Hierarchical spatial modelling for applied population and community ecology

Jeffrey W. Doser, Marc Kéry, Gesa von Hirschheydt

24-27 June 2024

Introduction to hierarchical spatial modelling

Jeffrey W. Doser

24-27 June 2024

What is spatial data?

What is spatial data?

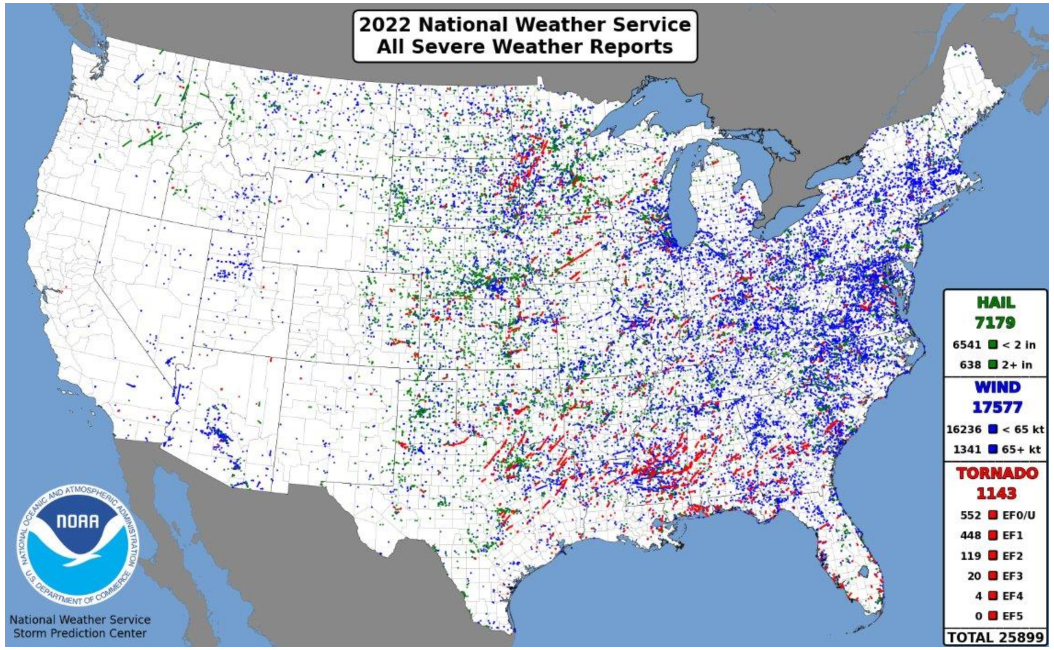
Any type of data that relates to a specific geographical area or location

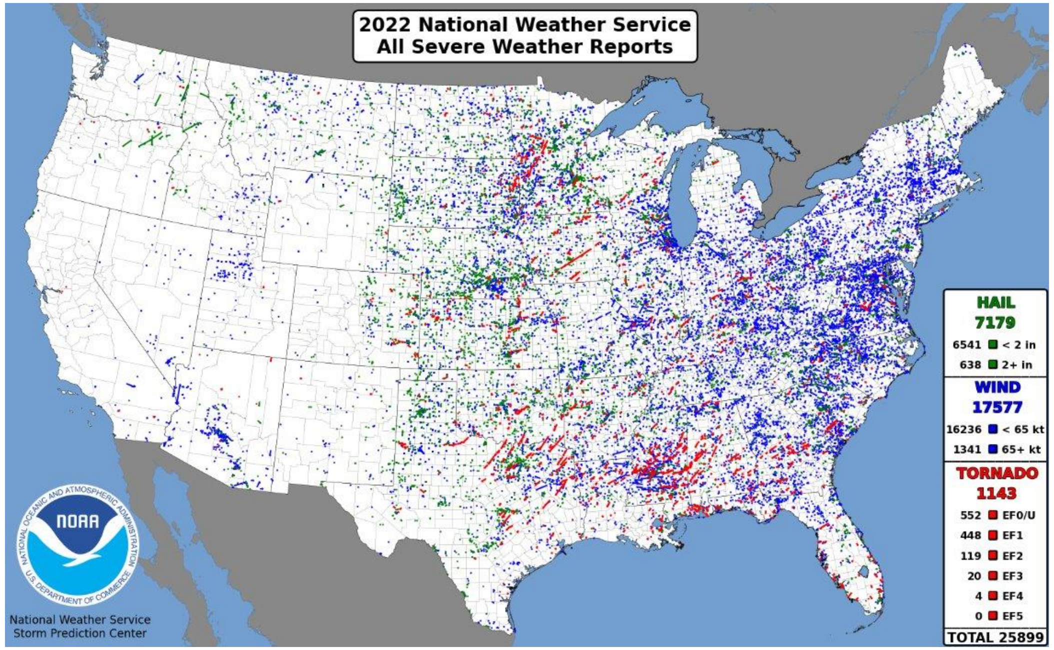
What is spatial data?

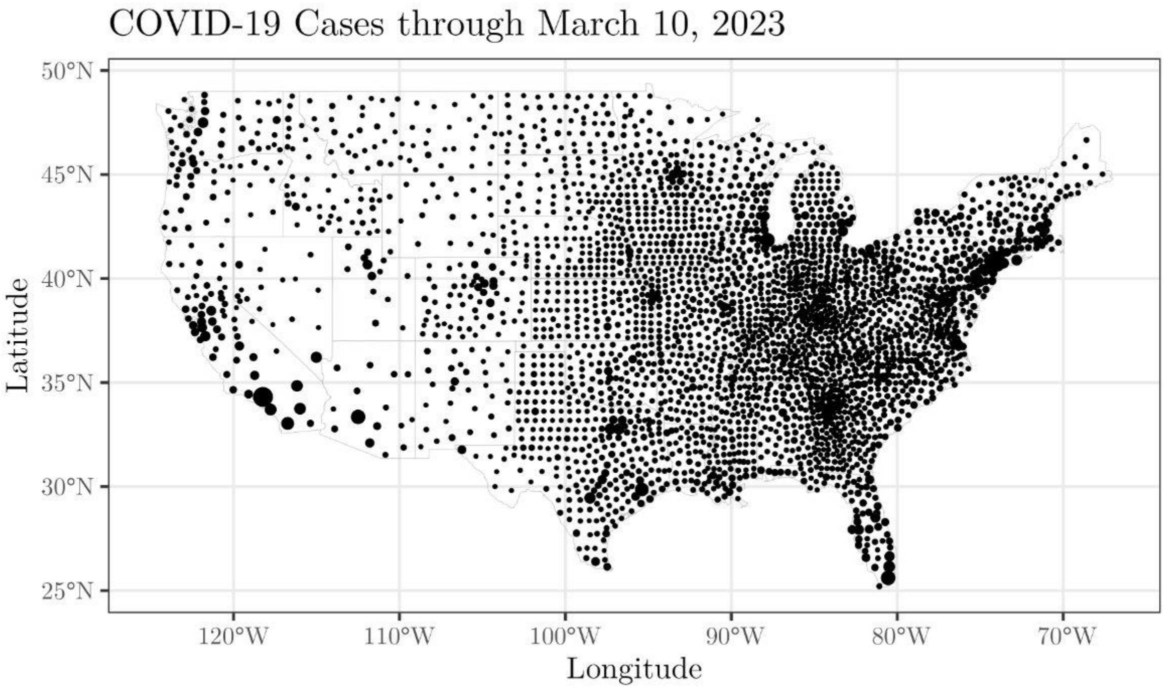
Any type of data that relates to a specific geographical area or location

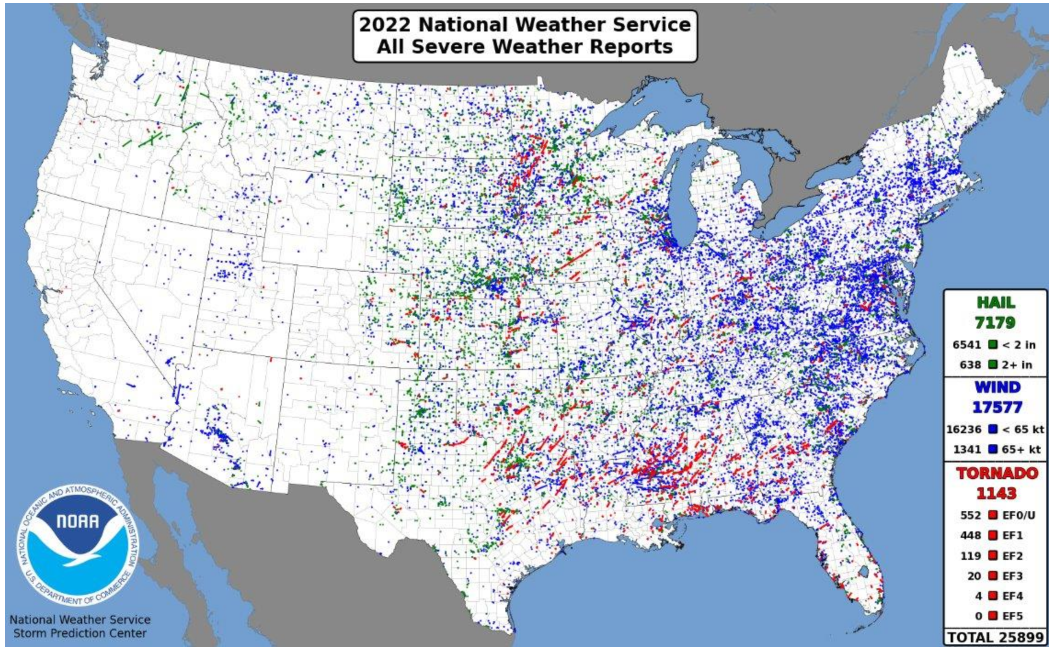
First Law of Geography

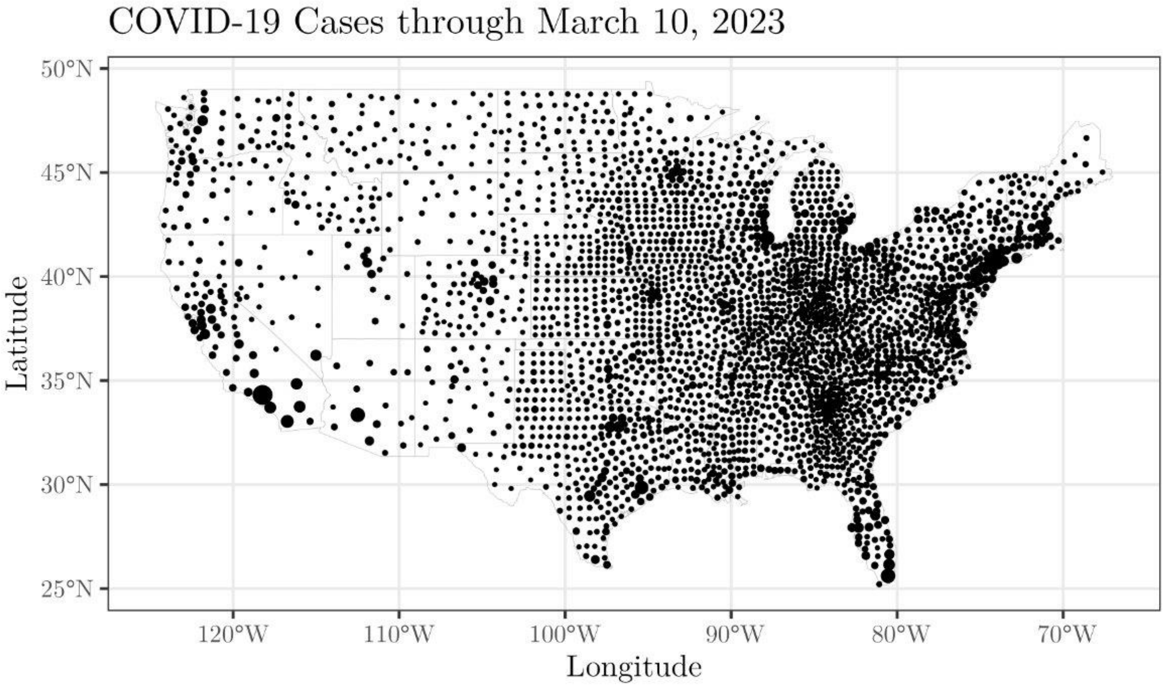
"Everything is related to everything else, but near things are more related than distant things." - Waldo Tobler

