

Day 1 – Spatial & Spatio-temporal Modelling

About me

- ▶ Completed a Bachelor's degree in Statistics at FUTA
- ▶ Master's in Mathematical Sciences at AIMS-Tanzania
- ▶ PhD in Statistics and Epidemiology at University of Lancaster
- ▶ Postdoc at University of Manchester
- ▶ Lecturer in Statistics at the University of Manchester

Overview of the 3 days

- ▶ Spatial and spatio-temporal analysis (different likelihoods)
- ▶ Joint modelling of multiple malaria processes
- ▶ Non-stationary spatial processes
- ▶ Hybrid machine learning + geostatistical models

Linear Regression

- ▶ Goal: Model a continuous response variable as a linear function of predictors.
- ▶ Model $Y = X\beta + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$
- ▶ Key Assumption
 - ▶ Linearity
 - ▶ Independence
 - ▶ Homoscedasticity (constant variance)
 - ▶ Normality of errors

Generalized Linear Models (GLMs)

- ▶ Extension of linear models to handle non-normal response distributions.
- ▶ Three components:
 - ▶ Random component: Distribution from the exponential family (e.g., Binomial, Poisson).
 - ▶ Systematic component: Linear predictor $\eta = X\beta$.
 - ▶ Link function: Relates $\mathbb{E}(Y) \sim \text{to} \sim \eta$.
- ▶ Examples:
 - ▶ Logistic regression for binary outcomes – logit link function
 - ▶ Poisson regression for count data – log link function

Why Spatial Statistics?

► **Spatial Dependence:**

Observations collected at nearby locations are often more similar than those farther apart.

► **Ignoring Spatial Structure:**

- ▶ Leads to biased parameter estimates.
- ▶ Underestimates uncertainty.
- ▶ Misses important spatial patterns.

► **Applications:**

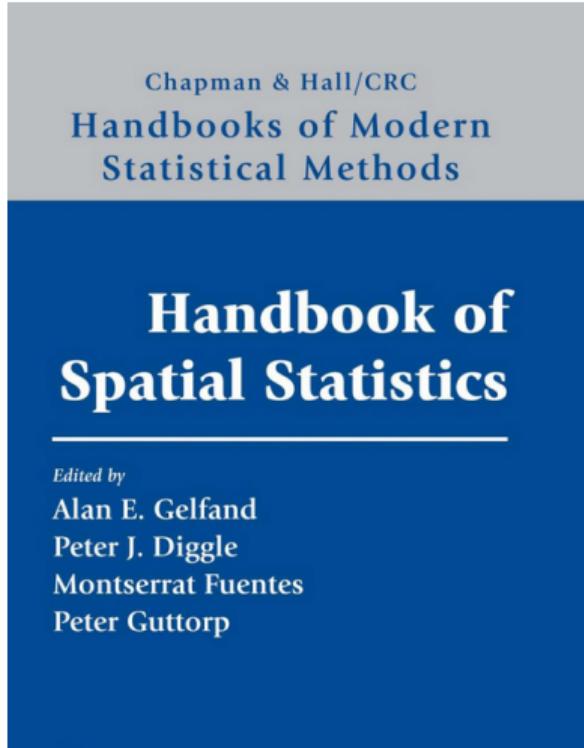
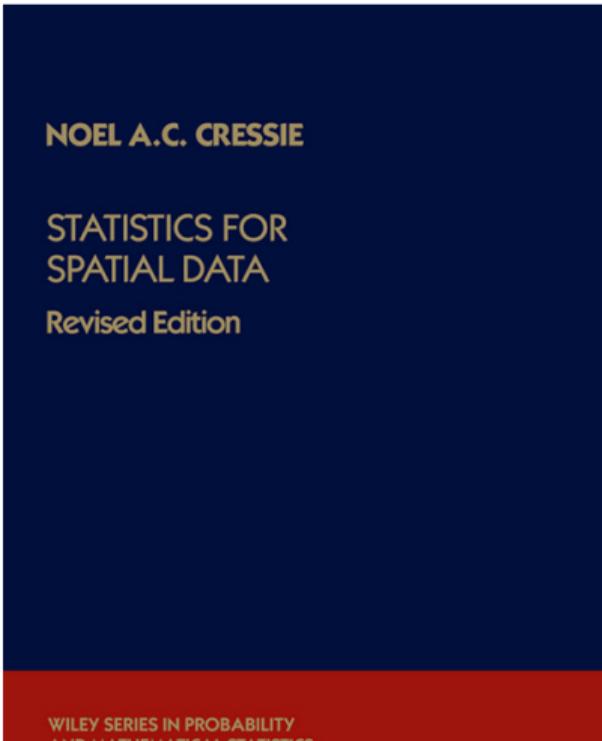
- ▶ Disease mapping
- ▶ Environmental monitoring
- ▶ Agricultural field trials

► **Goal:**

Account for spatial correlation to improve prediction and inference.

