Affiliation
Big Data Institute,
Nuffield Department of
Population Health,
University of Oxford,
Oxford, UK

Tel. & Skype
07881 306097
TomMaullin

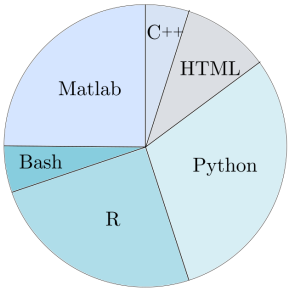
Mail
Thomas.Maullin-Sapey@
bdi.ox.ac.uk

Website
TomMaullin.com

GitHub
github.com/TomMaullin

LinkedIn
linkedin.com/in/TomMaullin/

Programming



Thomas J. Maullin-Sapey

Researcher in Neuroimaging Statistics

About

At present, I am a postdoctoral researcher at the Big Data Institute, University of Oxford. My research interests include distributed machine learning, spatial statistics, linear mixed models, and random field theory. Most recently, I have worked on establishing probabilistic bounds for random excursion sets and introducing novel derivations for mixed model parameter estimation.

For detailed information, publication lists, and regular updates, please visit my academic website at TomMaullin.com.

Research Overview

Linear Mixed Models: The analysis of longitudinal, heterogeneous or unbalanced clustered data is central to neuroimaging analysis. Such analyses are typically conducted using the Linear Mixed Model (LMM) but, over time, the breadth of models considered to fall under the umbrella term 'LMM' has expanded to include more complex, multi-factored designs. In my research on this topic, I extend classic methods for LMM parameter estimation and inference to encompass the modern-day definition. Applications include large-scale computation, privacy-protected algorithms, and a clinical study investigating the link between alcohol consumption and brain function ([1,2,3,4]).

Random Field Theory: A random field is a continuously differentiable function which varies over space, like the surface of the ocean or heat in a weather map. Random fields are central to a range of applications such as climatology, neuroscience and cosmology. In forthcoming work in this area, in the context of simple signal plus noise models, we derive minimal regularity conditions under which estimated topological features of random fields (such as Morse points and Betti numbers) resemble those of truth, and provide novel central limit theorems for the size of random clusters.

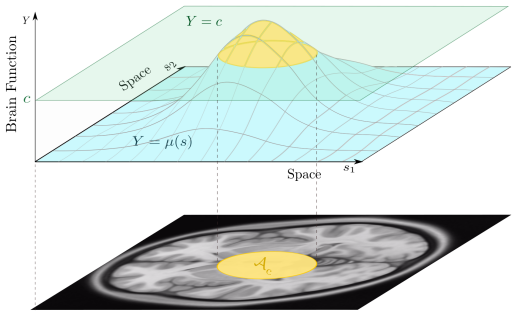


Figure 1: Illustration of an excursion set, $\mathcal{A}_c$, being derived from a random field $\mu(s)$. In fMRI, $\mathcal{A}_c$ is projected onto a T_1 image of brain anatomy to assess where activation occurs within the brain.

Confidence Regions: Researchers often are interested in questions of the form "At which locations does a signal, μ, exceed a certain value c?". For example, "Where did large temperature changes occur in a geospatial heatmap?" or "Where do changes in blood flow in the brain suggest cognition?". Such questions translate formally into the study of excursion sets, sets of the form $\mathcal{A}_c := \{s : \mu(s) \geq c\}$ (Fig. 1). Confidence Regions (CRs) illustrate the variability in estimating such regions, providing bounds on their shape and size. Presently, I am involved in a range of CR-themed projects including; (i) generating CRs for piece-wise continuous processes, (ii) using CRs to assist presurgical planning and (iii) deriving 'cluster-specific' CR coverage statements. In previous work, I derived novel CR methodology for bounding intersections and unions of excursion sets ([5]).