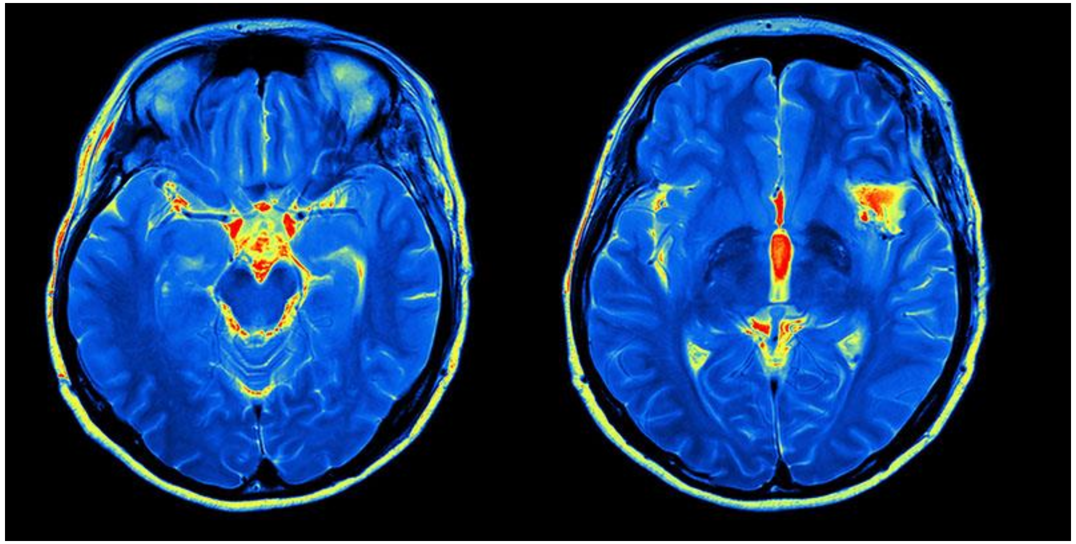

2022 National Weather Service
All Severe Weather Reports
HAIL
7179
6541 < 2 in
638 2+ in
WIND
17577
16236 < 65 kt
1341 65+ kt
TORNADO
1143
552 EF0/U
448 EF1
119 EF2
20 EF3
4 EF4
0 EF5
TOTAL 25899
NOAA
National Weather Service
Storm Prediction Center

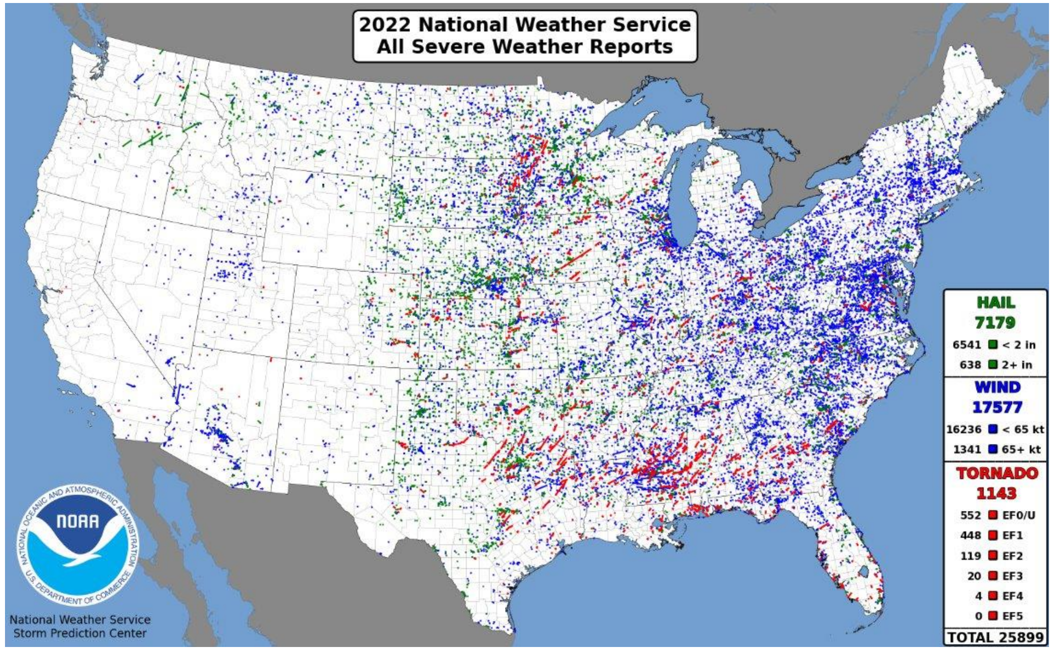

2022 National Weather Service
All Severe Weather Reports
HAIL
7179
6541 < 2 in
638 2+ in
WIND
17577
16236 < 65 kt
1341 65+ kt
TORNADO
1143
552 EF0/U
448 EF1
119 EF2
20 EF3
4 EF4
0 EF5
TOTAL 25899
NOAA
National Weather Service
Storm Prediction Center

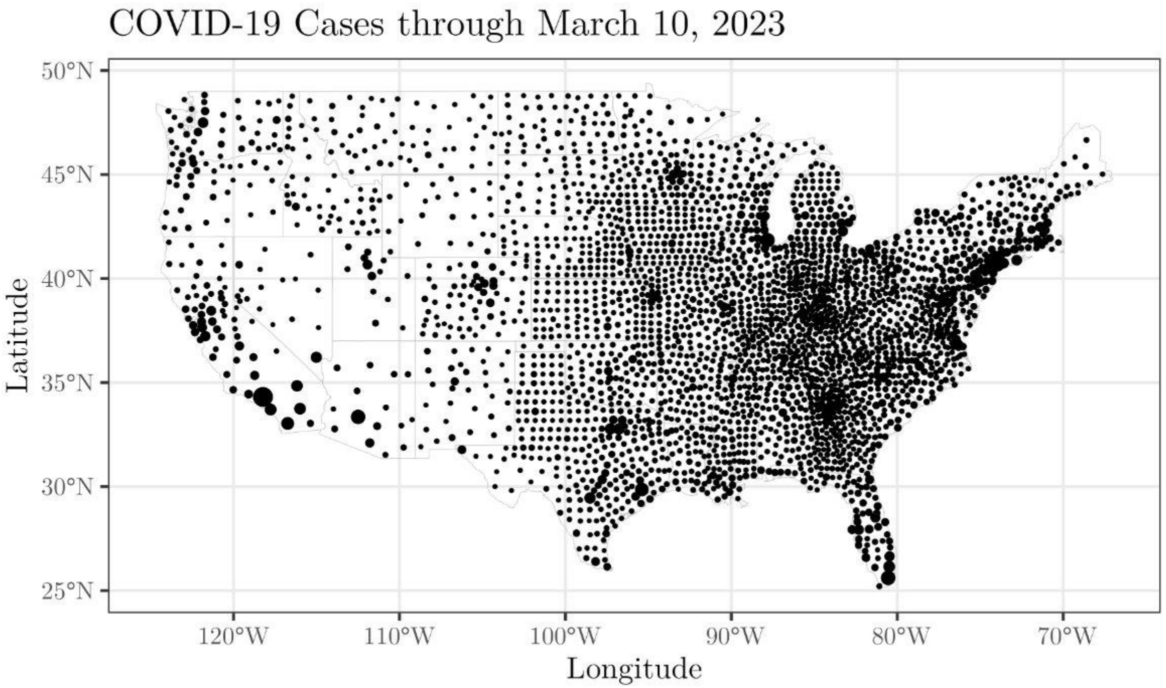

COVID-19 Cases through March 10, 2023
Latitude
50°N
45°N
40°N
35°N
30°N
25°N
120°W
110°W
100°W
90°W
80°W
70°W
Longitude

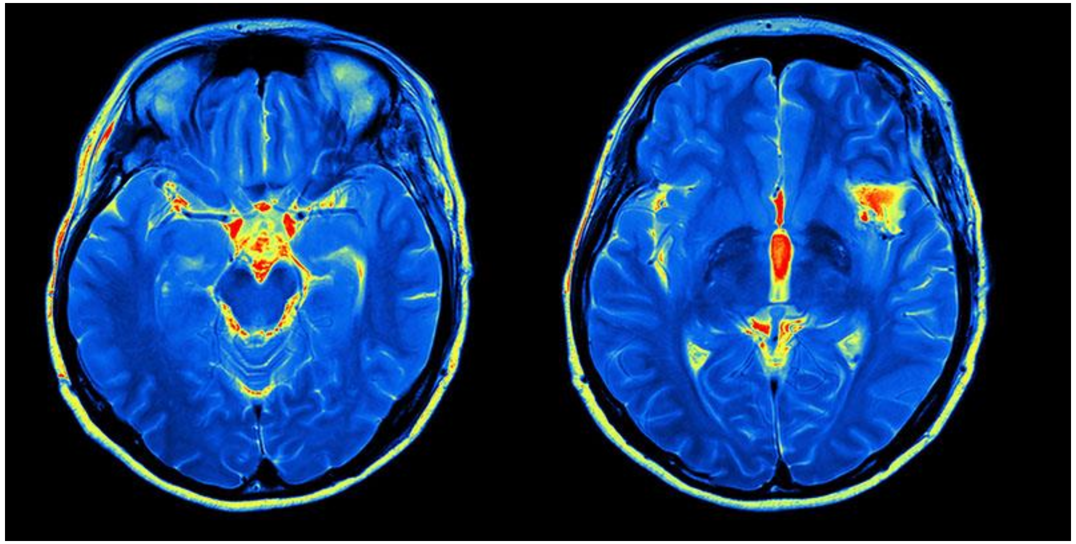

2022 National Weather Service
All Severe Weather Reports
HAIL
7179
6541 < 2 in
638 2+ in
WIND
17577
16236 < 65 kt
1341 65+ kt
TORNADO
1143
552 EF0/U
448 EF1
119 EF2
20 EF3
4 EF4
0 EF5
TOTAL 25899
NOAA
National Weather Service
Storm Prediction Center

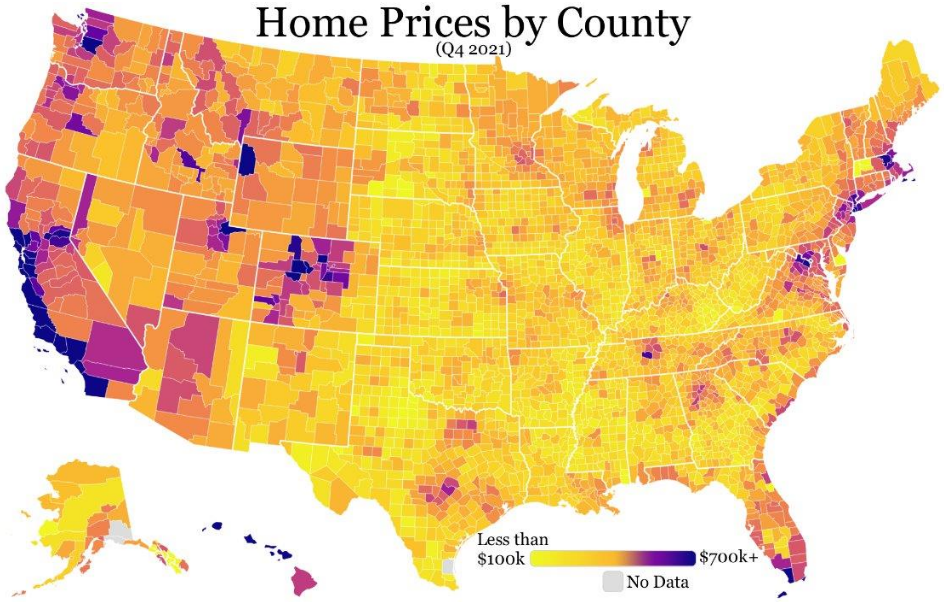

COVID-19 Cases through March 10, 2023
Latitude
50°N
45°N
40°N
35°N
30°N
25°N
120°W
110°W
100°W
90°W
80°W
70°W
Longitude




