
SENIOR SCIENTIST

Aug. 2021 – Aug. 2022
Redesign Science, New York, USA
SENIOR SCIENTIST (PART-TIME)

Aug. 2021 – Aug. 2022
Antonia Mey research group, University of Edinburgh, UK
POST-DOCTORAL RESEARCH FELLOW (PART-TIME)

Feb. 2021 – Jul. 2021
John Chodera Research Group, MSKCC, New York, USA
SOFTWARE SCIENTIST

Jun. 2018 – Jan. 2021
Bristol, UK
FREELANCE DATA SCIENTIST

Apr. 2012 – Sep. 2015
London, UK
PRIVATE TUTOR

Apr. 2008 – Jan. 2012
Economic Consulting Associates Ltd., London, UK
ECONOMICS ANALYST

Sep. 2007 – Mar. 2008
Campbell Carr Ltd., Bucks., UK
ANALYST

Mar. 2007 – Aug. 2007
London Borough of Havering Council, Essex, UK
POLICY ANALYST

Sep. 2000 - Jul. 2001
DSTL (formerly DERA), MoD, Wilts., UK
RESEARCH ASSISTANT, Organic Chemistry Laboratory

EDUCATION

Aug. 2021
UNIVERSITY OF BRISTOL, UK
PhD in computational Chemistry, title: *Markov models of biochemical systems*

Mar. 2007
HARVARD UNIVERSITY, USA
SM[1] in Applied Physics

Jul. 2005
OXFORD UNIVERSITY, UK
MChem in Chemistry (First Class). Dissertation title: *Quantum mechanical effects in liquid He-4*

[1] Equivalent to a UK MSc

ROBERT ARBON

https://orcid.org/0000-0001-6163-3029 | github.com/robertarbon | robert.arbon@gmail.com

SKILLS

Statistics and machine learning:

- Frequentist & Bayesian penalized regression
- Generalized linear & cumulative probability models
- Markov state models, Hidden Markov models
- Machine learning: random forests, support vector machines, multi-task learning, Gaussian process regression, feed forward neural networks
- Clustering and Gaussian mixture models
- Bayesian optimization and stochastic scheduling
- Dimensionality reduction and signal processing
- Time-series analysis
- Randomized Control Trial analysis
- Sensitivity analysis and model selection
- Missing data imputation

Computational chemistry:

- Protein and enzyme modelling
- Protein conformational dynamics and kinetics
- Unbiased molecular dynamics simulations
- Meta-dynamics simulations
- Adaptive sampling
- Markov and Koopman models of molecular dynamics
- Path-integral methods

Programming languages:

- Python (9 years' experience)
- R (6 years' experience)
- Visual Basic for Applications (VBA) (7 years' experience)
- Fortran (4 years' experience)
- SQL (<1 year experience)

Scientific computing:

- **Molecular dynamics**: AMBER, OpenMM and CHARMM
- **Cheminformatics:** PyMol, VMD, Deeptime and PyEmma, mdtraj, RDKit, OEChem TK.
- **Data analysis in Python**: SciPy, Scikit-learn, Pandas, Matplotlib, PyMC, Plotly, PyTorch, Dask.
- **Statistical computing in R**: Tidyverse (ggplot, dplyr etc.), Shiny, BRMS (probabilistic programming) and RMS (multivariable regression modelling)
- **HPC**: MPI for Fortran; Slurm PBS/Torque; and Prefect.

Software contributions:

- My own package **Celerity** (github.com/RobertArbon/online_kinetics). **Celerity** (WIP) is a deep 'online' estimator for models of molecular kinetics.
- New features and bug fixes for **Osprey** (github.com/msmbuilder/osprey). Osprey is a Python library for hyper-parameter optimization.
- New feature for **CHARMM** (charmm.org). CHARMM is a Fortran code for running molecular dynamics simulations on multi-node CPU and GPU computational clusters.
- Bug fixes for **PyEMMA** (github.com/markovmodel/PyEMMA). PyEMMA is a Python/C library of Markov model algorithms
- Bug fixes for **MSMBuilder** (github.com/msmbuilder/msmbuilder). MSBuilder is a Python/C library of Markov model algorithms.

TEACHING EXPERIENCE

1. *Private tutor* 2012 – 2019. I have tutored GCSE and A-level students professionally full-time for four years and part time for three years.
2. *Robust and Open Analysis in R*. 2018 – 2020. Co-designed with Natalie Thurlby a 1-day course on reproducible workflows in R which has been delivered four times in Bristol and was scheduled to be delivered in March 2020 (now pending). Supported by UK Reproducibility Network and Cancer Research UK.
3. *Style Transfer for Biochemists.* 2020. Designed a half-day course on using deep learning techniques for novel visualization of biochemical structures. Supported by CCPBiosim.