Academic Writing

- Basodi, S., Rajikha, R., Gazula, H., Romero, J., Panta ,S., Maullin-Sapey T., Nichols, T.E. & Calhoun, V. (2024). Decentralized Mixed Effects Modeling in COINSTAC. Neuroinformatics, 22(2). https://doi.org/10.1007/s12021-024-09657-7
- Maullin-Sapey, T., Schwartzman, A., & Nichols, T.E. (2023). Spatial Confidence Regions for Combinations of Excursion Sets in Image Analysis. Journal of the Royal Statistical Society Series B: Statistical Methodology 86(1). https://doi.org/10.1093/jrsssb/qkad104
- Maullin-Sapey, T. (2022). Multi-level Modelling and Spatial Inference for Large-Scale Neuroimaging Data. Postgraduate Thesis. https://ora.ox.ac.uk/objects/uuid:425fac44-6413-47cc-b173-0a78c26aaf1c
- Topiwala, A., Ebmeier, K.P., Maullin-Sapey, T., & Nichols, T.E. (2022). Alcohol consumption and MRI markers of brain structure and function: Cohort study of 25,378 UK Biobank participants. NeuroImage: Clinical, 35, 103066. https://doi.org/10.1016/j.nicl.2022.103066
- Maullin-Sapey, T., & Nichols, T.E. (2022). BLMM: Parallelised computing for big linear mixed models. NeuroImage, 264, 119729. https://doi.org/10.1016/j.neuroimage.2022.119729
- Maullin-Sapey, T, & Nichols, T.E. (2021). Fisher Scoring for Crossed Factor Linear Mixed Models. Statistics and Computing, 31, 53. https://doi.org/10.1007/s11222-021-10026-6
- Maullin-Sapey, T. (2017). Visualisation and Integration of Multiple Brain Imaging Studies. Undergraduate Dissertation. https://doi.org/10.6084/m9.figshare.21893061.v1

Conference Posters

- Maullin-Sapey, T., Schwartzman, A., & Nichols, T. E. (2022). Spatial Confidence Regions for Conjunctions of fMRI Effects. OHBM & RSS Conference Poster. https://doi.org/10.6084/m9.figshare.21836901.v1
- Maullin-Sapey, T, & Nichols, T.E. (2021). Fisher Scoring for Crossed Factor Linear Mixed Models. Statistics and Computing, 31, 53. https://doi.org/10.1007/s11222-021-10026-6
- Maullin-Sapey, T., & Nichols, T.E. (2021). BLMM: Parallelised Computing for Big Linear Mixed Models. SMI Conference Poster.https://doi.org/10.6084/m9.figshare.21842097.v1
- Maullin-Sapey, T., & Nichols, T. E. (2020). BLMM: Parallelised Computing for Big Linear Mixed Models. OHBM Conference Poster. https://doi.org/10.6084/m9.figshare.21842172.v1
- Maullin-Sapey, T., & Nichols, T. E. (2019). BLM: Parallelized Computing for Big Linear Models. OHBM Conference Poster. https://doi.org/10.6084/m9.figshare.21842208.v1
- Maullin-Sapey, T., Maumet, C., & Nichols, T.E. (2018). Detecting and Interpreting Heterogeneity and Publication Bias in Image-Based MetaAnalyses. OHBM Conference Poster. https://doi.org/10.6084/m9.figshare.21842349.v1
- Maullin-Sapey, T. (2017). Visualisation and Integration of Multiple Brain Imaging Studies. Undergraduate Dissertation. https://doi.org/10.6084/m9.figshare.21893061.v1
- Maullin-Sapey, T. et al. (2017). Viewing FSL results with SPM and vice versa. OHBM Conference Poster. https://doi.org/10.6084/m9.figshare.21861906.v1

Selected Talks

- Maullin-Sapey, T. (2024, April 26). Spatial Confidence Regions and Uncertainty in fMRI. Statistics Seminar Series at the School of Mathematics, University of Bristol. https://doi.org/10.6084/m9.figshare.25709775.v1
- Maullin-Sapey, T. (2024, January 18). Small Sample Studies in fMRI Neuroscience. Thematic Einstein Seminar (TES) series on small data analysis, Zuse Institute Berlin (ZIB). https://doi.org/10.6084/m9.figshare.25709778.v1
- Maullin-Sapey, T. (2022, November 10). BLMM: A toolbox for parameter estimation and inference on big linear mixed models on NeuroImaging data. University of California Spatial Statistics Reading Group. https://doi.org/10.6084/m9.figshare.21892947.v1
- Maullin-Sapey, T. (2022, September 13). Rapid Fire Introduction to Spatial Confidence Regions. Royal Statistical Society Rapid Fire Sessions. https://doi.org/10.6084/m9.figshare.21892980.v1
- Maullin-Sapey, T. (2022, June 6). Spatial Confidence Regions for Combinations of Excursion Sets in Image Analysis. FMRIb Reading Group. https://doi.org/10.6084/m9.figshare.21892995.v1

