Bayesian hierarchical modeling and analysis of spatial data
Project III

Advisor: Georgios P. Karagiannis

Academic year 2026-2027 @ Durham University

General description

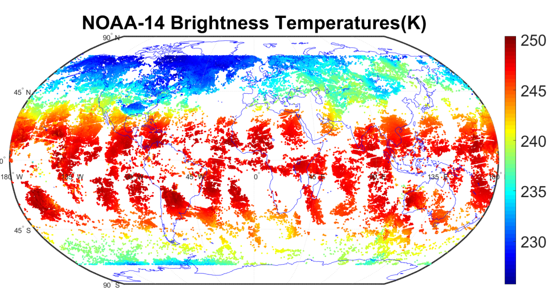
This project focuses on statistical methodology and computational tools for the analysis of spatial data, primarily within a Bayesian framework.

Spatial datasets record both measured values and the locations at which those values were collected. Such data arise in environmental science, climatology, weather forecasting, epidemiology, engineering, and biology. Spatial statistical methods use these data to construct probabilistic models for estimation, prediction, and scientific interpretation.

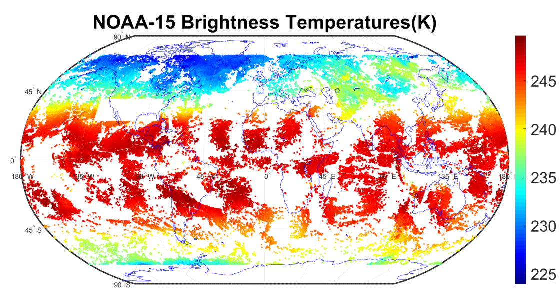
A Bayesian hierarchical modelling approach is particularly useful because it allows uncertainty to be represented explicitly at multiple levels of a model, while also supporting flexible dependence structures across space.

Project-specific intended learning outcomes

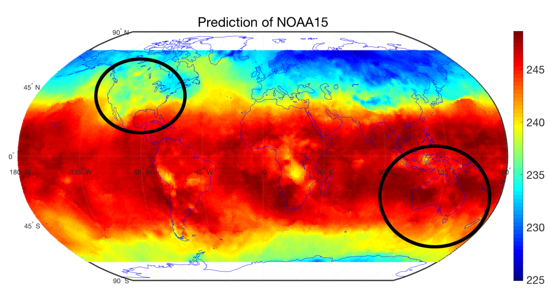
By the end of this project, students should be able to design suitable Bayesian hierarchical models over spatial or spatio-temporal stochastic domains; apply these models to real spatial datasets; and implement computational tools for fitting, prediction, and uncertainty quantification using appropriate statistical software.



(1)



(2)



(3)

Borrowed from: Cheng, S., Konomi, B. A., Matthews, J. L., Karagianis, G., & Kang, E. L. (2020). Hierarchical Bayesian Nearest Neighbor Co-Kriging Gaussian Process Models; An Application to Intersatellite Calibration. arXiv preprint arXiv:2004.01341.

Group project

The group project will build a shared foundation in the theory and practice of Bayesian modelling for spatial data. Students will work together to understand how dependence across space can be represented statistically, and how hierarchical models can be used for prediction, smoothing, calibration, and inference in real applications.

By the end of the group project you will have learned:

- **Spatial data:** Types of spatial data * Spatial data in R * The sf and terra packages for spatial data * Making maps with R * R packages to download open spatial data
- **Areal data:** Spatial neighborhood matrices * Spatial autocorrelation * Bayesian spatial models * Disease risk modeling * Areal data issues
- **Geostatistical data:** Geostatistical data * Spatial interpolation methods * Kriging * Model-based geostatistics * Methods assessment
- **Inference:** Bayesian models for spatial and spatial processes * INLA

By the end of the group project you will have practiced:

- reading and discussing methodological literature in Bayesian spatial statistics;
- formulating and comparing candidate models for spatial datasets;
- implementing model-fitting workflows in appropriate statistical software;
- interpreting fitted models, predictions, and uncertainty summaries;
- communicating statistical ideas clearly in written and oral form.

Mode of Operation and Evidence of Learning:

- The group project will operate through a combination of guided reading, collaborative discussion, mathematical modelling, data analysis, and computational implementation. Evidence of learning will come from engagement with the core literature, development of appropriate hierarchical model formulations, practical analysis of spatial datasets, interpretation of outputs, and clear communication of methods and findings in both written and oral formats.

Individual project

The individual project will build on the common foundation developed in the group component. Each student will choose a more focused direction and investigate it in greater depth from a methodological, computational, or applied perspective. Possible directions include:

- multivariate spatial methods, where several related variables are observed and their joint dependence must be modelled;
- large-data methods, where computational efficiency becomes central because of image-scale or high-resolution datasets;
- multi-resolution or multi-source data methods, where information from sources with different quality or resolution must be combined;
- analysis of real datasets, for example in disease mapping, atmospheric variables, CO_2 concentration, or remote sensing.

Mode of Operation and Evidence of Learning:

- The individual project will involve independent reading, sustained critical investigation, and a substantial piece of statistical work on a focused topic. Depending on the topic selected, this may include methodological comparison, simulation design, real-data analysis, algorithmic implementation, or a structured review of the literature. Evidence of learning will be demonstrated through a coherent written report, supported where appropriate by figures, tables, code, and critical discussion of modelling choices, assumptions, limitations, and conclusions.

Requirements

- Bayesian statistics and/or Statistical modelling (co-requisites)

References

- Moraga, P. (2023). Spatial Statistics for Data Science: Theory and Practice with R. CRC Press.
- Gómez-Rubio, V. (2020). Bayesian inference with INLA. Chapman and Hall/CRC.
- Blangiardo, M., & Cameletti, M. (2015). Spatial and spatio-temporal Bayesian models with R-INLA. John Wiley & Sons.
- Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2014). Hierarchical modeling and analysis for spatial data. CRC press.

Contact details

For further information, feel free to contact

Georgios Karagiannis (Office MCS3088)
Email: georgios.karagiannis@durham.ac.uk
Telephone: +44 (0) 1913342718