2022 National Weather Service
All Severe Weather Reports
HAIL
7179
6541 < 2 in
638 2+ in
WIND
17577
16236 < 65 kt
1341 65+ kt
TORNADO
1143
552 EF0/U
448 EF1
119 EF2
20 EF3
4 EF4
0 EF5
TOTAL 25899
NOAA
National Weather Service
Storm Prediction Center


COVID-19 Cases through March 10, 2023
Latitude
50°N
45°N
40°N
35°N
30°N
25°N
120°W
110°W
100°W
90°W
80°W
70°W
Longitude




Home Prices by County
(Q4 2021)
Less than
$100k
$700k+
No Data

Three types of spatial data

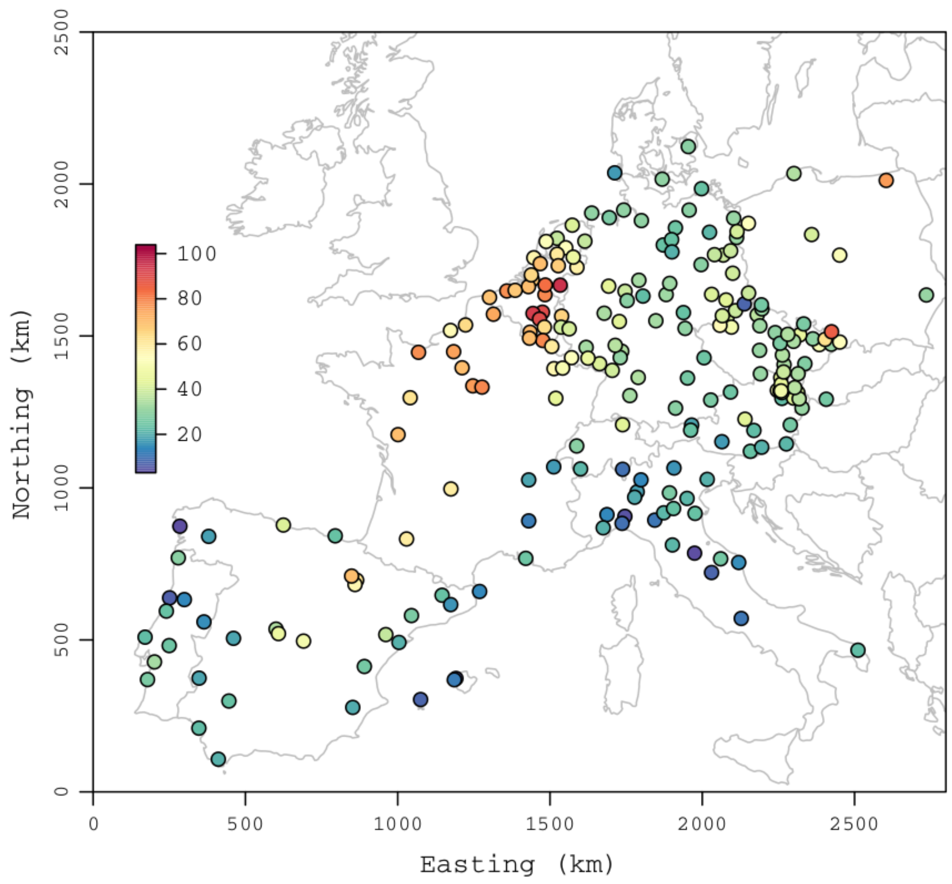
Point-referenced data

Areal data

Point pattern data

Point-referenced data

- Each observation is associated with a location (point)
- Data represents a sample from a continuous spatial domain
- Also referred to as **geocoded** or **geostatistical** data

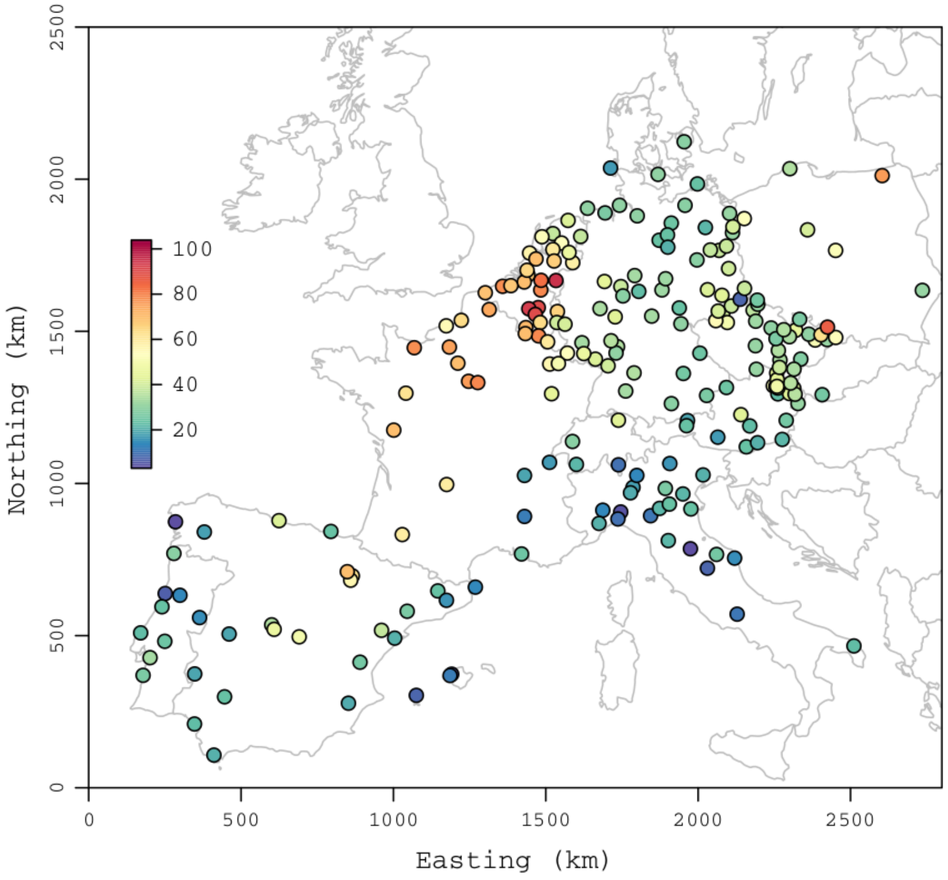


Pollutant levels in Europe in March 2009

Point-referenced data

- Each observation is associated with a location (point)
- Data represents a sample from a continuous spatial domain
- Also referred to as **geocoded** or **geostatistical** data

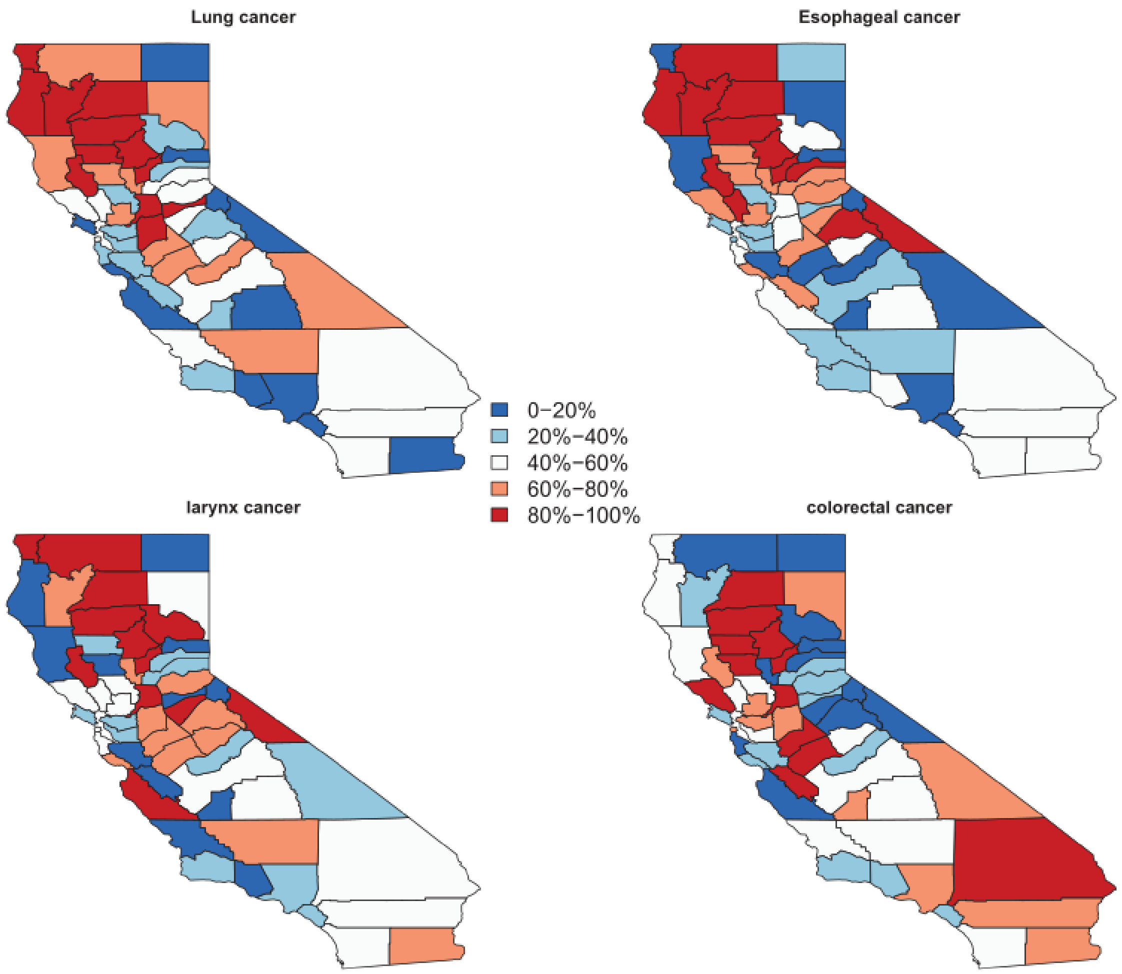
We will focus on methods for point-referenced data.



Pollutant levels in Europe in March 2009

Areal data

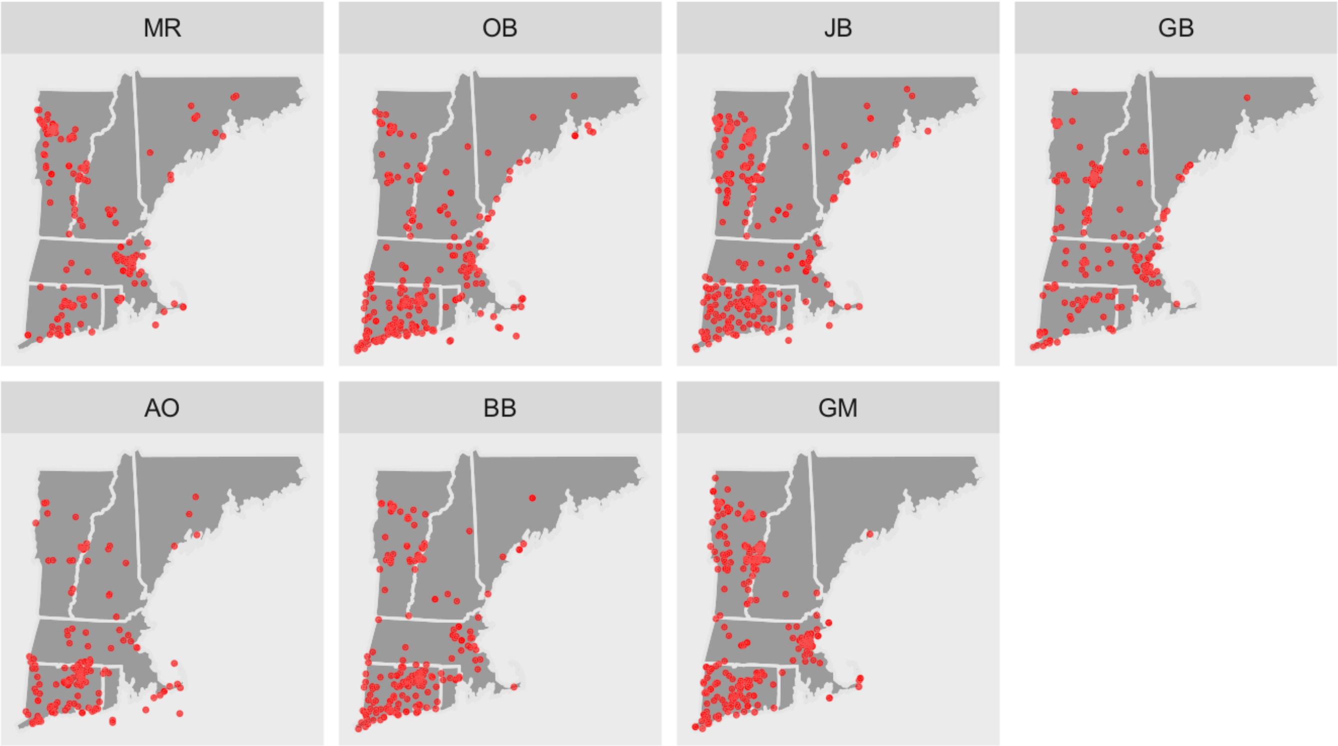
- Each observation is associated with an areal unit (of regular or irregular shape)
- Data represents a sample from a discrete spatial domain
- Also referred to as **lattice** data
- Example: disease incidence rates in California, USA