Spatial Analysis

Spatial Statistics



Classification of spatial statistics

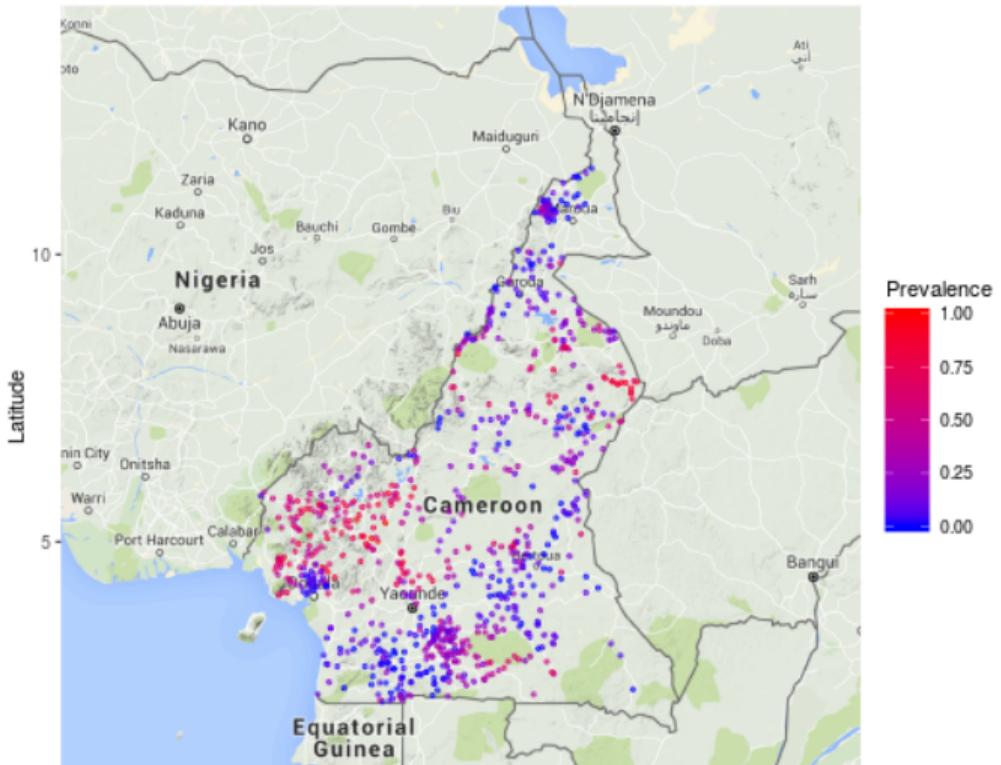
Cressie's book classifies spatial statistics according to **data format**:

1. Geostatistical data
2. Lattice data
3. Point patterns

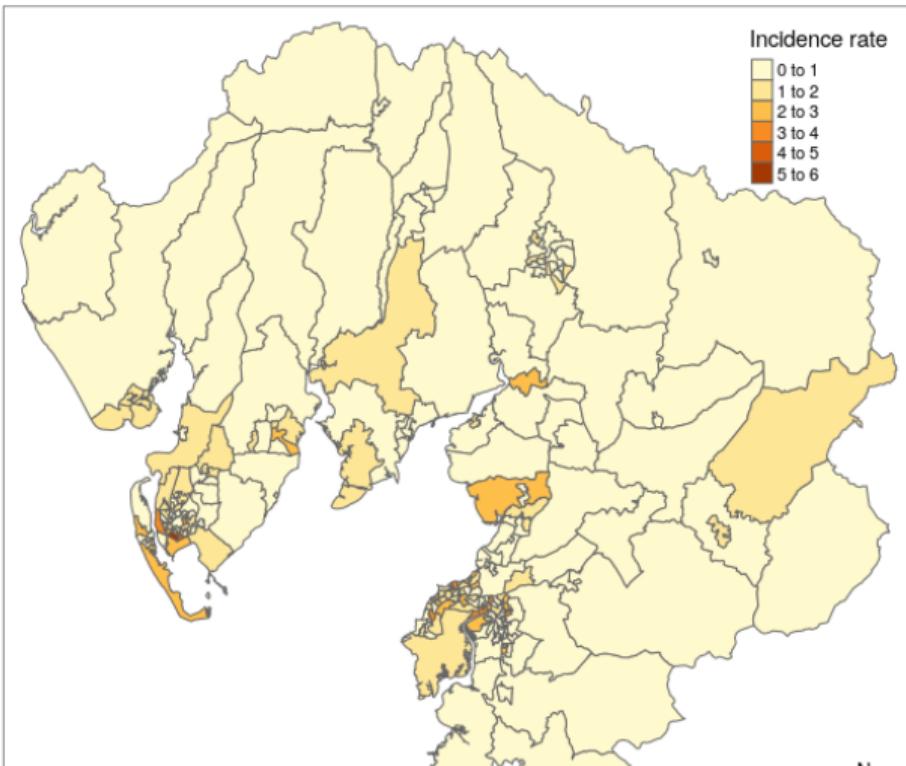
Gelfand's book classifies spatial statistics according to **spatial variation**:

