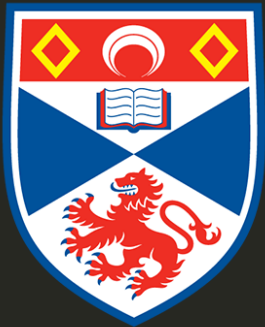


# Lecture 1: distance sampling & density surface models

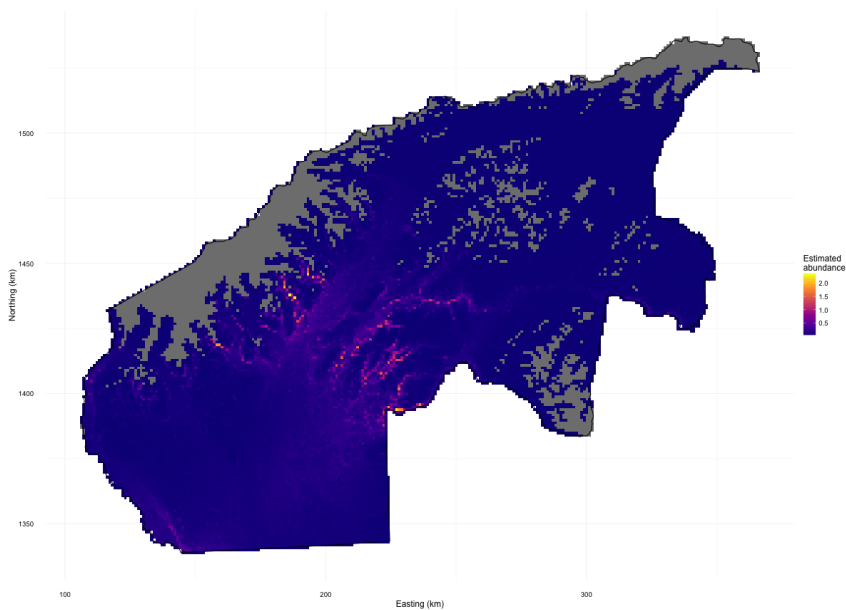


University of  
St Andrews

Why model abundance spatially?

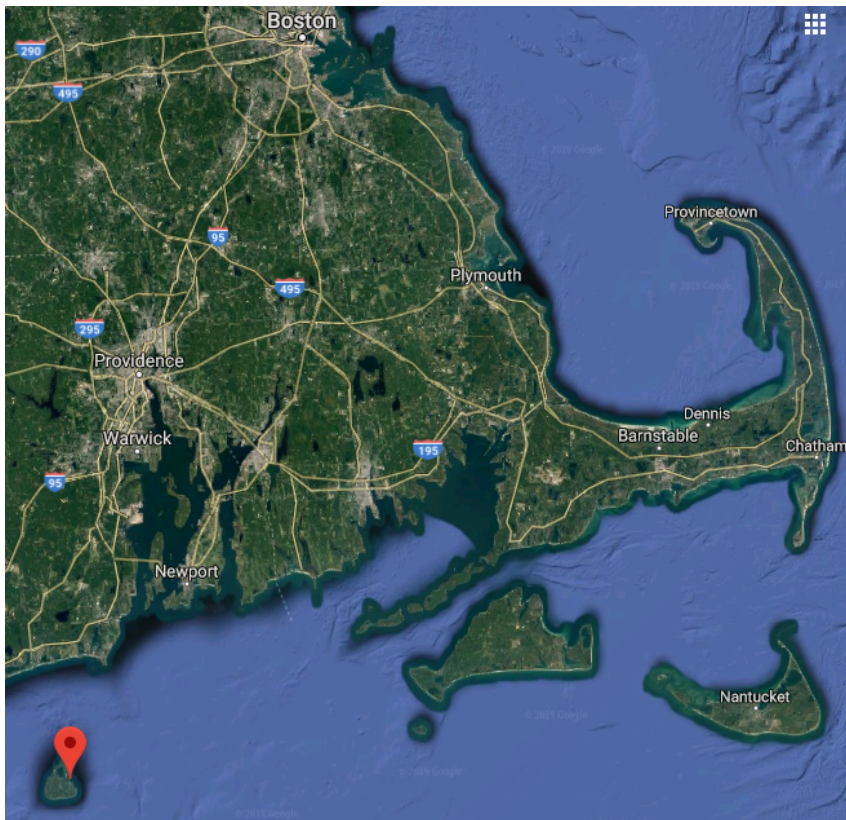
# Maps

- Black bears in Alaska
- Heterogeneous spatial distribution

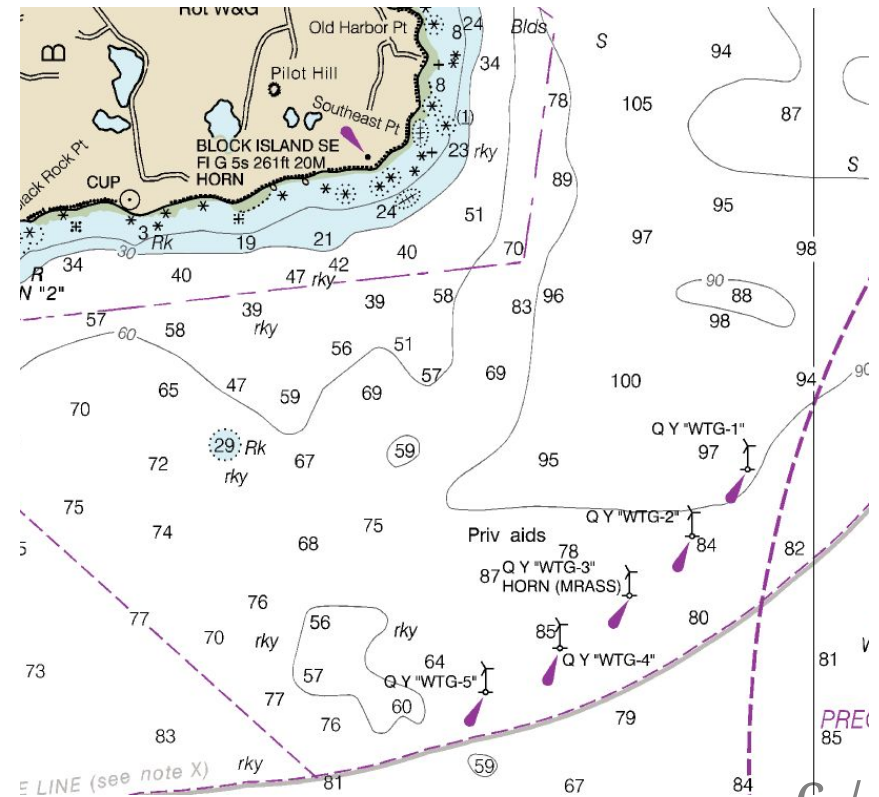
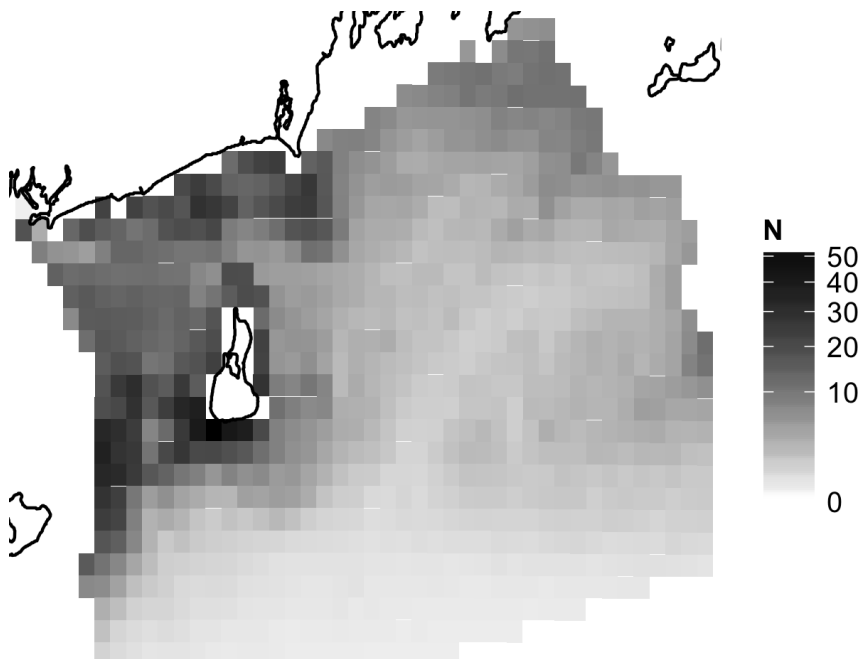