Gao, Banerjee, Ritz (2022)

Point pattern data

- Locations of the data points themselves are random
- Data set is often just the locations of the points
- Also called **point process data**
- If there is covariate information along with the points, called a **marked point process**
- Example: most presence-only data sets



Gelfand and Shirota (2022) *Ecological Monographs*

Point-referenced spatial modelling

- Modelling of point-referenced data collected at locations referenced by coordinates (e.g., lat-long, Easting-Northing).
- Also relevant for modelling of areal data (e.g., use the areal unit centroid), but not always ideal.
- We will model detection-nondetection and count data using point-referenced spatial regression models.

Why point-referenced modelling?

- Occupancy and abundance are inherently defined within some area (e.g., 250m radius circle, 1x1km grid cell).
- Why not use areal models (e.g., CAR, SAR)?

Why point-referenced modelling?

- Occupancy and abundance are inherently defined within some area (e.g., 250m radius circle, 1x1km grid cell).
- Why not use areal models (e.g., CAR, SAR)?
  1. Point-referenced models are more flexible, particularly for big data.

Why point-referenced modelling?

- Occupancy and abundance are inherently defined within some area (e.g., 250m radius circle, 1x1km grid cell).
- Why not use areal models (e.g., CAR, SAR)?
  1. Point-referenced models are more flexible, particularly for big data.
  2. Area sampled is not well-defined given animal movement (see Kéry and Royle, 2015).

Why point-referenced modelling?

- Occupancy and abundance are inherently defined within some area (e.g., 250m radius circle, 1x1km grid cell).
- Why not use areal models (e.g., CAR, SAR)?
  1. Point-referenced models are more flexible, particularly for big data.
  2. Area sampled is not well-defined given animal movement (see Kéry and Royle, 2015).
  3. Interest often lies in generating prediction surfaces of occupancy probability.

Why point-referenced modelling?

- Occupancy and abundance are inherently defined within some area (e.g., 250m radius circle, 1x1km grid cell).
- Why not use areal models (e.g., CAR, SAR)?
  1. Point-referenced models are more flexible, particularly for big data.
  2. Area sampled is not well-defined given animal movement (see Kéry and Royle, 2015).
  3. Interest often lies in generating prediction surfaces of occupancy probability.
  4. Area of interest is often large relative to the actual sampled area.

Introduction to Bayesian geostatistics

Let's start with a simple linear model

Introduction to Bayesian geostatistics

Let's start with a simple linear model

$$y(\boldsymbol{s}_j) \sim \mathrm{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$
$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \beta_2 \cdot x_2(\boldsymbol{s}_j) + \cdots + \beta_p \cdot x_p(\boldsymbol{s}_j)$$

Introduction to Bayesian geostatistics

Coordinates for site *j = 1, 2, …, J*

Let's start with a simple linear model

$$y(\boldsymbol{s}_j) \sim \mathrm{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$

$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \beta_2 \cdot x_2(\boldsymbol{s}_j) + \cdots + \beta_p \cdot x_p(\boldsymbol{s}_j)$$

Introduction to Bayesian geostatistics

Coordinates for site *j = 1, 2, …, J*

Let's start with a simple linear model

$$
\begin{aligned}
y(\boldsymbol{s}_j) &\sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2) \\
\mu(\boldsymbol{s}_j) &= \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \beta_2 \cdot x_2(\boldsymbol{s}_j) + \cdots + \beta_p \cdot x_p(\boldsymbol{s}_j) \\
\beta_k &\sim \text{Normal}(0, 100) \\
\tau^2 &\sim \text{Inverse-gamma}(0.01, 0.01)
\end{aligned}
$$

Introduction to Bayesian geostatistics

Coordinates for site *j = 1, 2, …, J*

Let's start with a simple linear model

$$y(\boldsymbol{s}_j) \sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$

$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \beta_2 \cdot x_2(\boldsymbol{s}_j) + \cdots + \beta_p \cdot x_p(\boldsymbol{s}_j)$$

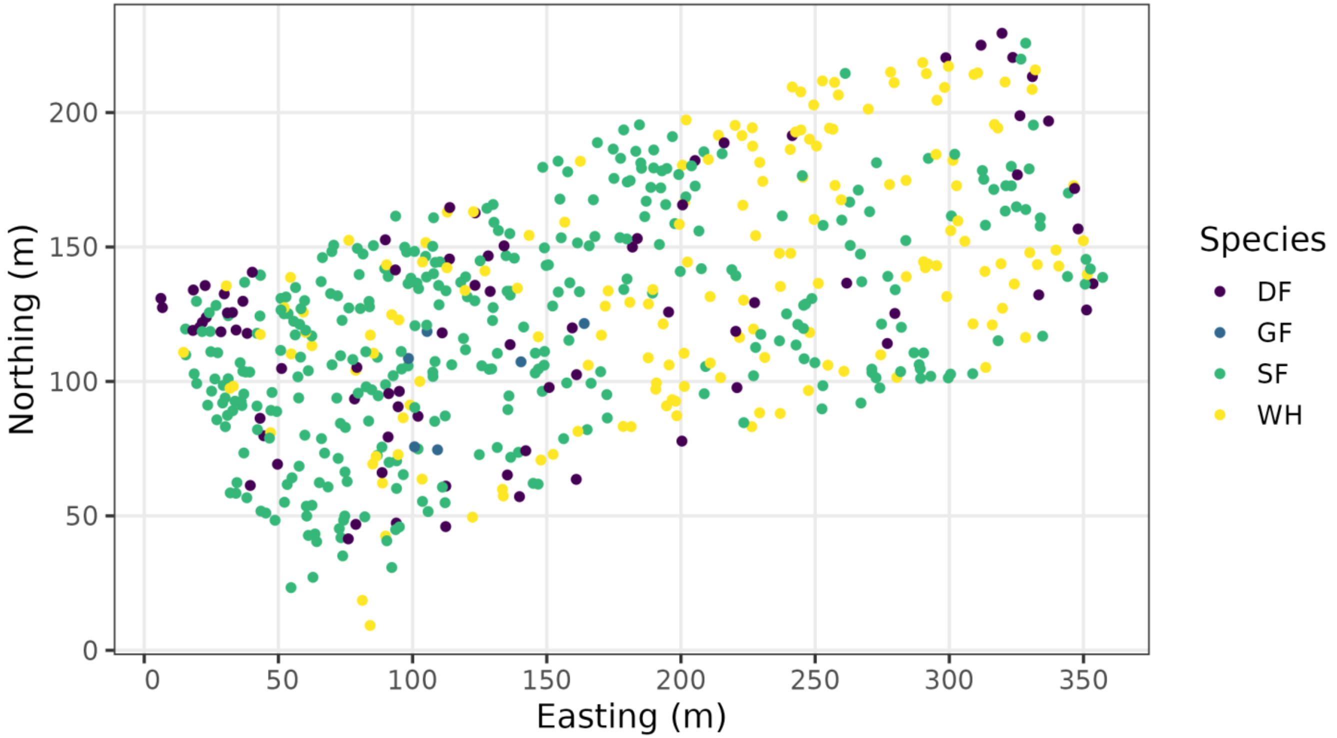
$$\beta_k \sim \text{Normal}(0, 100)$$

$$\tau^2 \sim \text{Inverse-gamma}(0.01, 0.01)$$

- Recall our basic assumption that the residuals in a linear model are IID (independent and identically distributed).
- Alternatively, can say that $y(\boldsymbol{s}_j)$ are conditionally independent given the covariates included in the model.

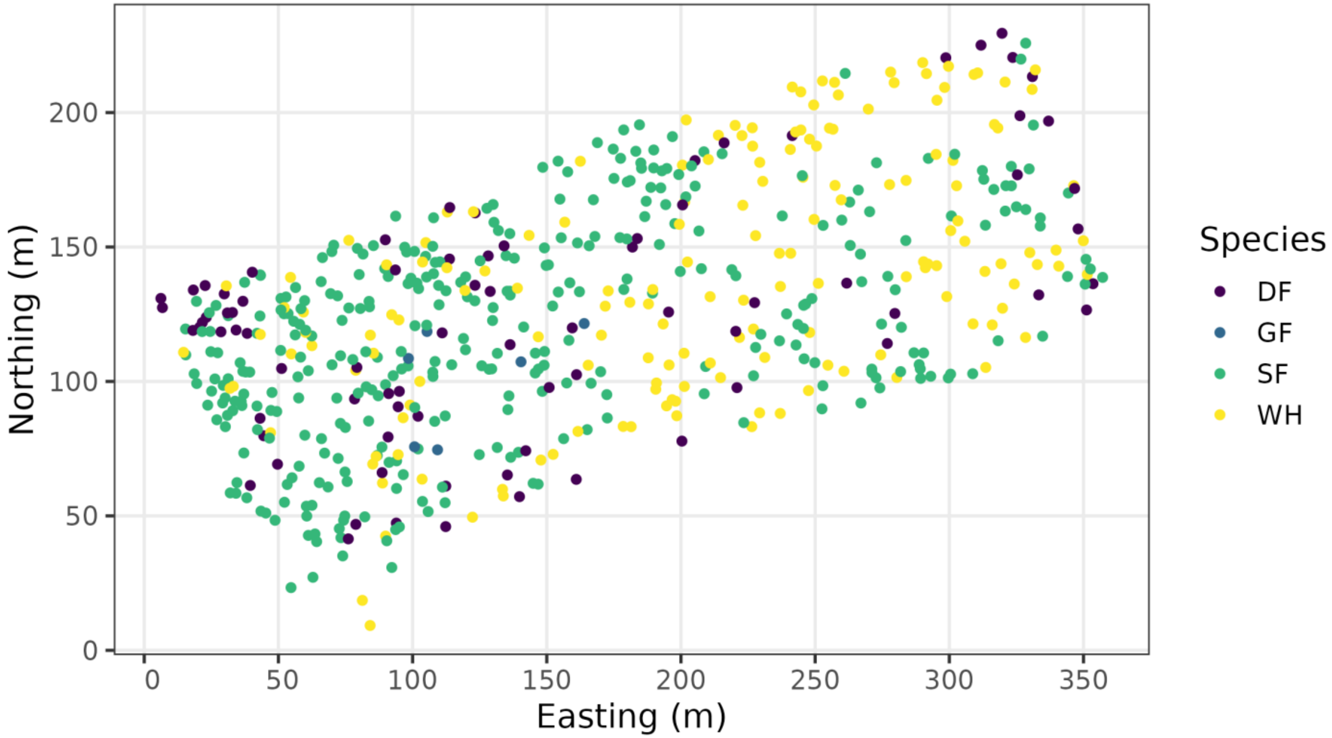
Example: Western Experimental Forest (WEF) data

- 600 trees measured across a 10 ha forest stand in Oregon, USA
- Objective: Predict diameter at breast height (DBH) of all trees across the stand
- Covariate: Tree species (categorical variable)