Qualifications

2018 - 2022: **Doctoral Degree, University of Oxford:**

My doctoral research pursued two distinct avenues of neuroimaging methods development. The first avenue focused on the derivation of novel expressions for estimation and inference of the Linear Mixed Model (LMM). These expressions were employed to perform computationally efficient large-scale fMRI LMM analyses. The second research direction centred on extending an existing theory of confidence regions to provide bounds for the intersection, or 'conjunction', of excursion sets that have been acquired across the same spatial region but under different study conditions. Such 'conjunctions' are of natural interest in image analysis as they correspond to the question "Where was brain activation observed in the same place across multiple study conditions?".

2014 - 2017: **Undergraduate Degree, University of Warwick:**

I hold a 2:1 in BSc Data Science (Hons) from the University of Warwick. This course was comprised of modules from the Computer Science, Mathematics and Statistics programs, respectively. Through my courses, I developed a broad range of skills relevant to statistics and computing, including the ability to efficiently program in Java, R and SQL, as well as logical deduction, proof methods and the understanding and application of several widely-used statistical models.

Selected Work Experience

2022-Present: **Postdoctoral Researcher in Neuroimaging Statistics**
Big Data Institute, University of Oxford, OX3 7LF

Roles: As a postdoctoral researcher, I was funded by an NIH R01 award for a project that focused on spatial confidence regions for fMRI results. During my time in this position, I have published multiple papers, attended conferences such as *Organisation for Human Brain Mapping* and *Royal Statistical Society*, presented posters and talks, and submitted multiple grant applications, the results of which I am currently waiting on.

Skills acquired: Experience publishing academic work, preparing grants and building research networks; strengthened knowledge of topology, probability theory and manifolds.

2020-Present: **Graduate Teaching for Health and Data Sciences CDT programme**
Big Data Institute, University of Oxford, OX3 7LF

Roles: Since 2020, I have taught on the Health and Data Sciences CDT programme at the Big Data Institute, University of Oxford. In particular, I organize and run the 'Introduction to Python' and 'Imaging with Nilearn' intensive crash courses, mentor small teams of students in statistical analysis on the 'Data Challenge' course, give the introductory fMRI lectures, and have previously assisted on the 'Computational Statistics' and 'Modern Statistical Methods' courses. For the 2020 'Data Challenge', my role additionally entailed extensive preparation and analysis of NHS Covid-19 data for the students to further investigate. A full list of my teaching, alongside course materials may be found here.

Skills acquired: Experience preparing educational content, experience teaching students across different backgrounds and with varying skill levels, effective communication skills, practical experience working with NHS Covid-19 data.

2019-Present: **Graduate Teaching for Interdisciplinary Bioscience CDT programme**
Doctoral Training Centre, University of Oxford, OX1 3NP

Roles: Since 2023, I have lead the 'Introduction to Machine Learning' and 'Statistical Methods' classes for the Interdisciplinary Bioscience CDT programme, before which I tutored on the previous iterations of the courses. I now design the curriculum for these short, intensive modules, prepare the course material and deliver the practicals and lectures. These are introductory classes for students to learn about CNN's and regression modelling. Both classes are given in the programming language R.

Skills acquired: Teaching experience, communication skills.

2017-2018: **Neuroimaging Research Assistant**
Big Data Institute, University of Oxford, OX3 7LF

Roles: Writing academic papers, creating and presenting conference posters, maintaining software.

Skills acquired: Experience with Bash, Python, CI testing and Docker containers. Understanding of random field theory, functional connectivity and time series methods used in neuroimaging.