# Spatial decision making

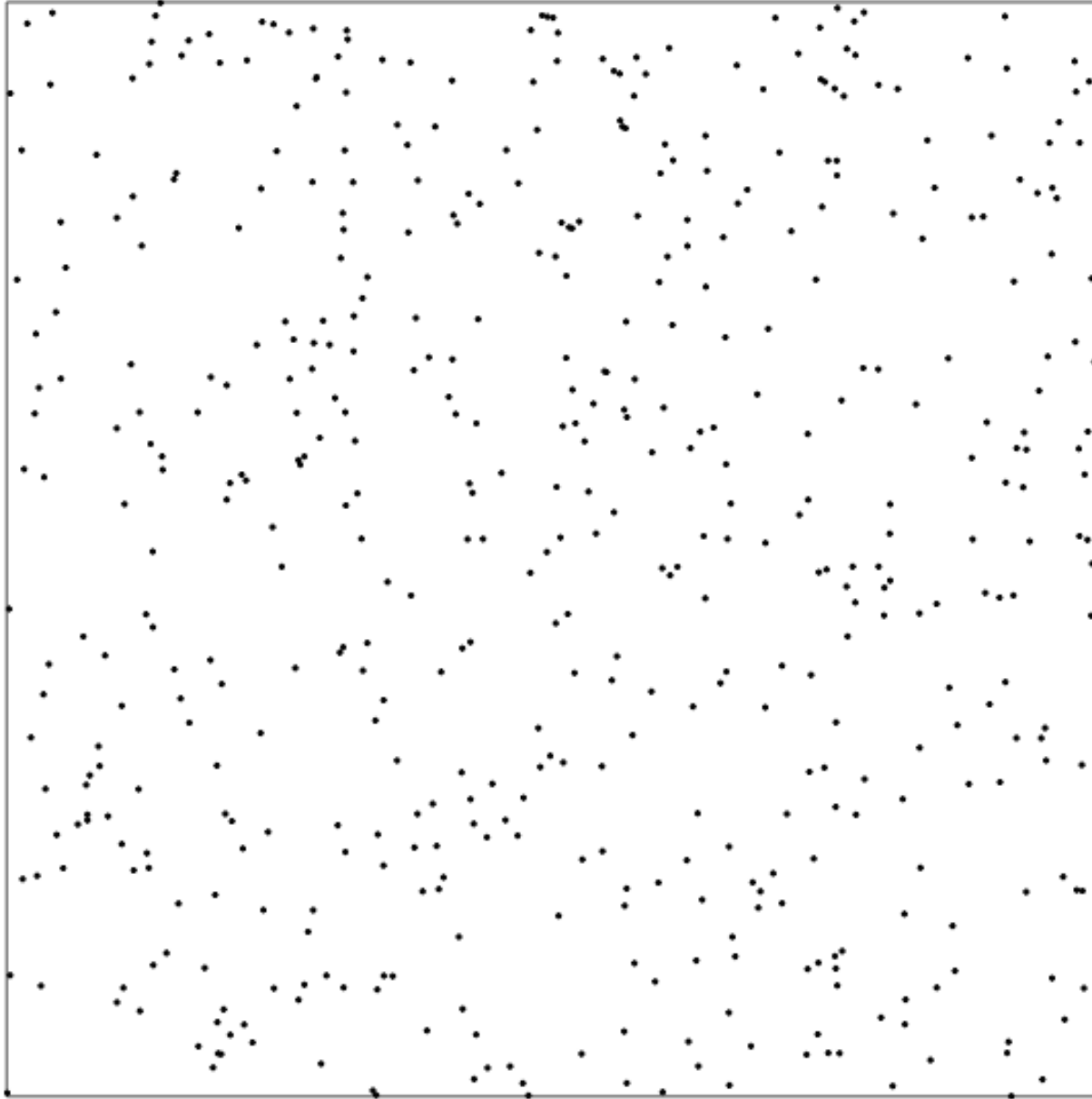


- Block Island, Rhode Island
- First offshore wind in the USA
- Spatial impact assessment

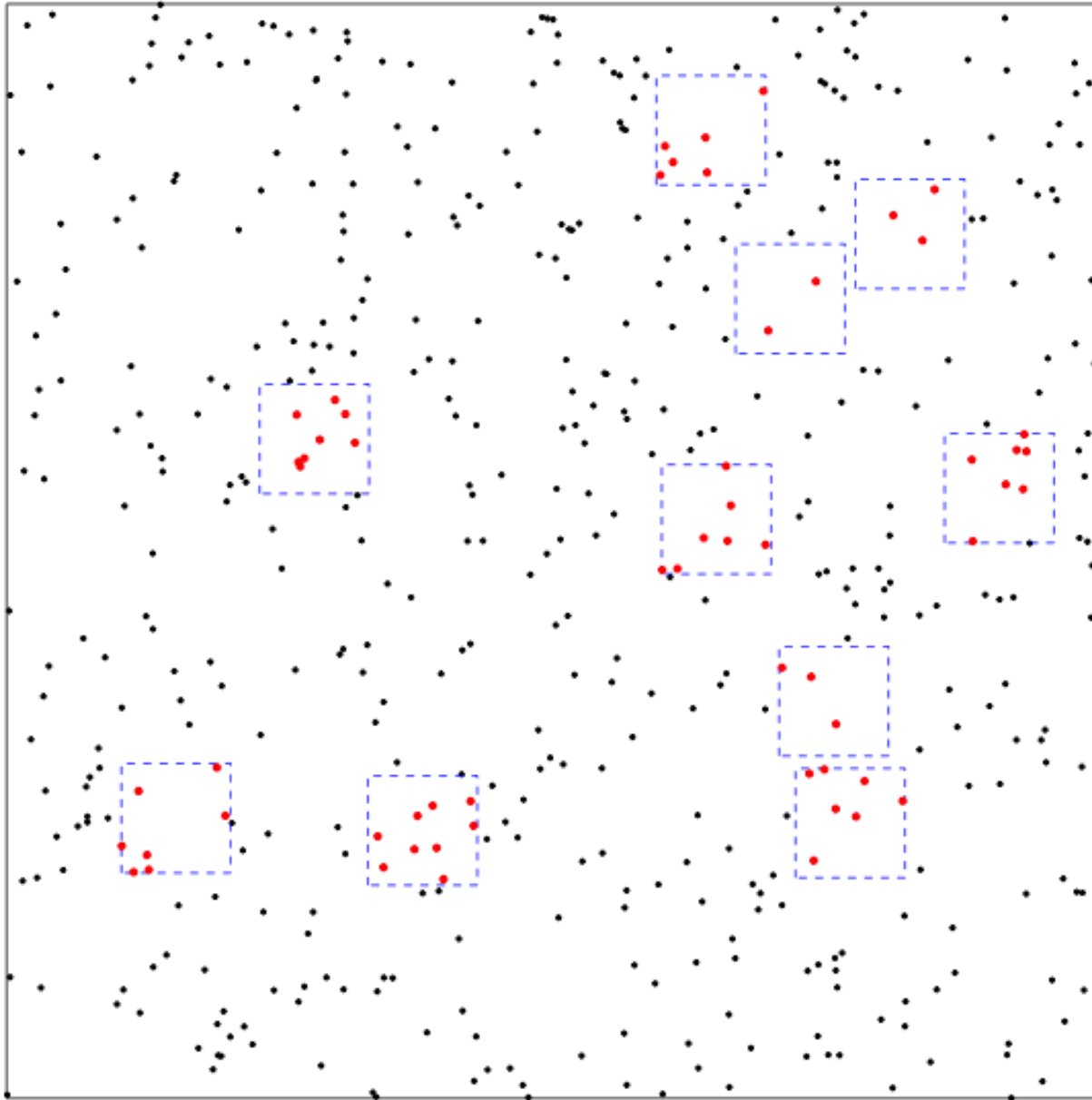


# A quick tour of distance sampling

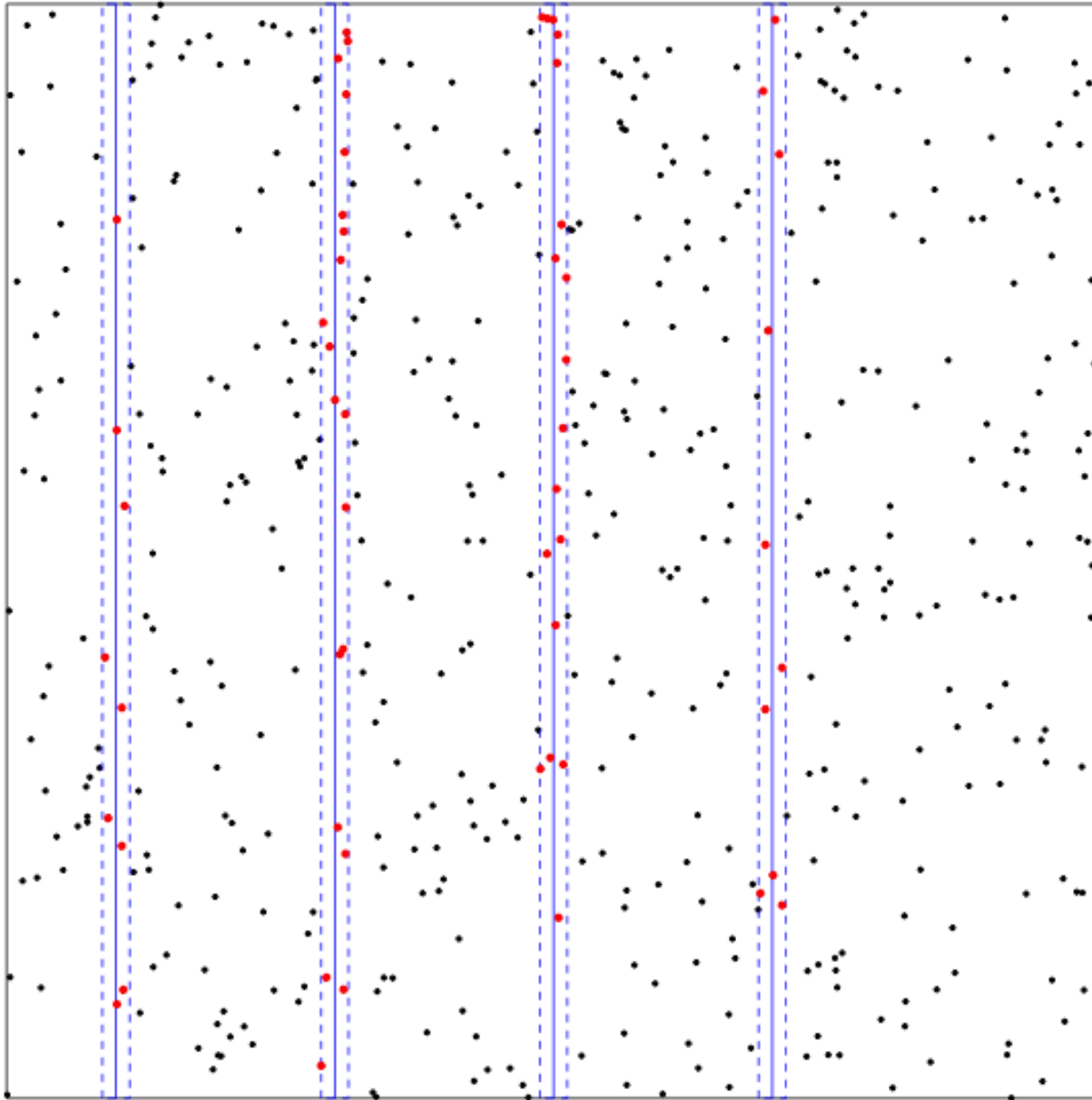
# How many animals are there? (500!)