Example: Western Experimental Forest (WEF) data

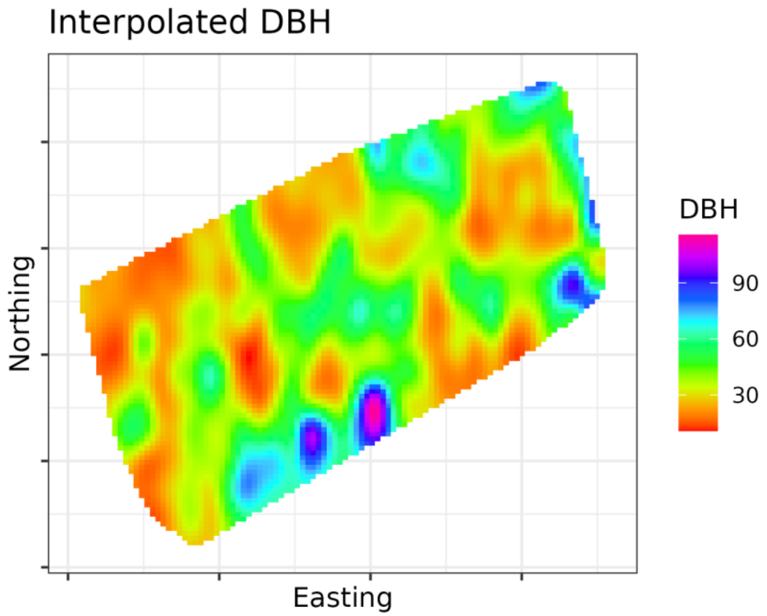
- 600 trees measured across a 10 ha forest stand in Oregon, USA
- Objective: Predict diameter at breast height (DBH) of all trees across the stand
- Covariate: Tree species (categorical variable)

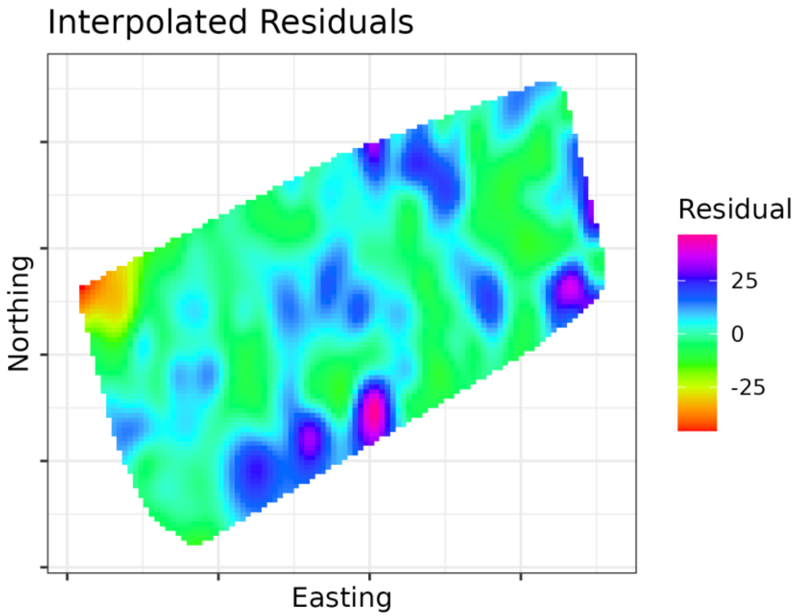


`02-spatial-linear-model-wef.R`

Classical geostatistics: residual plots

- Surface plots of the residuals help to identify any spatial patterns left unexplained by the covariates
- Local spatial patterns are still evident in residual plots. Basic linear model does not seem sufficient





Classical geostatistics: semivariogram

- More formal EDA approach for identifying spatial patterns

Classical geostatistics: semivariogram

- More formal EDA approach for identifying spatial patterns
- Overall concept: if the sqaured difference between any two data points gets larger as the distance between points increases, this implies a spatial correlation.

Classical geostatistics: semivariogram

- More formal EDA approach for identifying spatial patterns
- Overall concept: if the squared difference between any two data points gets larger as the distance between points increases, this implies a spatial correlation.

$$(y(\boldsymbol{s}_j) - y(\boldsymbol{s}_k))^2$$

Classical geostatistics: semivariogram

- More formal EDA approach for identifying spatial patterns
- Overall concept: if the squared difference between any two data points gets larger as the distance between points increases, this implies a spatial correlation.

$$(y(\boldsymbol{s}_j) - y(\boldsymbol{s}_k))^2$$

If the value on the left gets larger as the distance between $\mathbf{s}_j$ and $\mathbf{s}_k$ increases, there is spatial correlation.

Classical geostatistics: empirical semivariogram

1. Define a set of intervals based on distances between points

Classical geostatistics: empirical semivariogram

1. Define a set of intervals based on distances between points
2. Assign each pair of points to the associated distance interval

Classical geostatistics: empirical semivariogram

1. Define a set of intervals based on distances between points
2. Assign each pair of points to the associated distance interval
3. Calculate the average squared difference between points within each interval

Classical geostatistics: empirical semivariogram

1. Define a set of intervals based on distances between points
2. Assign each pair of points to the associated distance interval
3. Calculate the average squared difference between points within each interval
4. Repeat for all intervals to give a set of values across distances that you can use for plotting

Classical geostatistics: empirical semivariogram

Midpoint of the r^{th} distance bin

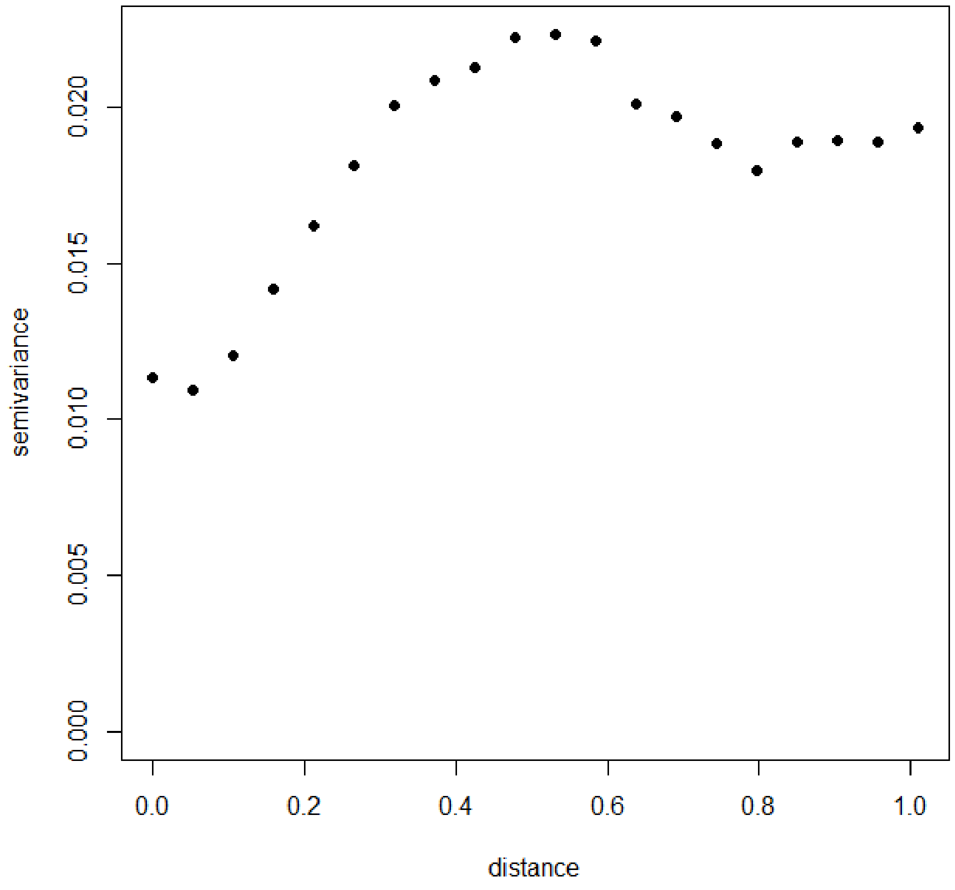
$$\gamma(t_r) = \frac{1}{2\mid N(t_r)\mid} \sum_{\boldsymbol{s}_j, \boldsymbol{s}_k \in N(t_r)} (y(\boldsymbol{s}_j) - y(\boldsymbol{s}_k))^2$$

Number of point pairs in the r^{th} distance bin

The "in" symbol. We are summing over all pairs of points in $N(t_r)$

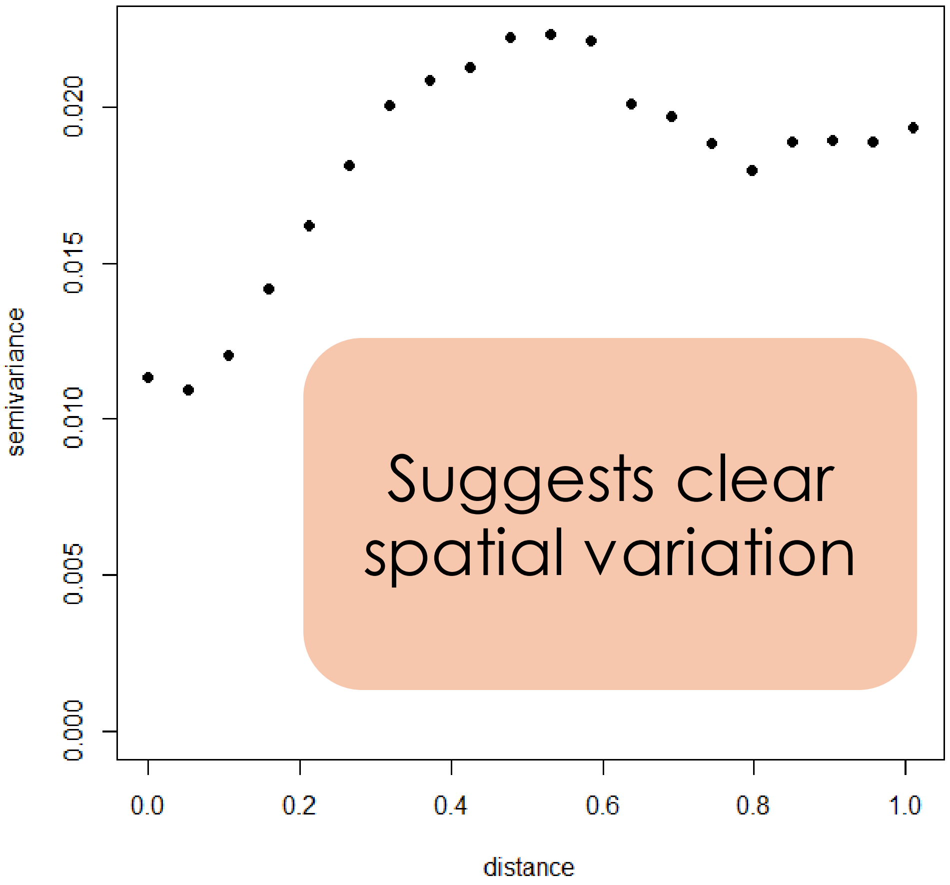
Classical geostatistics: empirical semivariogram

- For spatial data, the empirical variogram is expected to roughly increase with distance
- A flat semivariogram would suggest little spatial variation

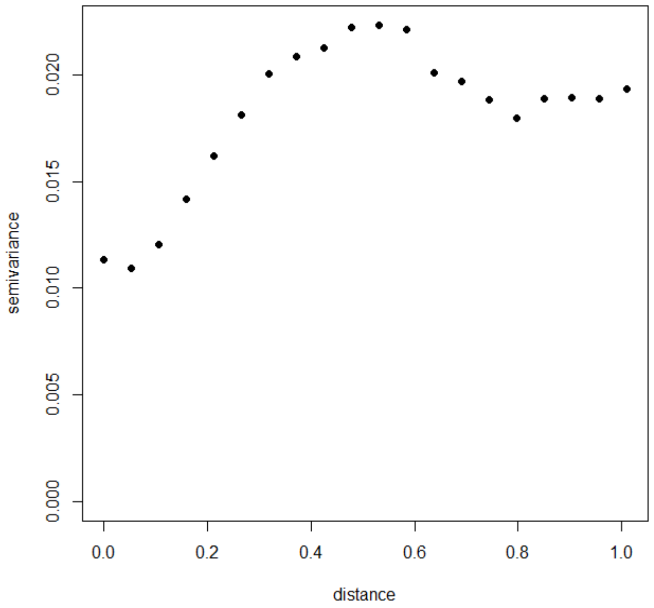


Classical geostatistics: empirical semivariogram

- For spatial data, the empirical variogram is expected to roughly increase with distance
- A flat semivariogram would suggest little spatial variation