1. Discrete spatial variation
2. Continuous spatial variation

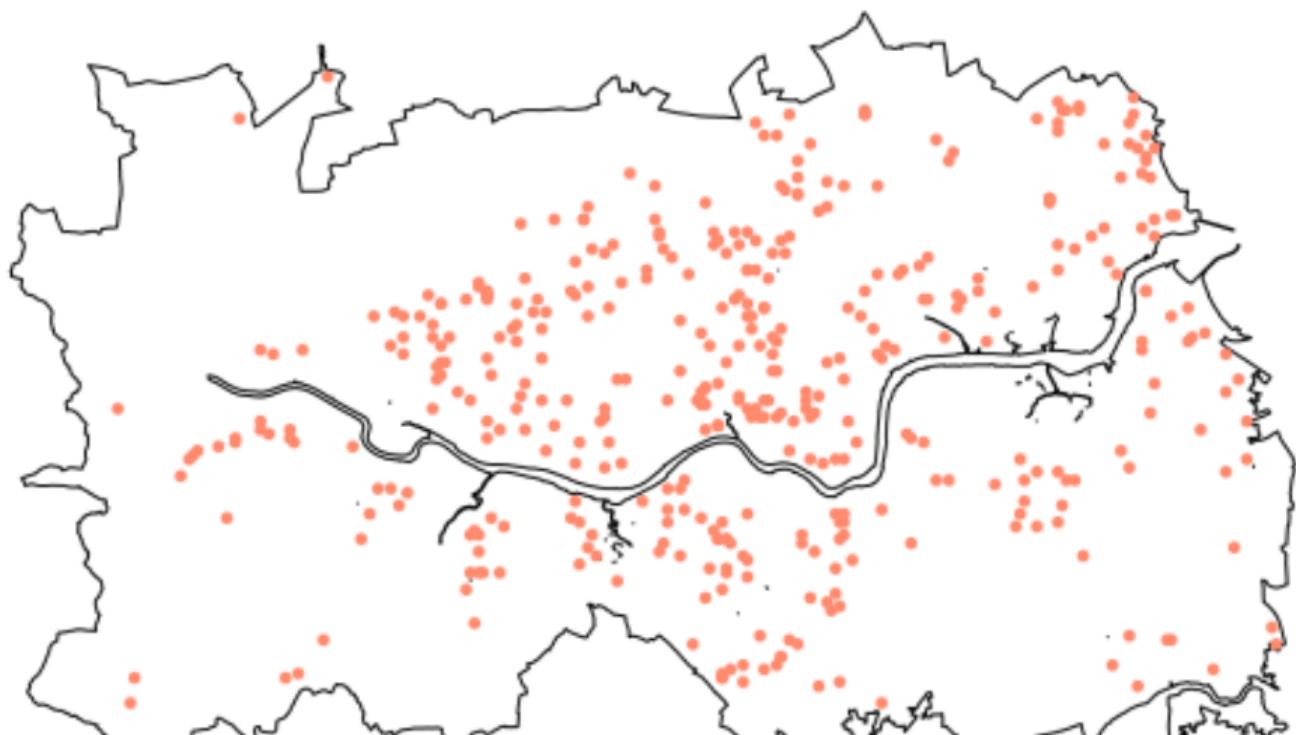
Geostatistical data: River blindness in Cameroon



Lattice Data: COPD emergency admission



Point pattern: Primary biliary cirrhosis data



Model-based Geostatistics

Modelling Geostatistical Data – Model-based Geostatistics

- ▶ The term **Model-based Geostatistics (MBG)** was coined by Peter Diggle in 1998 (Diggle, Tawn, and Moyeed 1998; Diggle and Giorgi 2019).
- ▶ MBG applies general principles of statistical modelling and inference to the analysis of geostatistical data.
- ▶ It emphasises the use of likelihood-based inference.
- ▶ and the use of a latent spatial process (Gaussian or stochastic process)

References

- Diggle, Peter, and Emanuele Giorgi. 2019. *Model-Based Geostatistics*. CRC Press.
- Diggle, Peter, Jonathan Tawn, and Rana Moyeed. 1998. "Model-Based Geostatistics." *Journal of the Royal Statistical Society: Series C* 47: 299–350.