# Plot sampling



# Strip transect



# Detectability matters!

- We've assumed certain detection so far
- This rarely happens in the field
- Distance to the **object** is important
- Detectability should decrease with increasing distance

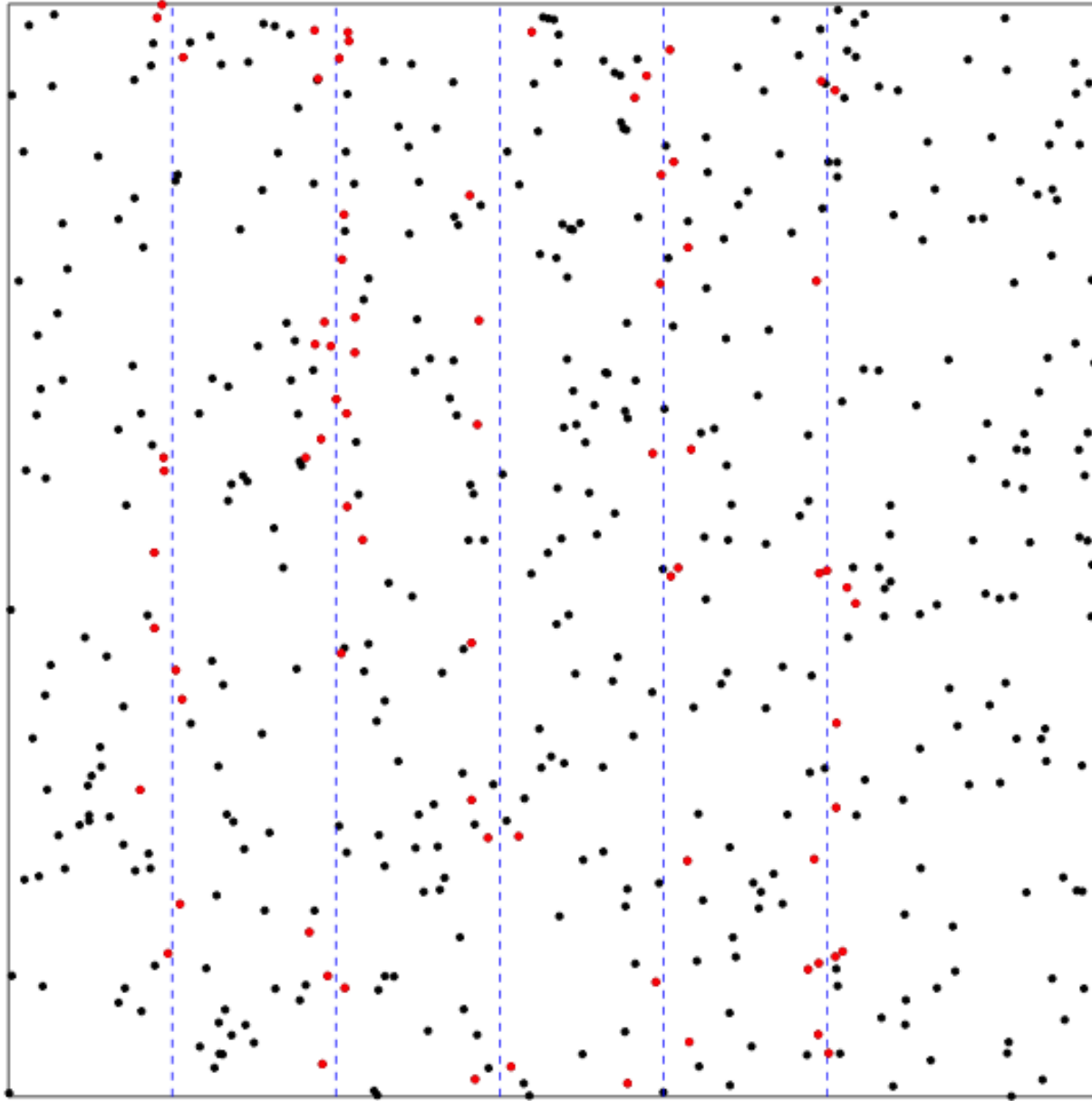
# Distance and detectability



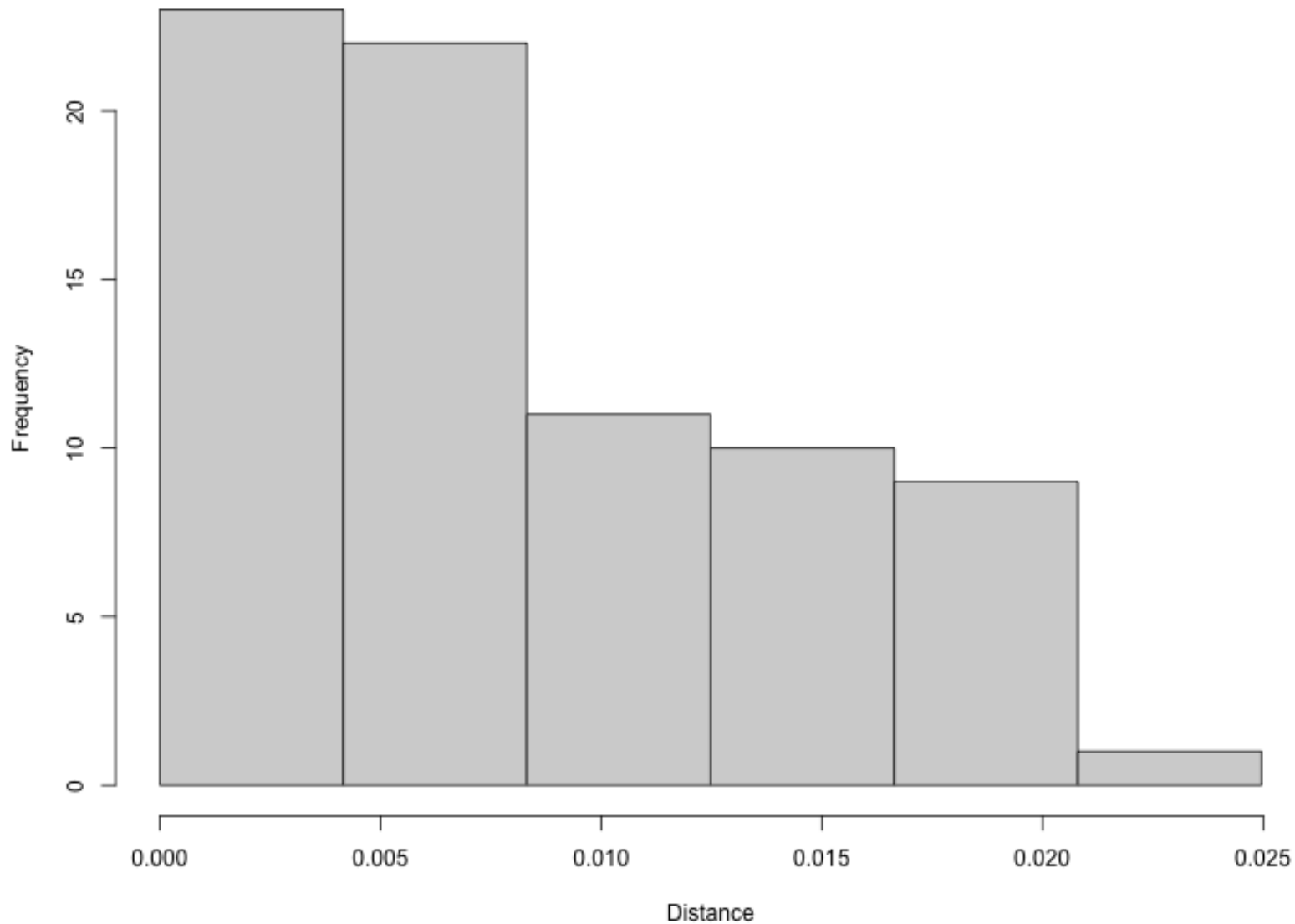
Credit **Scott and Mary Flanders**



# Line transect

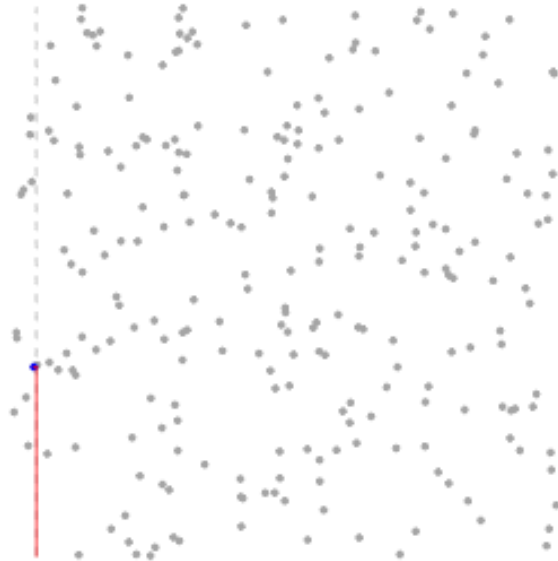


# Line transects - distances

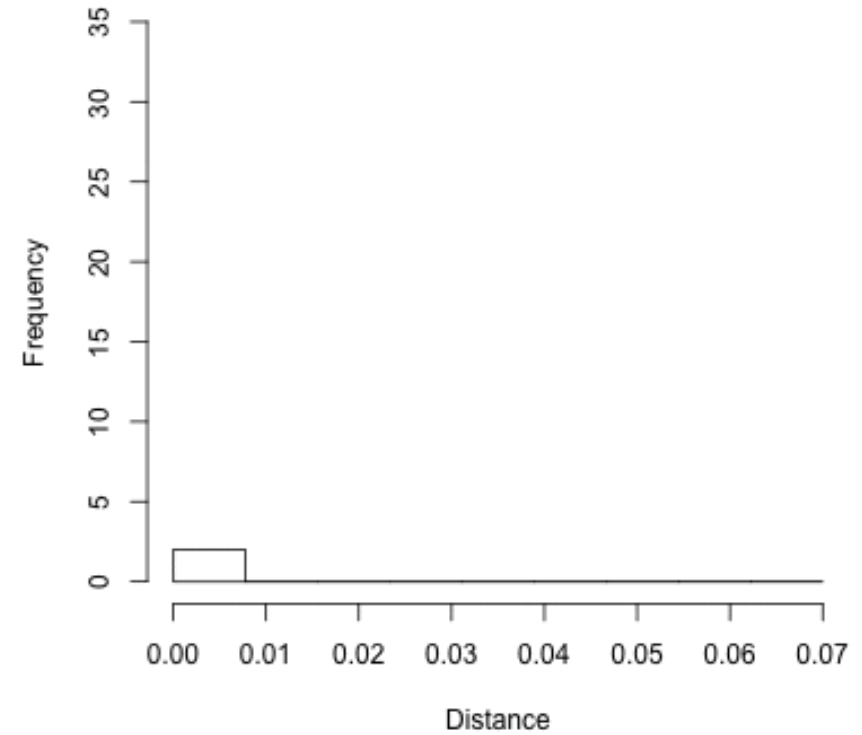


# Distance sampling animation

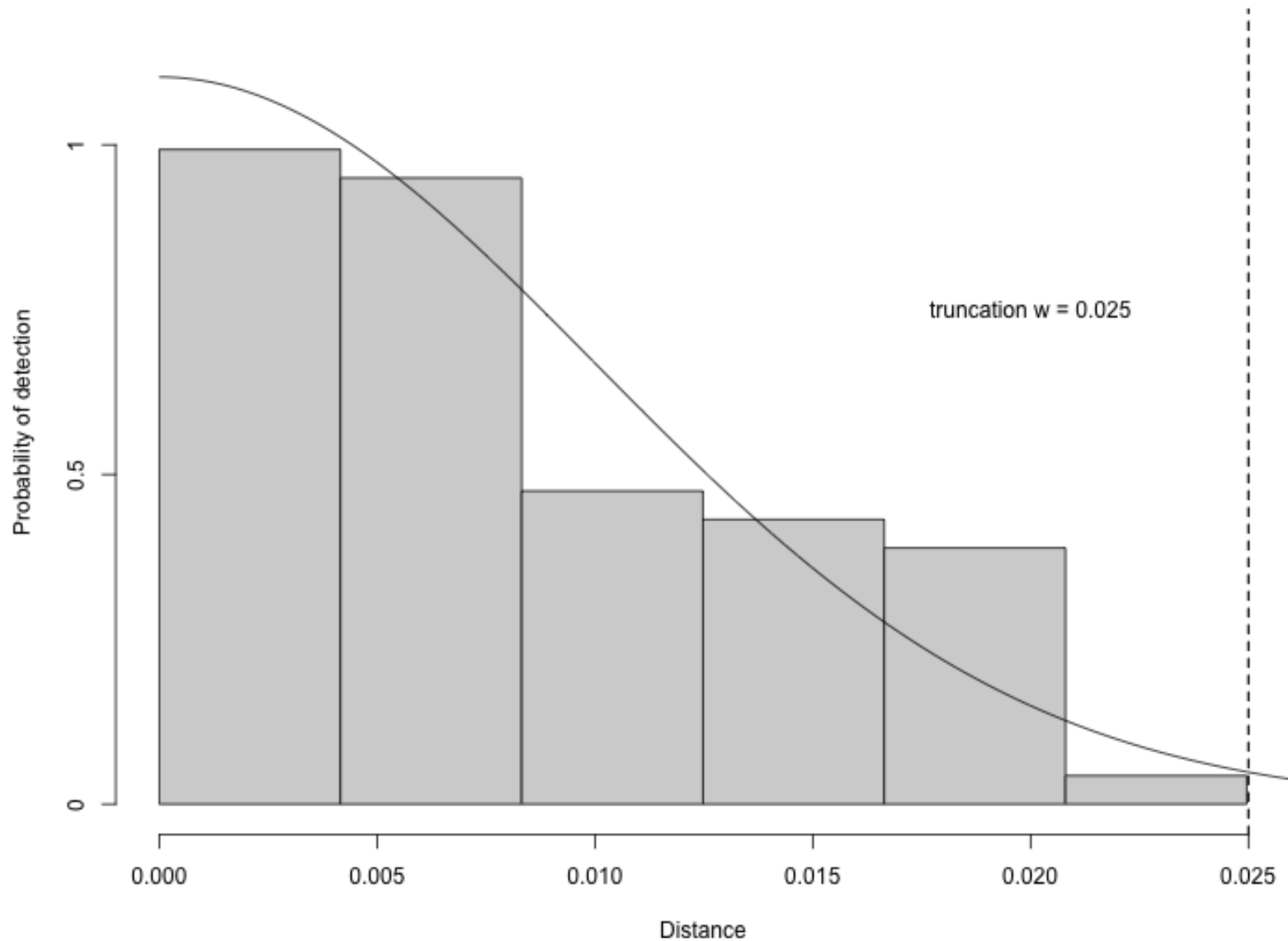
