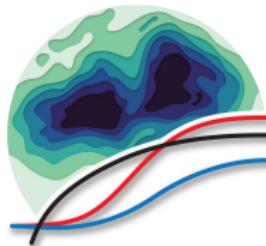


Geospatial Methodologies in Toxicology

Linking exposure, toxicity, and disease profiles to identify U.S.
regions at elevated health risks

Kyle P Messier, PhD

National Institute of Environmental Health Sciences - Division of Translational
Toxicology - Predictive Toxicology Branch



Spatiotemporal Exposures
and Toxicology (SET) Group

Introduction

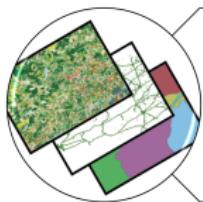
About Us

Spatiotemporal Exposures and Toxicology {SET} group

- Spatiotemporal Exposure Mapping
- Chemical and Stressor Mixtures Prediction
- Mechanistically Informed Geospatial Risk Assessment

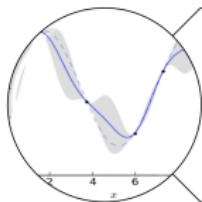


Geospatial Methods



Translational

- Maps
- Integrative



Predictive

- Interpolation
- Uncertainty

Objective

- Provide an overview of geospatial methods, data, applications, and future directions in toxicology and risk assessment

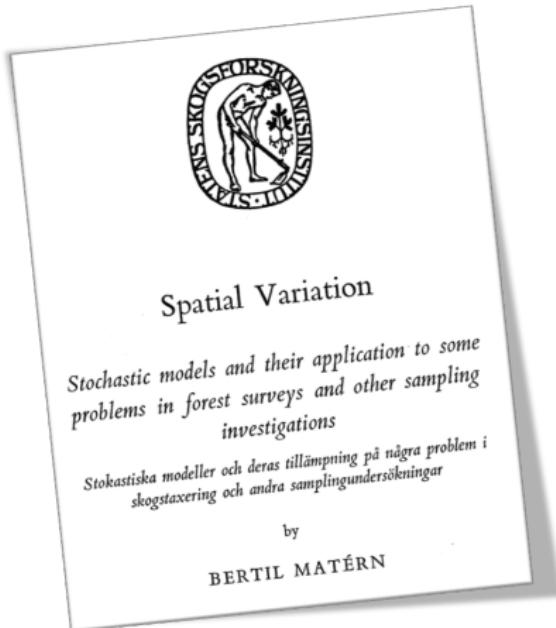
History

Mining



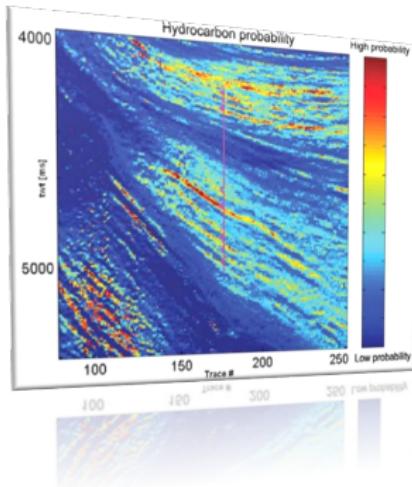
- Matheron and Krige developed geostatistical methods to predict ore content from core samples
- Matheron coined the term “Kriging” after Krige
- “Nugget” is a term used to random noise because predicting where gold nuggets were was so difficult

Forestry



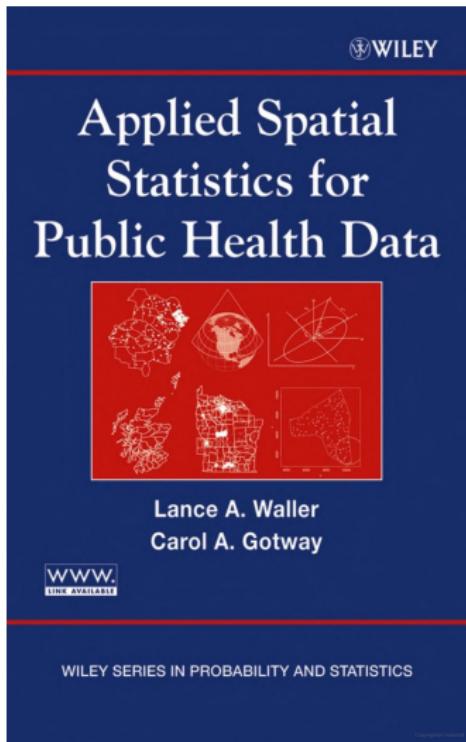
- Matérn developed correlation models for spatial variation for applications in Forestry
- To this day, we use the “Matérn” covariance function

Petroleum Engineering



- Used to evaluate the oil and gas field reservoirs
- Uses geology and seismic data

Public Health



- Cressie, 1990: Statistics for Spatial Data
- Waller and Gotway, 2004: Applied Statistics for Public Health Data
- Wide scale adoption for statisticians and engineers in ecological and human exposure and risk applications

Toxicology



- **Toxicology is a new frontier for geospatial methods**
- Aggregate Exposure Pathways
- Adverse Outcome Pathway
- GeoTox
- Source-to-Outcome

Classical Models

Uses in Public Health

- Estimate exposure to air pollutants, water contaminants, and other environmental stressors
- Geocode patient addresses to link to environmental exposures
- Estimate the spatial distribution of disease rates
- Estimate the spatial distribution of health risk factors

Land Use Regression

Linear regression for spatial data

$$Y(s) = X(s)\beta + \varepsilon$$

where $Y(s)$ is the response variable, $X(s)$ are the predictor variables, β are the regression coefficients, ε is the iid error term, and (s) denotes the spatial location.

Not a terrible idea for spatial data, but it directly violates the assumption of independence of observations.

Kriging

Kriging and spatial models provide an explicit term for spatial correlation. A reasonable approach is a random-effect model:

$$Y(s) = \mu(s) + \varepsilon + \eta(s)$$

where $\eta \sim N_n(0, \Sigma_\theta)$

and Σ_θ is a covariance matrix with parameters, θ , that accounts for correlation between spatial and temporal locations

Data Sources

Common Data Sources and Types

- U.S. Census Bureau
- U.S. Environmental Protection Agency
- U.S. Geological Survey
- National Aeronautics and Space Administration
- National Oceanic and Atmospheric Administration
- U.S. Department of Agriculture
- Land cover data
- Health statistics
- Population characteristics
- Infrastructure data
- Air quality data
- Water quality data
- Satellite imagery