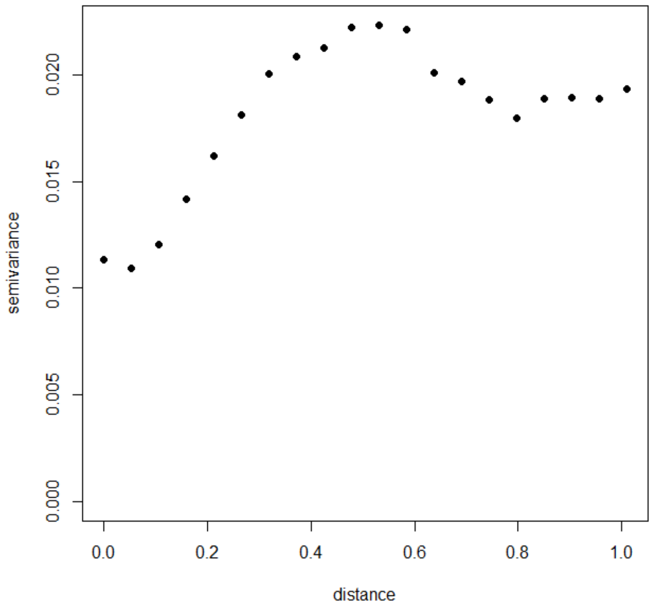


Classical geostatistics: empirical semivariogram

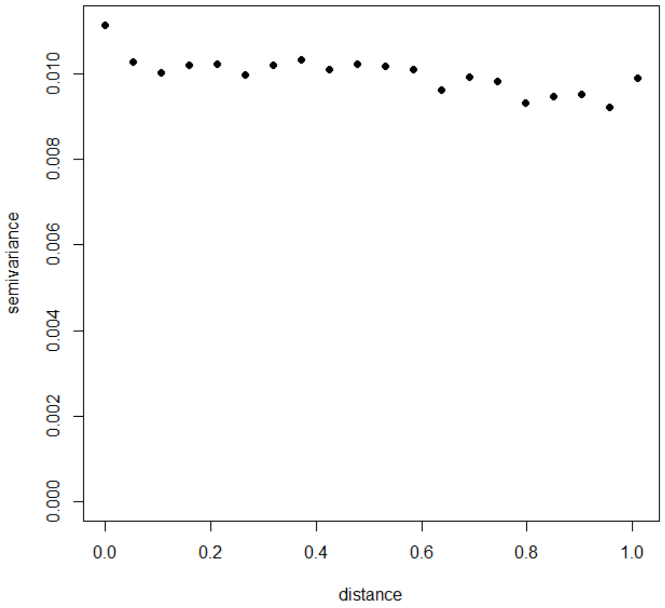


Response variable

Classical geostatistics: empirical semivariogram

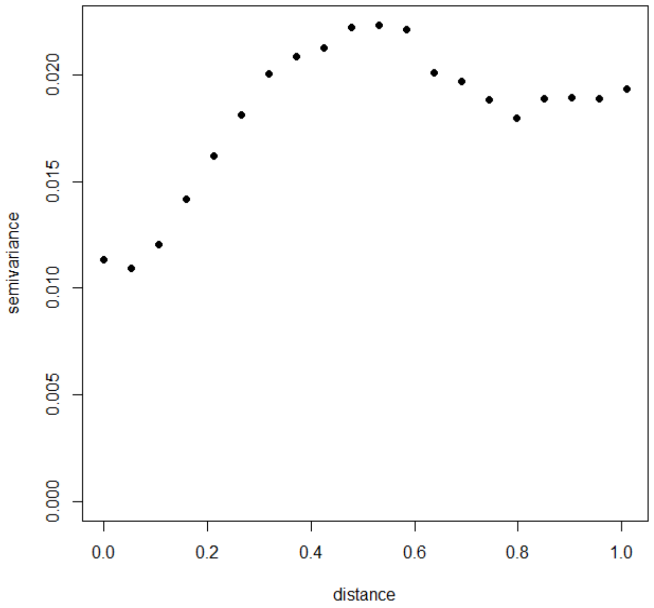


Response variable

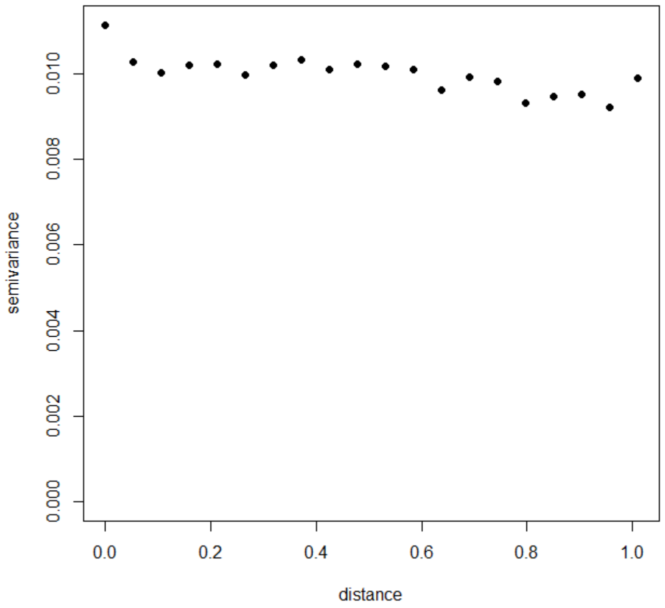


Residuals

Classical geostatistics: empirical semivariogram



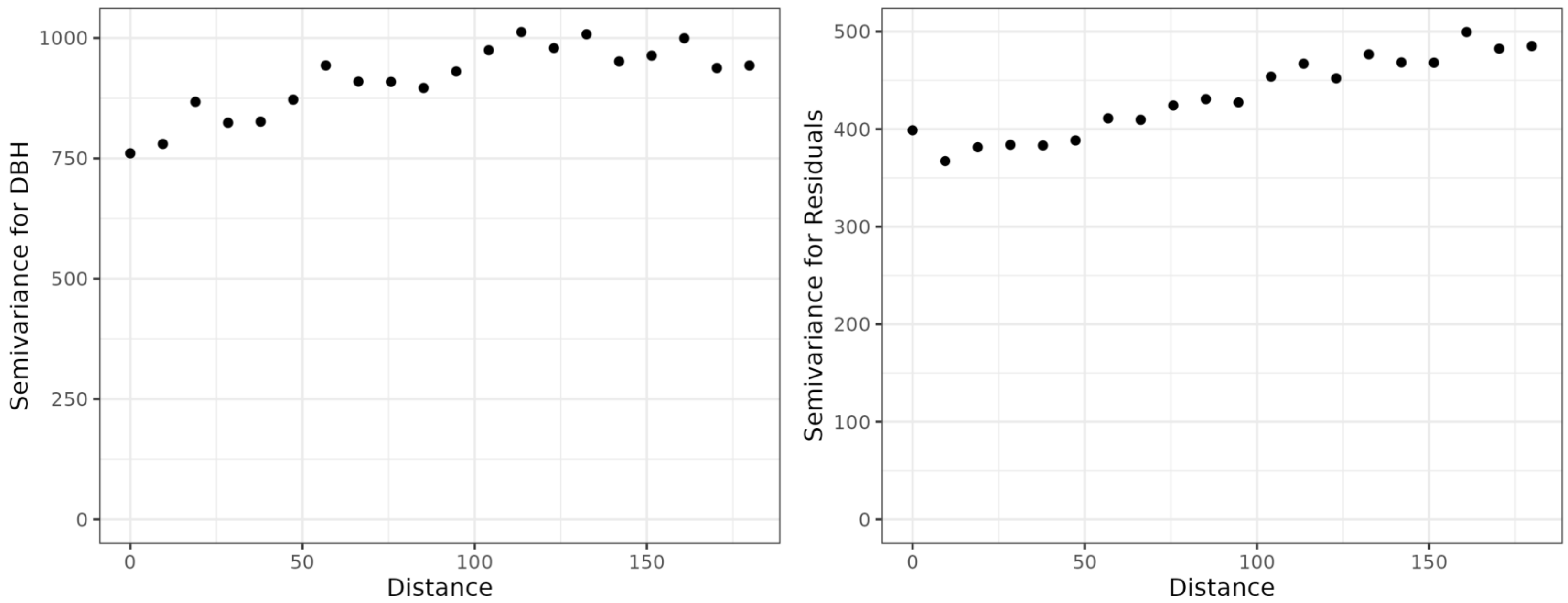
Response variable



Residuals

Residuals show little pattern. A simple linear model is adequate.

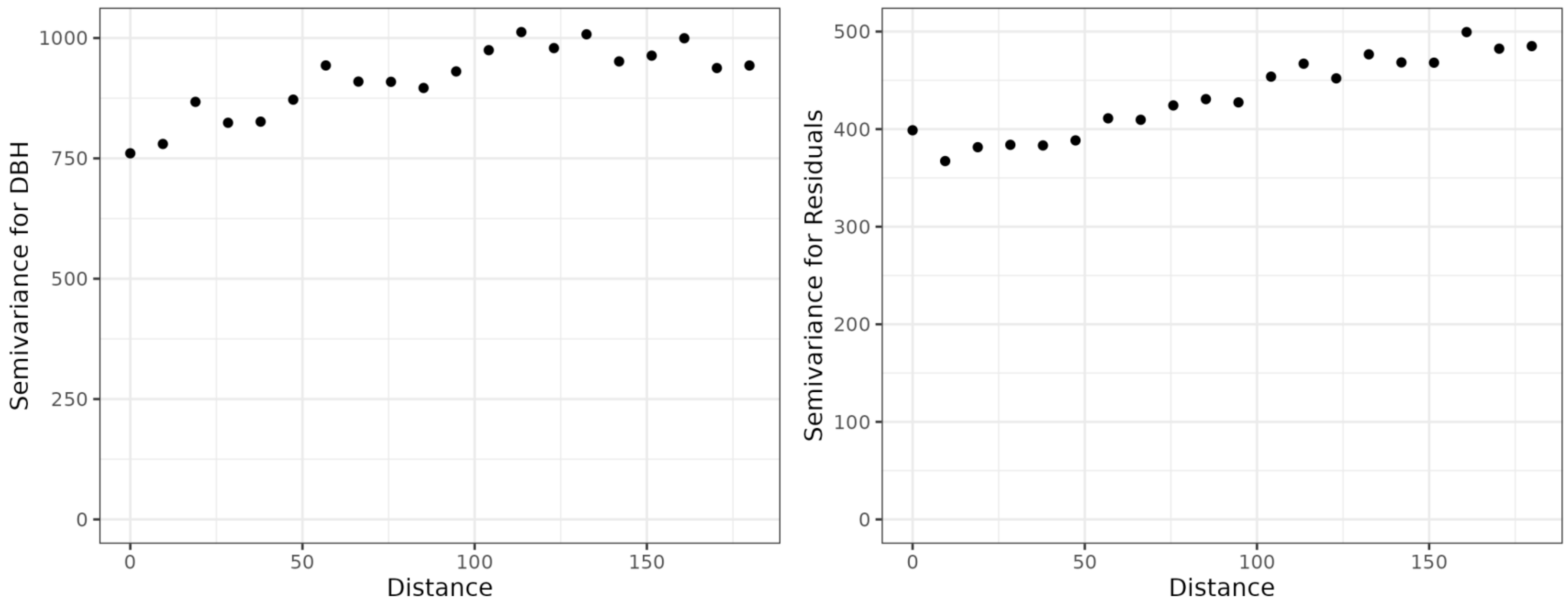
WEF Example



Semivariogram of the residuals shows unexplained spatial variation. We need a spatial model!