Survey area



Histogram of observed distances



# Detection function



# Distance sampling estimate

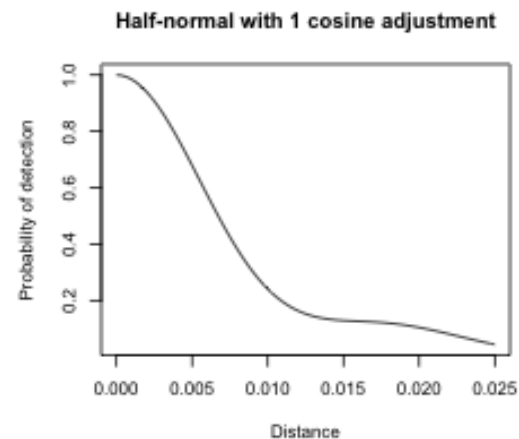
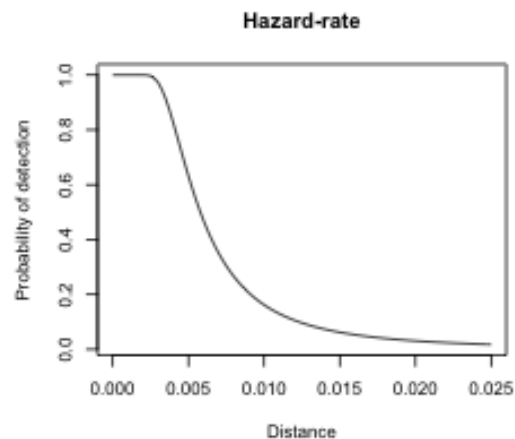
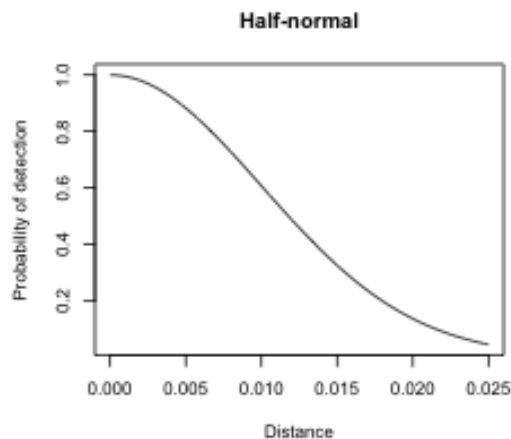
- Surveyed 5 lines (each area  $1 * 2 * 0.025$ )
  - Total covered area  $a = 5 * 1 * (2 * 0.025) = 0.25$
- Probability of detection  $\hat{p} = 0.546$
- Saw  $n = 76$  animals
- Inflate to  $n/\hat{p} = 139.198$
- Estimated density  $\hat{D} = \frac{n/\hat{p}}{a} = 556.8$
- Total area  $A = 1$
- Estimated abundance  $\hat{N} = \hat{D}A = 556.8$