4. *Thermodynamics*. 2006 – 2007. Teaching assistant for 1 semester course on third year thermodynamics at Harvard university.

PUBLICATIONS

1. (Under review) **Arbon, R.**; Zhu, Y.; Mey, A. S. J. S; *Markov state models: to optimize or not to optimize.* https://doi.org/10.26434/chemrxiv-2023-rrpkw
2. Bell, K.; Hickel, J.; **Arbon, R**.; Zoomkawala, H. *Which Direction for Sustainable Development? A Time Series Comparison of the Impacts of Redistributive versus Market Policies in Bolivia and South Korea*. Sustainable Development **2023**, 1– 20.
3. Matthews, J. C.; Navasumrit, P.; Wright, M. D.; Chaisatra, K.; Chompoobut, C.; **Arbon, R.;** Khan, M. A. H.; Ruchirawat, M.; Shallcross, D. E. *Aerosol Mass and Size-Resolved Metal Content in Urban Bangkok, Thailand.* Environmental Science and Pollution Research **2022**, 29 (52), 79025–79040.
4. Matheson, E. L.; Smith, H. G.; Lewis-Smith, H.; **Arbon, R. E.;** Diedrichs, P. C. *Game on! A Randomised Controlled Trial Evaluation of Playable Technology in Improving Body Satisfaction and Negative Affect among Adolescents*. New Media & Society **2022**, 24 (12), 2635–2658.
5. **Arbon, R. E**. *Markov Models of Biomolecular Systems*. PhD, University of Bristol, **2021**.
6. Schubert, H.; Morley, K.; Puddy, E. F.; **Arbon, R.**; Findlay, J.; Mounsey, O.; Gould, V. C.; Vass, L.; Evans, M.; Rees, G. M.; Barrett, D. C.; Turner, K. M.; Cogan, T. A.; Avison, M. B.; Reyher, K. K. *Reduced Antibacterial Drug Resistance and BlaCTX-M β-Lactamase Gene Carriage in Cattle-Associated Escherichia Coli at Low Temperatures, at Sites Dominated by Older Animals and on Pastureland: Implications for Surveillance.* Applied and Environmental Microbiology **2021,** 87 (6), e01468-20.
7. Consortium, C. M.; Achdout, H.; Aimon, A.; Alonzi, D. S.; **Arbon, R.**; Bar-David, E.; Barr, H.; Ben-Shmuel, A.; Bennett, J.; Bilenko, V. A.; others. *Open Science Discovery of Potent Non-Covalent SARS-CoV-2 Main Protease Inhibitors*. BioRxiv **2020**, 2020–10.
8. Song, Y.-C., Ingram, S., **Arbon, R.E**., Topping, D.O., Glowacki, D.R., Reid, J.P.,. *Transient cavity dynamics and divergence from the Stokes–Einstein equation in organic aerosol*. Chem. Sci. **2020**, 11, 2999–3006.
9. **Robert E Arbon**, Alex J Jones, Lars A. Bratholm, Tom Mitchell, David R. Glowacki. *Sonifying stochastic walks on biomolecular energy landscapes***.** Proceedings of the International Conference on Audio Display, **2018**.
10. Christopher M. Timperley, **Robert E. Arbon**, Sally A. Saunders, Matthew J. Waters. *Fluorinated phosphorus compounds: Part 6. The synthesis of bis(fluoroalkyl) phosphites and bis(fluoroalkyl) phosphorohalidates***.** Journal of Fluorine Chemistry, **2002** 113 (1), 65-78.
11. Christopher M. Timperley, **Robert E. Arbon**, Michael Bird, Stuart A. Brewer, Malcolm W. Parry, David J. Sellers, Colin R. Willis *Bis(fluoroalkyl)acrylic and methacrylic phosphate monomers, their polymers and some of their properties***.** Journal of Fluorine Chemistry. **2003**, 121 (1), 23-31.

OTHER OUTPUT

1. **Robert E. Arbon**, George Holloway, Alex J. Jones, Peter Bennett, The Ligeti Quartet, 2019. *Metastable impressions: Artistic representations of Molecular Dynamics*. https://www.youtube.com/watch?v=zmXO6AZMOYU
2. **Arbon, R**., Drax, K., Thurlby, N., Timpson, N. J., Northstone, K., Brown, K. R., … Munafo, M. R. *(2019, February 27). MAPS: Mapping the Analytical Paths of a Crowdsourced Data Analysis.* Retrieved from osf.io/7vq3b. Experimental preregistration.