WEF Example

02-spatial-linear-model-wef.R



Semivariogram of the residuals shows unexplained spatial variation. We need a spatial model!

What do we mean by "spatial model"?

- A model that explicitly seeks to leverage information from the *locations* of the data points, in addition to using covariates.

What do we mean by "spatial model"?

- A model that explicitly seeks to leverage information from the *locations* of the data points, in addition to using covariates.
- When purely covariate-based models do not suffice, we should leverage the spatial information of the points.

What do we mean by "spatial model"?

- A model that explicitly seeks to leverage information from the *locations* of the data points, in addition to using covariates.
- When purely covariate-based models do not suffice, we should leverage the spatial information of the points.
- Allow us to account for **Residual Spatial Autocorrelation** (i.e., spatial correlation in the response that is not explained by the covariates).

Spatial linear model

Spatial random effect

$$y(\boldsymbol{s}_j) \sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$
$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \text{w}(\boldsymbol{s}_j)$$

Spatial linear model

Spatial random effect

$$y(\boldsymbol{s}_j) \sim \mathrm{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$
$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j)$$

- How to choose the form of w(**s**)?
- We want to choose w(**s**) such that we can easily predict and make smooth maps.
- We will choose w(**s**) to be a *surface*.

Gaussian Processes (GPs)

Gaussian Processes (GPs)

- "Gold standard" for modelling spatial data.

Gaussian Processes (GPs)

- "Gold standard" for modelling spatial data.
- A bit of stats jargon: A GP is a *stochastic process* that describes all the infinitely many locations within some area of interest.

Gaussian Processes (GPs)

- "Gold standard" for modelling spatial data.
- A bit of stats jargon: A GP is a *stochastic process* that describes all the infinitely many locations within some area of interest.
- Why is it good for us:
    - Completely characterized by a mean and covariance function.
    - Any given set of spatial locations follows a multivariate normal distributions.
    - Multivariate normal distributions are nice.

Gaussian Processes (GPs)

- "Gold standard" for modelling spatial data.
- A bit of stats jargon: A GP is a *stochastic process* that describes all the infinitely many locations within some area of interest.
- Why is it good for us:
  - Completely characterized by a mean and covariance function.
  - Any given set of spatial locations follows a multivariate normal distributions.
  - Multivariate normal distributions are nice.
- We almost always work with GPs that have a mean of 0.

Gaussian Processes (GPs)

- "Gold standard" for modelling spatial data.
- A bit of stats jargon: A GP is a *stochastic process* that describes all the infinitely many locations within some area of interest.
- Why is it good for us:
  - Completely characterized by a mean and covariance function.
  - Any given set of spatial locations follows a multivariate normal distributions.
  - Multivariate normal distributions are nice.
- We almost always work with GPs that have a mean of 0.
- So, what does all this mean for our random effects?

Spatial linear model

$$
\begin{aligned}
y(\boldsymbol{s}_j) &\sim \mathrm{Normal}(\mu(\boldsymbol{s}_j), \tau^2) \\
\mu(\boldsymbol{s}_j) &= \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j) \\
\mathbf{w} &\sim \text{Multivariate Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))
\end{aligned}
$$

Spatial linear model

Spatial covariance matrix

$$y(\boldsymbol{s}_j) \sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$

$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j)$$

$$\mathbf{w} \sim \text{Multivariate Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))$$

Spatial linear model

Spatial covariance matrix

$$y(\boldsymbol{s}_j) \sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$

$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \text{w}(\boldsymbol{s}_j)$$

$$\mathbf{w} \sim \text{Multivariate Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))$$

- Spatial random effects are completely determined by a covariance matrix

Spatial linear model

Spatial covariance matrix

$$y(\boldsymbol{s}_j) \sim \mathrm{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$

$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j)$$

$$\mathbf{w} \sim \mathrm{Multivariate\ Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))$$

- Spatial random effects are completely determined by a covariance matrix
- Covariance between two sites is determined by:

Spatial linear model

Spatial covariance matrix

$$y(\boldsymbol{s}_j) \sim \mathrm{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$

$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j)$$

$$\mathbf{w} \sim \mathrm{Multivariate\ Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))$$

- Spatial random effects are completely determined by a covariance matrix
- Covariance between two sites is determined by:
    - Distance between the sites

Spatial linear model

Spatial covariance matrix

$$y(\boldsymbol{s}_j) \sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$

$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j)$$

$$\mathbf{w} \sim \text{Multivariate Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))$$

- Spatial random effects are completely determined by a covariance matrix
- Covariance between two sites is determined by:
  - Distance between the sites
  - A covariance function

Spatial covariance functions

- Tons of options.

Spatial covariance functions

- Tons of options.
- Isotropic vs. Anisotropic functions.
  - **Isotropic**: covariance between two locations only depends on distance.
  - **Anisotropic**: covariance between two locations depends on distance and direction.

Spatial covariance functions

- Tons of options.
- Isotropic vs. Anisotropic functions.
  - **Isotropic**: covariance between two locations only depends on distance.
  - **Anisotropic**: covariance between two locations depends on distance and direction.
- We will focus on isotropic functions because of their simplicity and interpretability.

Spatial covariance functions

- Tons of options.
- Isotropic vs. Anisotropic functions.
    - **Isotropic**: covariance between two locations only depends on distance.
    - **Anisotropic**: covariance between two locations depends on distance and direction.
- We will focus on isotropic functions because of their simplicity and interpretability..
- Four covariance functions supported in `spOccupancy` and `spAbundance`: exponential, spherical, Gaussian, Matérn

Spatial covariance functions

- Tons of options.
- Isotropic vs. Anisotropic functions.
  - o **Isotropic**: covariance between two locations only depends on distance.
  - o **Anisotropic**: covariance between two locations depends on distance and direction.
- We will focus on isotropic functions because of their simplicity and interpretability.
- Four covariance functions supported in `spOccupancy` and `spAbundance`: exponential, spherical, Gaussian, Matérn.
- We will focus primarily on using the exponential covariance function.

Exponential covariance function

- Covariance between site A and site B using exponential covariance function:

$$\boldsymbol{C}(d_{A,B}, \sigma^2, \phi)) = \sigma^2 \exp(-\phi d_{A,B})$$

Intuition on exponential spatial covariance

$$\boldsymbol{C}(d_{A,B}, \sigma^2, \phi)) = \sigma^2 \exp(-\phi d_{A,B})$$

$d_{A,B}$ Euclidean (linear) distance between site A and B.

Intuition on exponential spatial covariance

$$\boldsymbol{C}(d_{A,B}, \sigma^2, \phi)) = \sigma^2 \exp(-\phi d_{A,B})$$

$d_{A,B}$ Euclidean (linear) distance between site A and B.

σ^2 Spatial variance. Controls magnitude of random effects.

Intuition on exponential spatial covariance

$$\boldsymbol{C}(d_{A,B}, \sigma^2, \phi)) = \sigma^2 \exp(-\phi d_{A,B})$$

$d_{A,B}$ Euclidean (linear) distance between site A and B.

σ^2 Spatial variance. Controls magnitude of random effects.

ϕ Spatial decay. Controls how quickly the correlation between sites decays across space.

Intuition on exponential spatial covariance

$$\boldsymbol{C}(d_{A,B}, \sigma^2, \phi)) = \sigma^2 \exp(-\phi d_{A,B})$$

$d_{A,B}$ Euclidean (linear) distance between site A and B.

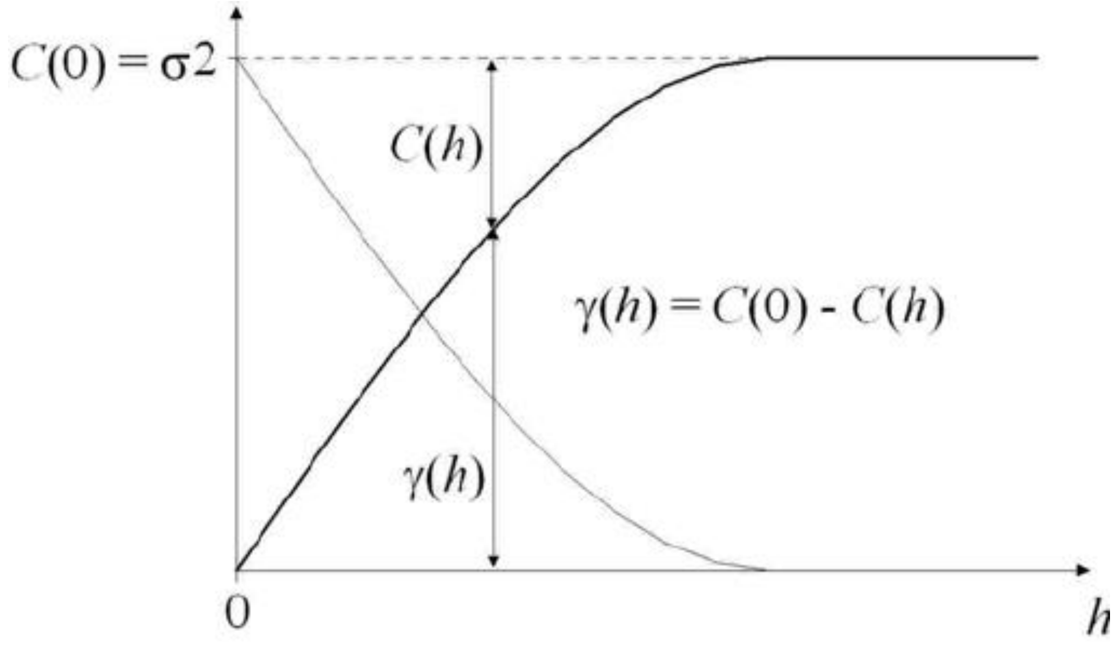
σ^2 Spatial variance. Controls magnitude of random effects.

ϕ Spatial decay. Controls how quickly the correlation between sites decays across space.

$\frac{3}{\phi}$ "Effective spatial range" when using an exponential covariance function. This is the distance at which the spatial correlation between two sites is essentially negligible (0.05)

Covariance functions and semivariograms

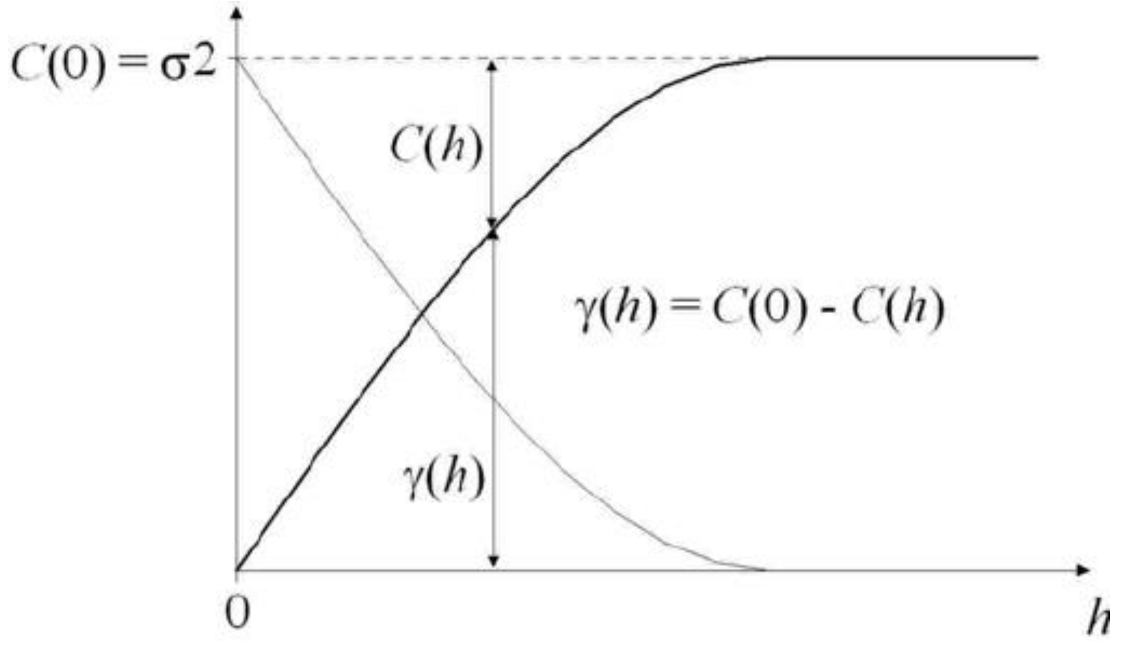
Every covariance function is associated with a semivariogram. They are inversely related.