# Distance sampling assumptions

1. Animals are distributed independent of lines
2. On the line, detection is certain
3. Distances are recorded correctly
4. Animals don't move before detection

# What are detection functions?

- $\mathbb{P}(\text{detection} \mid \text{animal at distance } x)$
- (But we want  $\mathbb{P}(\text{detection}) = \hat{p}$ )
- "Integrate out distance" == "area under curve" ==  $\hat{p}$
- Many different forms, depending on the data
- All share some characteristics



# Fitting detection functions (in R!)

- Using the package `Distance`
- Function `ds()` does most of the work

```
library(Distance)  
df_hn <- ds(distdata, truncation=6000)
```

More on this in the practical!



# Horvitz-Thompson-like estimators

- Once we have  $\hat{p}$  how do we get  $\hat{N}$ ?
- Rescale the (flat) density and extrapolate

$$\hat{N} = \frac{\text{study area}}{\text{covered area}} \sum_{i=1}^n \frac{s_i}{\hat{p}_i}$$

- $s_i$  are group/cluster sizes
- $\hat{p}_i$  is the detection probability (from detection function)

# Why spatial modelling?

## Horvitz-Thompson limitations

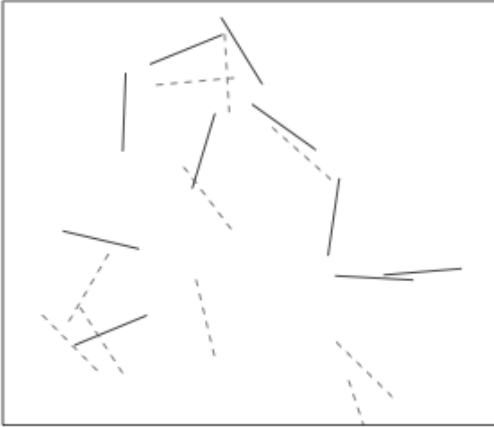
# Hidden in this formula is a simple assumption

- Probability of sampling every point in the study area is equal
- Is this true? Sometimes.
- If (and only if) the design is randomised

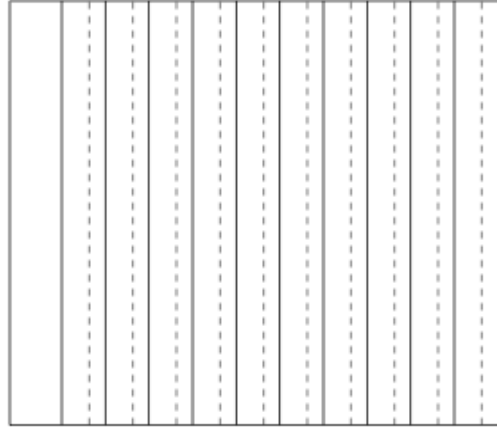
$$\hat{N} = \frac{\text{study area}}{\text{covered area}} \sum_{i=1}^n \frac{s_i}{\hat{p}_i \mathbb{P}(\text{included})}$$

# Many faces of randomisation

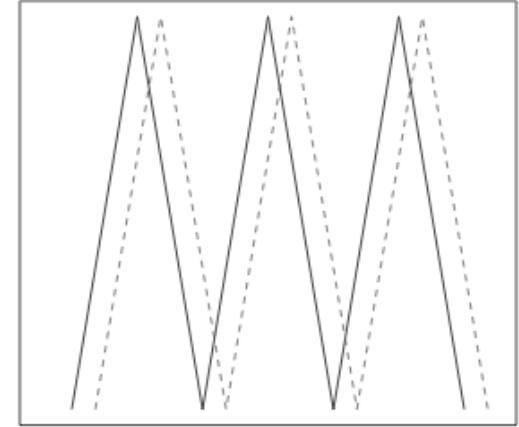
**random placement**



**random offset parallel lines**

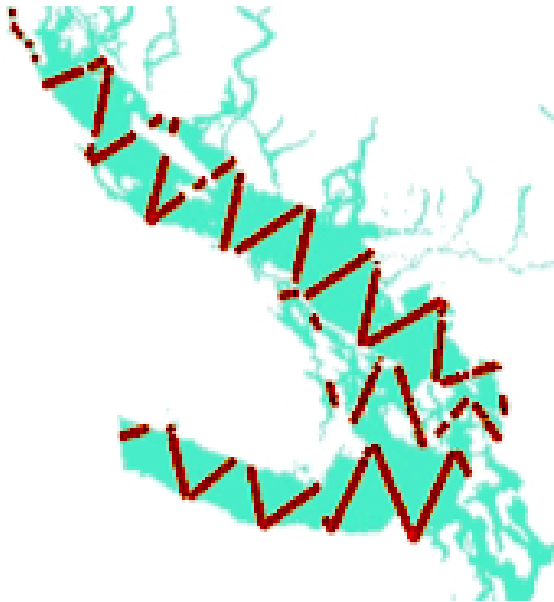


**random offset zigzag**



# Randomisation & coverage probability

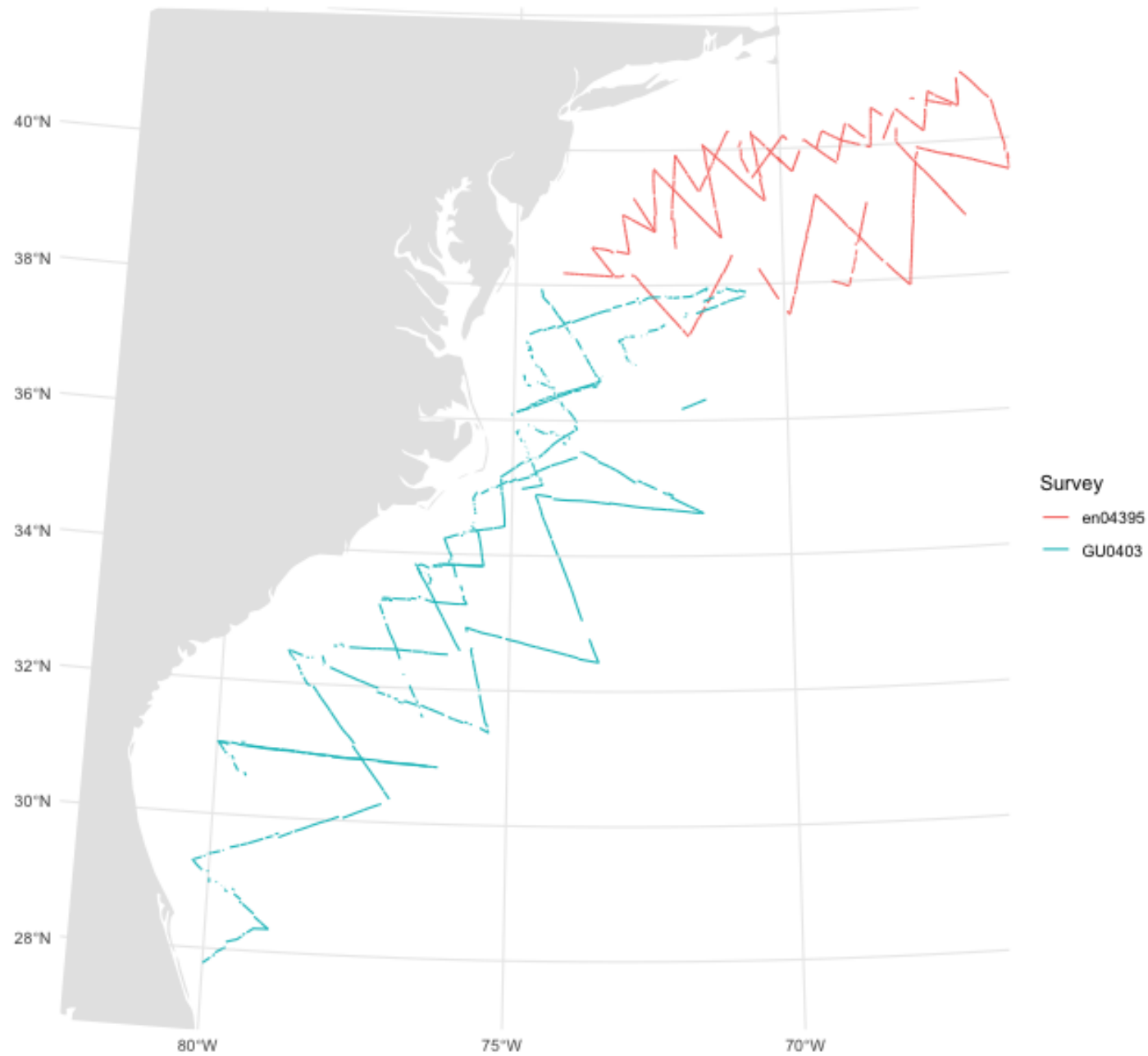
- H-T equation above assumes even coverage
- (Distance for Windows can estimate this)



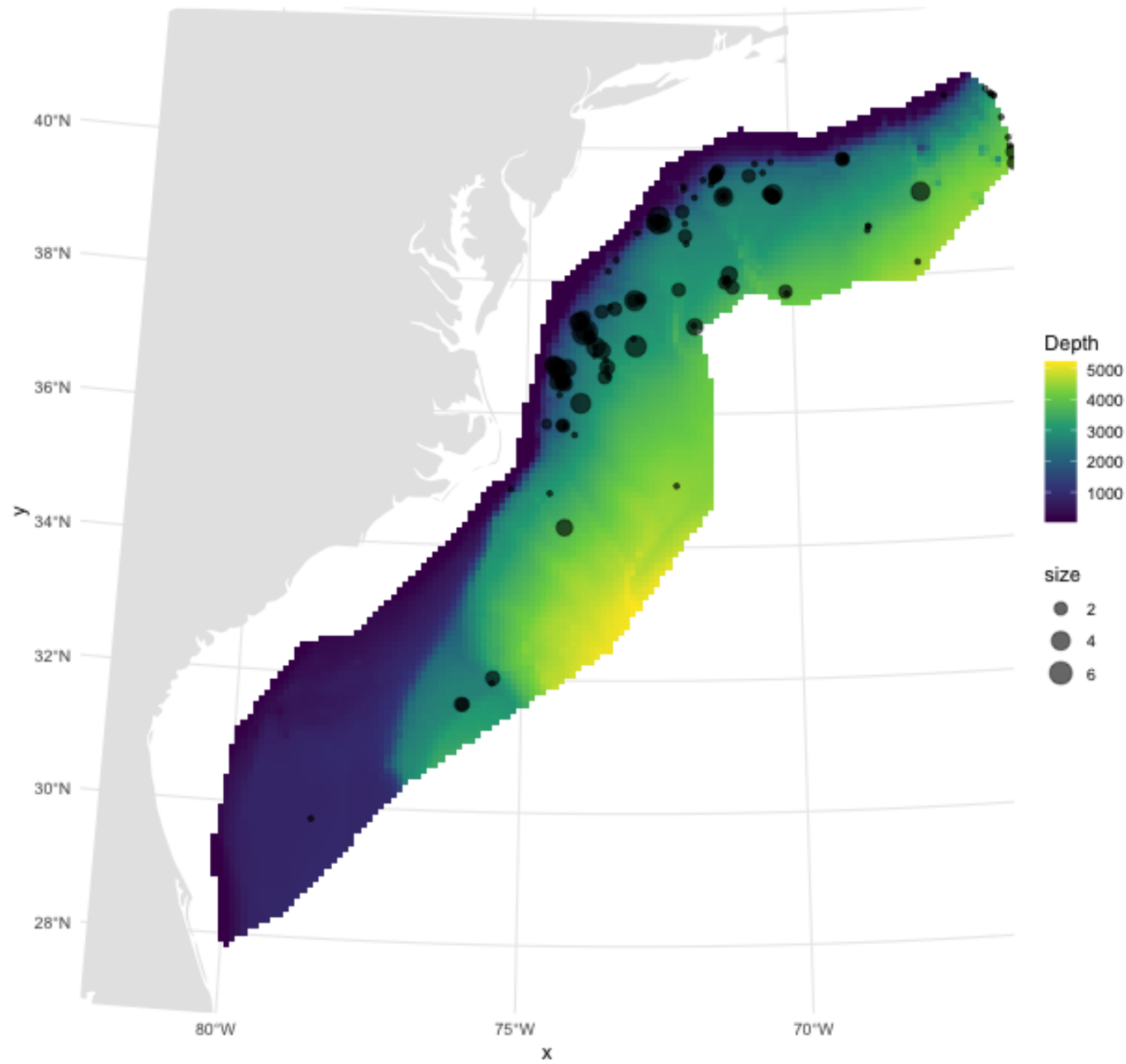
# Why spatial modelling?