Covariance functions and semivariograms

Every covariance function is associated with a semivariogram. They are inversely related.

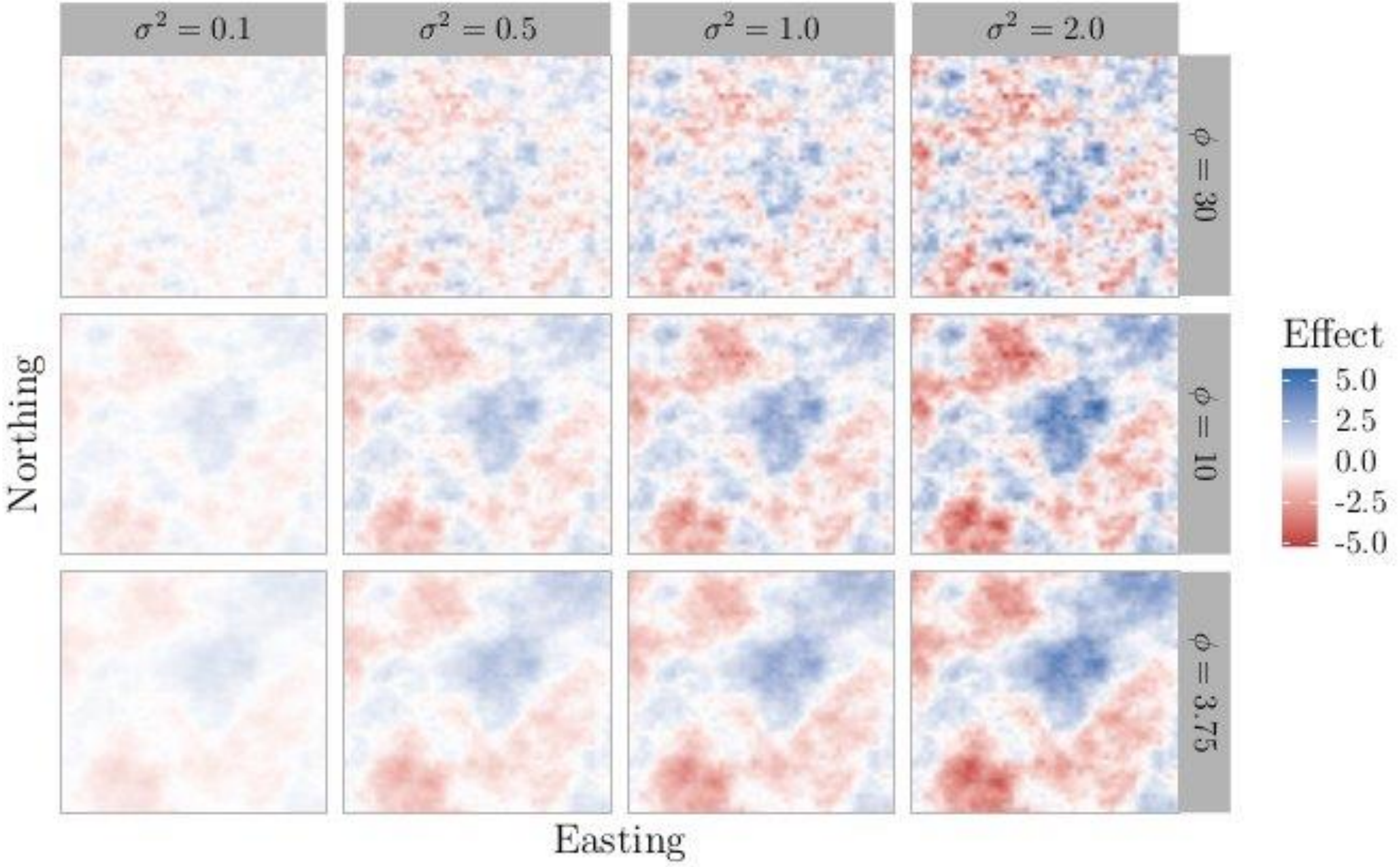
Covariance functions are intuitive: things closer together have higher correlation, things farther away have lower correlation.



Spatial Gaussian processes

$$\boldsymbol{C}(d_{A,B}, \sigma^2, \phi)) = \sigma^2 \exp(-\phi d_{A,B})$$

Maps of the resulting spatial random effects (**w**) under different parameter values



How to predict at new locations?

How to predict at new locations?

- By using GPs, estimates of the random effects at new locations also follow a multivariate normal distribution!!

How to predict at new locations?

- By using GPs, estimates of the random effects at new locations also follow a multivariate normal distribution!!
- In other words, GPs allow us to easily predict at new locations and generate nice maps.

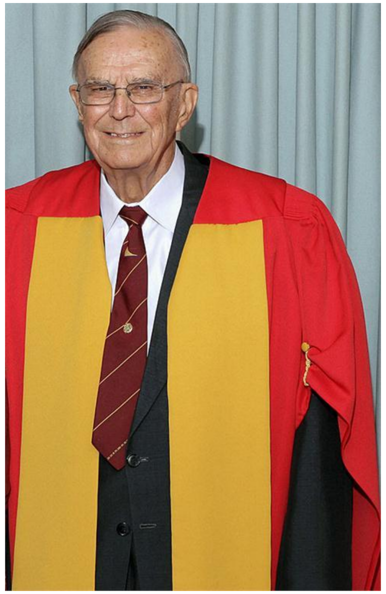
How to predict at new locations?

- By using GPs, estimates of the random effects at new locations also follow a multivariate normal distribution!!
- In other words, GPs allow us to easily predict at new locations and generate nice maps.
- Predicting using spatial models (particularly frequentist models) is often called **kriging.**

How to predict at new locations?

- By using GPs, estimates of the random effects at new locations also follow a multivariate normal distribution!!
- In other words, GPs allow us to easily predict at new locations and generate nice maps.
- Predicting using spatial models (particularly frequentist models) is often called **kriging.**

Danie
Krige



How to estimate spatial models?

$$
\begin{aligned}
y(\boldsymbol{s}_j) &\sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2) \\
\mu(\boldsymbol{s}_j) &= \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j) \\
\mathbf{w} &\sim \text{Multivariate Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))
\end{aligned}
$$

How to estimate spatial models?

$$y(\boldsymbol{s}_j) \sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$
$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j)$$
$$\mathbf{w} \sim \text{Multivariate Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))$$

- As per usual, we can go a Bayesian route or a frequentist route.

How to estimate spatial models?

$$y(\boldsymbol{s}_j) \sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$
$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j)$$
$$\mathbf{w} \sim \text{Multivariate Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))$$

- As per usual, we can go a Bayesian route or a frequentist route.
- The Bayesian approach is highly preferred (and I promise it's not just me saying that!).

How to estimate spatial models?

$$y(\boldsymbol{s}_j) \sim \text{Normal}(\mu(\boldsymbol{s}_j), \tau^2)$$
$$\mu(\boldsymbol{s}_j) = \beta_0 + \beta_1 \cdot x_1(\boldsymbol{s}_j) + \mathrm{w}(\boldsymbol{s}_j)$$
$$\mathbf{w} \sim \text{Multivariate Normal}(\mathbf{0}, \boldsymbol{C}(d, \sigma^2, \phi))$$

- As per usual, we can go a Bayesian route or a frequentist route.
- The Bayesian approach is highly preferred (and I promise it's not just me saying that!).
- Frequentist approaches (ML, REML) do not adequately account for uncertainty in the estimates of the spatial variance and spatial decay parameters. Leads to overly precise predictions.

Bayesian spatial linear model

- Assign prior distributions to all parameters and fit the model using MCMC.
- The two spatial parameters (ϕ and σ^2) are notoriously difficult to estimate separately from each other (Zhang, 2004 *JASA*).
- Priors become important. In `spOccupancy` and `spAbundance`, we use the following:
  - ϕ: A mildly informative uniform prior
  - σ^2: A vague inverse-gamma prior

Default priors on spatial parameters

$$\phi \sim \text{Uniform}(\frac{3}{d_{\text{max}}}, \frac{3}{d_{\text{min}}})$$

Maximum inter-site distance

Minimum inter-site distance

- Default priors allow the effective spatial range to be anywhere from the maximum inter-site distance in the data set to the minimum inter-site distance.
- Fairly uninformative: allows for broad-scale and fine-scale spatial autocorrelation.
- Can be made more informative by changing d_{max} and d_{min} to other values.

Default priors on spatial parameters

$$\sigma^2 \sim \underbrace{\text{inverse-gamma}(2,}_{\text{Shape}} \underbrace{1)}_{\text{Scale}}$$

- Sets the prior mean of the spatial variance to 1 (moderate amount of spatial variation) with an infinite variance.
- When the shape = 2, the prior variance is infinite and the prior mean is equal to the scale.

Exercise: Western Experimental Forest

`02-spatial-linear-model-wef.R`

Pros to fitting spatial models

- More accurate predictions at new locations (i.e., more accurate maps)
- More accurate uncertainty estimates
- Visualizing spatial random effects can provide insight on underlying drivers
- Help to generate new hypotheses

Cons to fitting spatial models

- Substantially slower (we will address this shortly).
- Convergence can be difficult to achieve without informative priors on spatial parameters.
- Generally requires more data (more spatial locations) than non-spatial models.
- Spatial confounding

Spatial confounding

- Covariates in your spatial model may be highly correlated with the spatial random effects, which can lead to difficulties when interpreting the covariate effect.
- Same concept as checking correlation between covariates between fitting a model (except one covariate is the spatial random effect).
- Some spatial models try to explicitly eliminate this problem (i.e., restricted spatial regression), but the benefits of this approach are highly debated.

Literature on spatial confounding

General

Adding Spatially-Correlated Errors Can Mess Up the Fixed Effect You Love

James S. Hodges & Brian J. Reich

Pages 325-334 | Received 01 Mar 2010, Published online: 01 Jan 2012

Listen

Statistical Practice

On Deconfounding Spatial Confounding in Linear Models

Dale L. Zimmerman & Jay M. Ver Hoef

Pages 159-167 | Received 26 Feb 2021, Accepted 12 Jun 2021, Published online: 26 Jul 2021

ECOGRAPHY

Research article

Spatial confounding in Bayesian species distribution modeling

Jussi Mäkinen, Elina Numminen, Pekka Niittynen, Miska Luoto and Jarno Vanhatalo

Things to consider when focusing on covariate effects in spatial models

1. Consider the range of spatial autocorrelation in the covariate. Covariates with long-range spatial autocorrelation are more likely to show spatial confounding than those with short-range autocorrelation (Mäkinen et al. 2022).

Things to consider when focusing on covariate effects in spatial models

1. Consider the range of spatial autocorrelation in the covariate. Covariates with long-range spatial autocorrelation are more likely to show spatial confounding than those with short-range autocorrelation (Mäkinen et al. 2022).
2. Plot the estimates of the spatial random effect and compare with the covariates. Are they highly correlated?

Things to consider when focusing on covariate effects in spatial models

1. Consider the range of spatial autocorrelation in the covariate. Covariates with long-range spatial autocorrelation are more likely to show spatial confounding than those with short-range autocorrelation (Mäkinen et al. 2022).
2. Plot the estimates of the spatial random effect and compare with the covariates. Are they highly correlated?
3. Compare estimated effects in spatial model to those in non-spatial model. Are there drastic differences?

Things to consider when focusing on covariate effects in spatial models

1. Consider the range of spatial autocorrelation in the covariate. Covariates with long-range spatial autocorrelation are more likely to show spatial confounding than those with short-range autocorrelation (Mäkinen et al. 2022).
2. Plot the estimates of the spatial random effect and compare with the covariates. Are they highly correlated?
3. Compare estimated effects in spatial model to those in non-spatial model. Are there drastic differences?
4. Use more informative priors to reduce the spatial confounding. Read Mäkinen et al. 2022 for specific details.