Extra spatial information

# Extra information

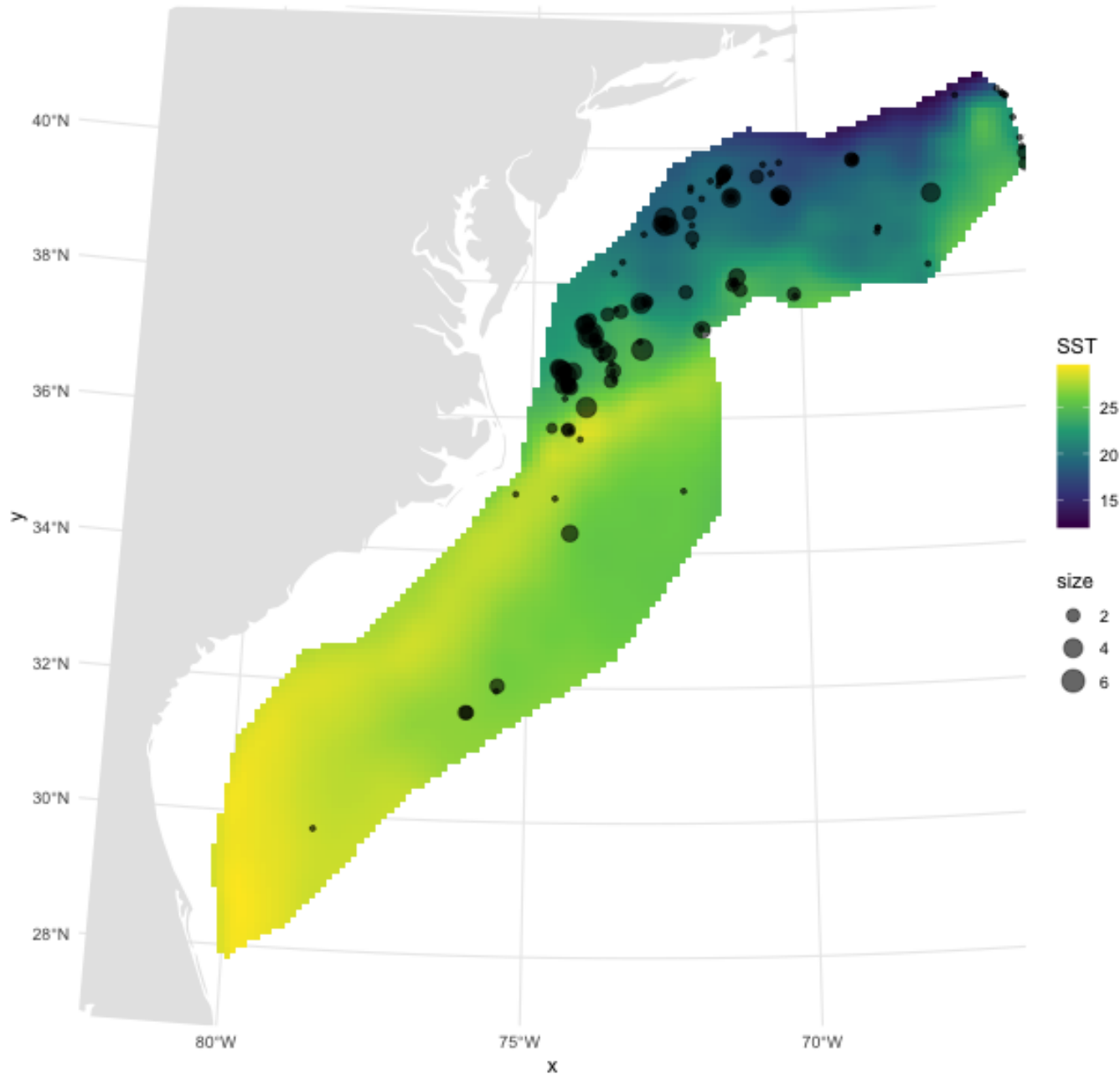


# Extra information - depth



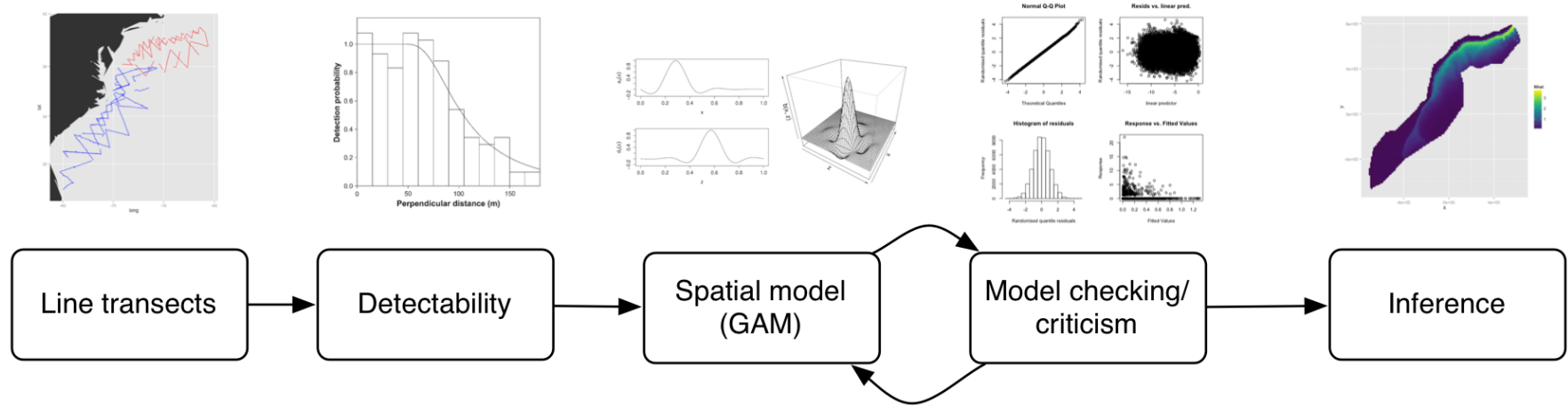


# Extra information - SST



# Density Surface Modelling overview

# Density Surface Modelling flow diagram

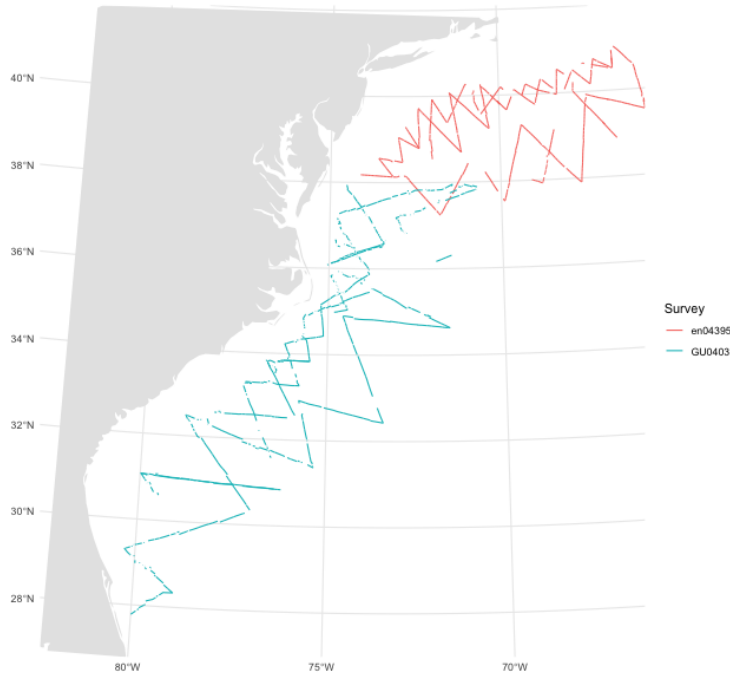


# Modelling requirements

- Account for effort
- Flexible/interpretable effects
- Predictions over an arbitrary area
- Include detectability

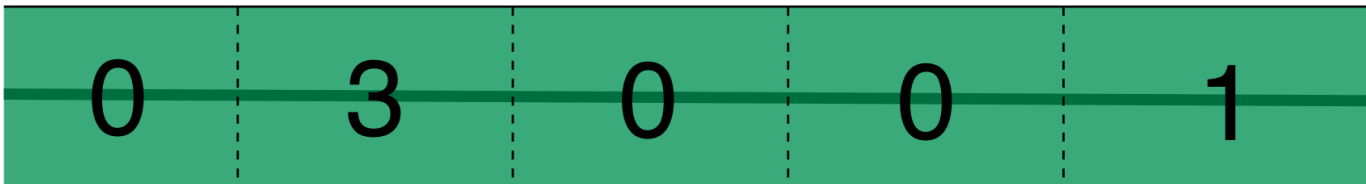
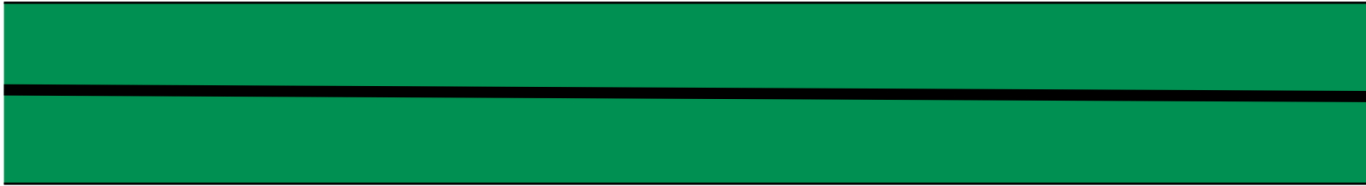
# Accounting for effort

# Effort



- Have transects
- Variation in counts and covars along them
- Want a sample unit w/ minimal variation
- "Segments": chunks of effort

# Chopping up transects

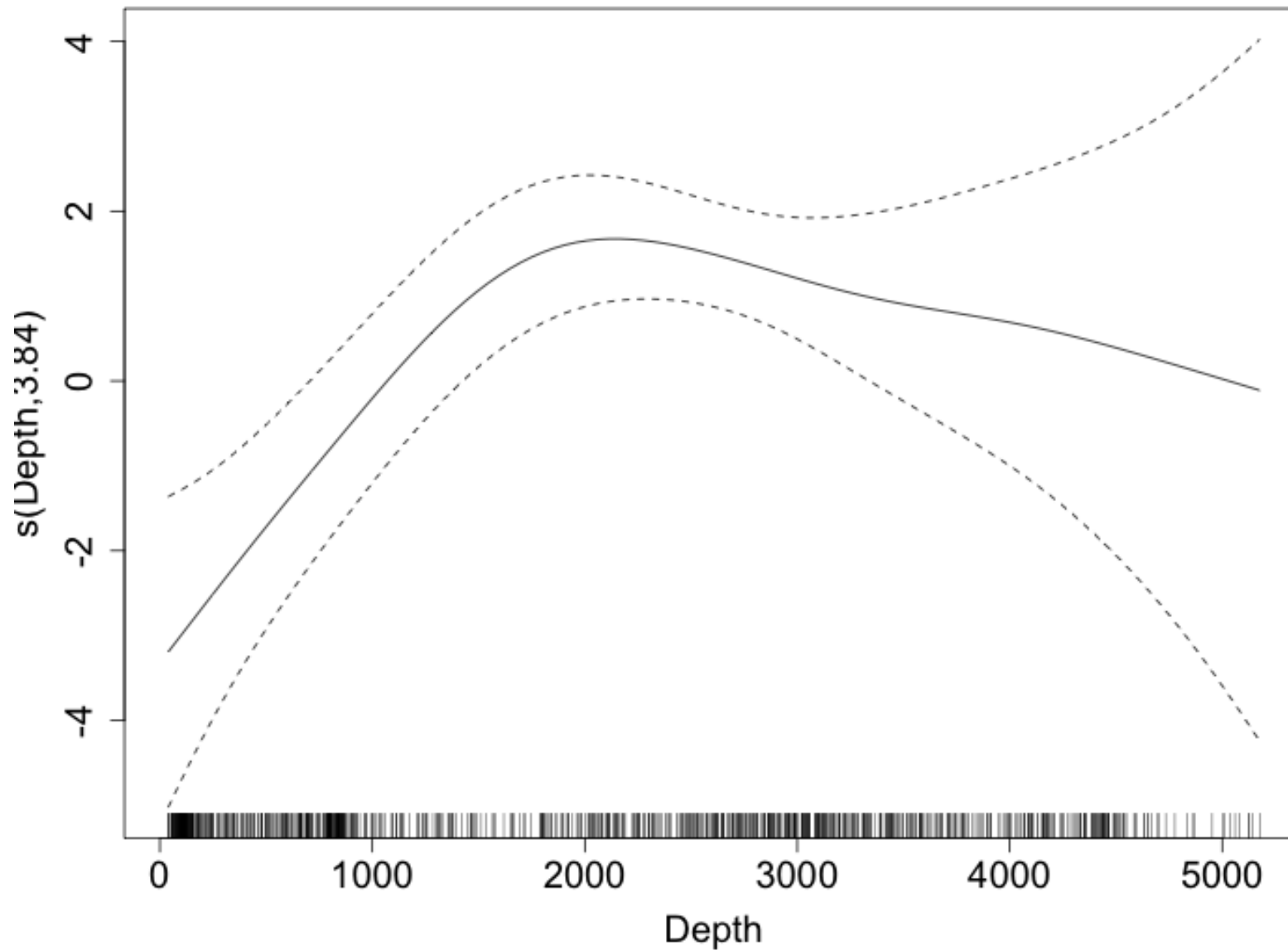


Physeter catodon by Noah Schlottman

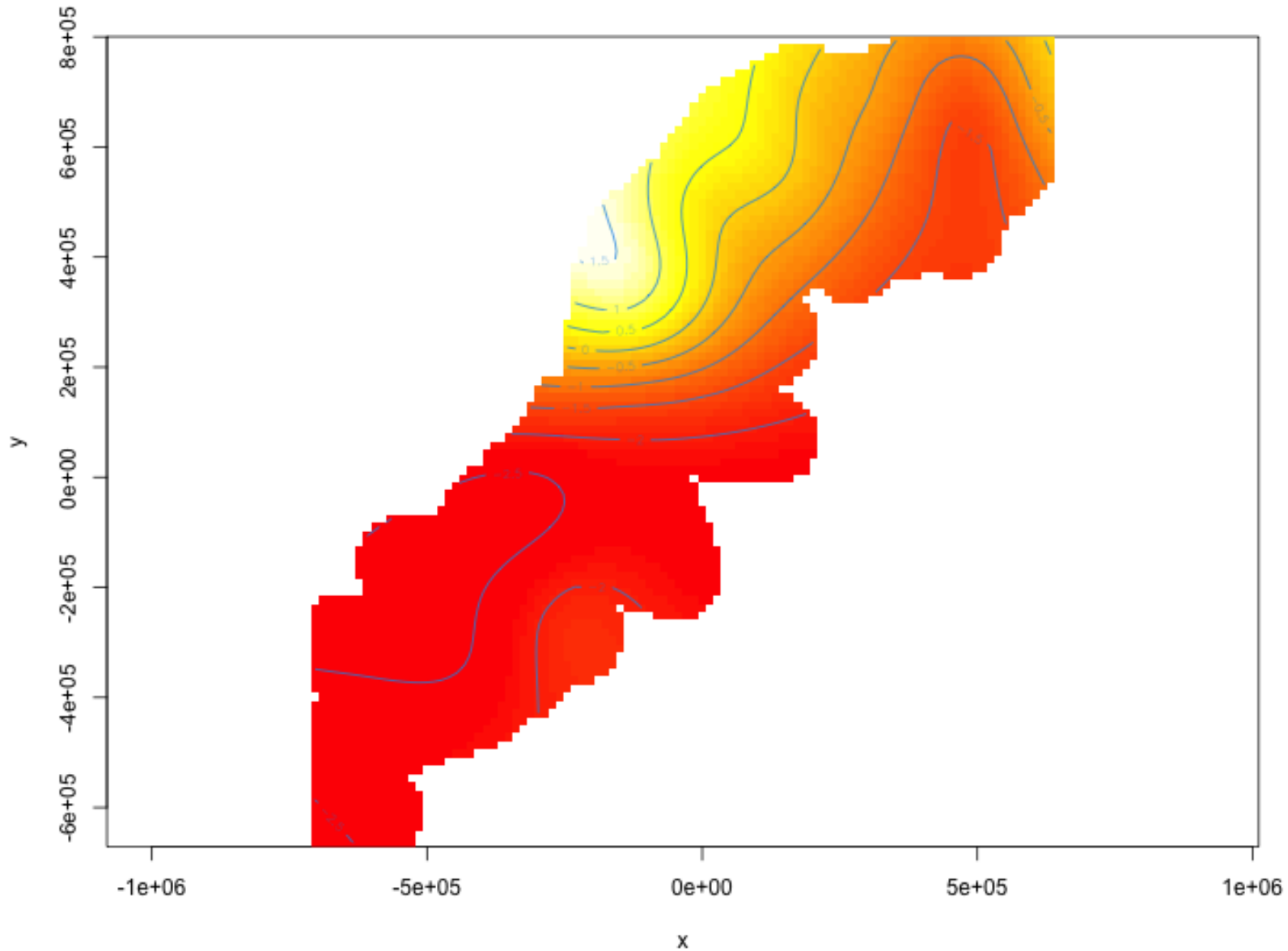
Flexible, interpretable effects



# Smooth response

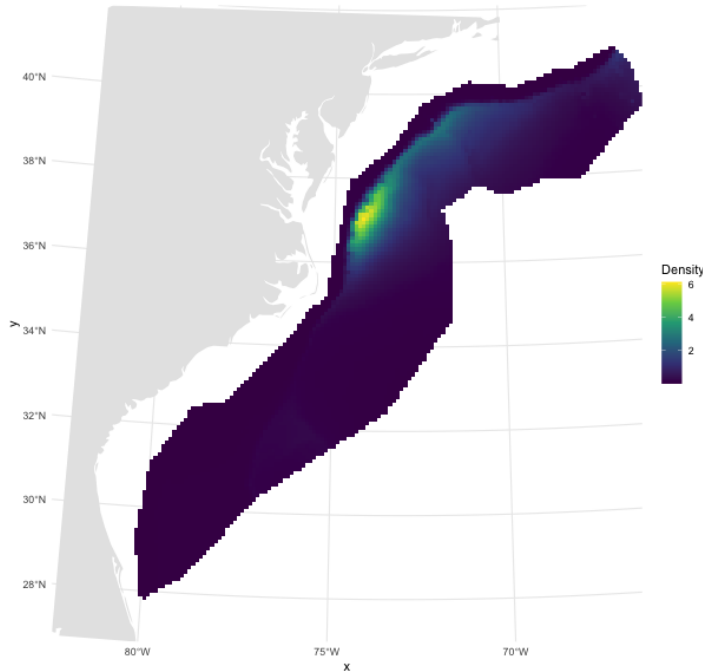


# Explicit spatial effects



# Predictions

# Predictions over an arbitrary area



- Don't want to be restricted to predict on segments
- Predict within survey area
- Extrapolate outside (with caution)
- Working on a grid of cells

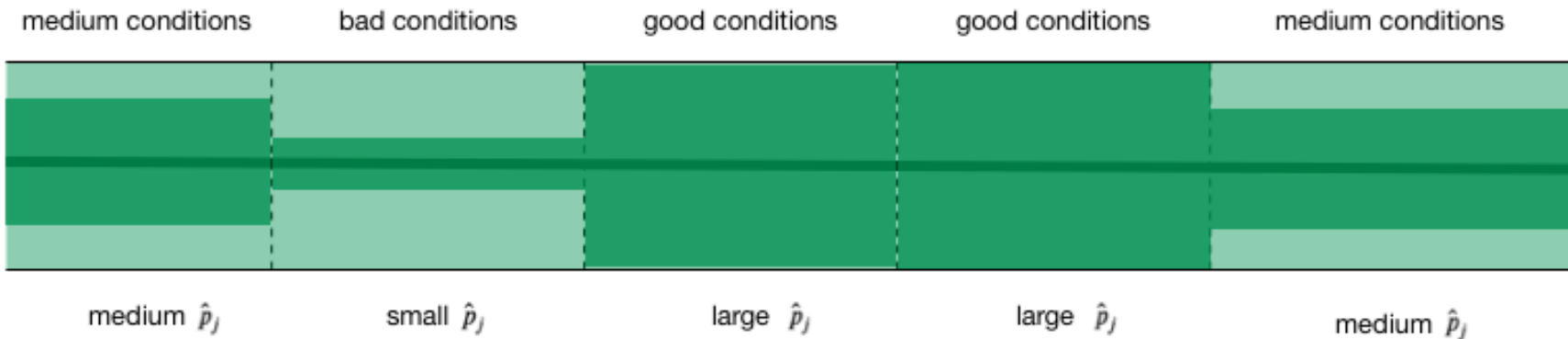
# Detection information

# Including detection information

- Two options:
  - adjust areas to account for **effective effort**
  - use **Horvitz-Thompson estimates** as response

# Count model

- Area of each segment,  $A_j$ 
  - use  $A_j \hat{p}_j$
- ☁ effective strip width (  $\hat{\mu} = w\hat{p}$  )
- Response is counts per segment
- "Adjusting for effort"

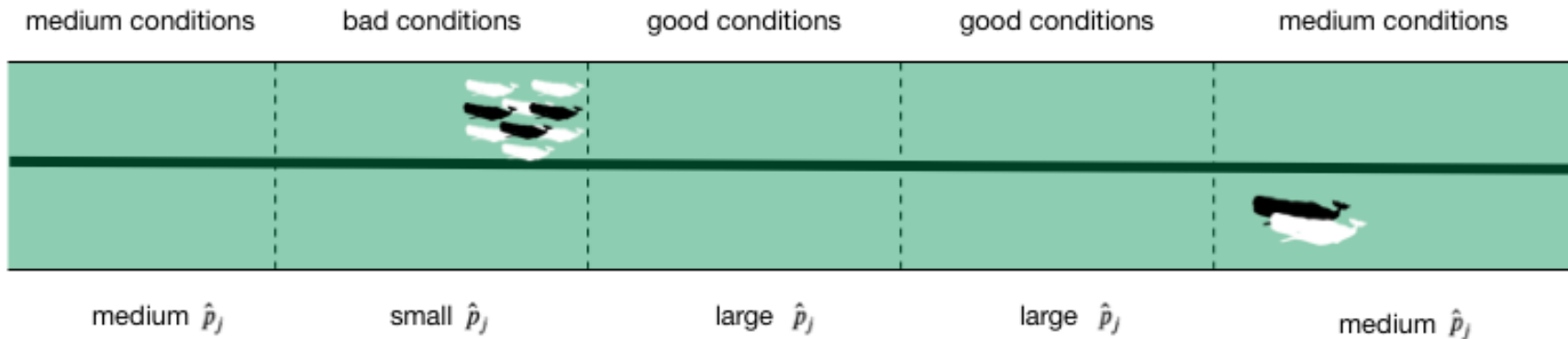


# Estimated abundance

- Effort is area of each segment
- Estimate H-T abundance per segment

$$\hat{n}_j = \sum_i \frac{s_i}{\hat{p}_i}$$

(where the  $i$  observations are in segment  $j$ )





# Detectability and covariates





- 2 covariate "levels" in detection function
  - "Observer"/"observation" -- change **within** segment
  - "Segment" -- change **between** segments
- "Count model" only lets us use segment-level covariates
- "Estimated abundance" lets us use either

# When to use each approach?

- Generally "nicer" to adjust effort
- Keep response (counts) close to what was observed
- **Unless** you want observation-level covariates

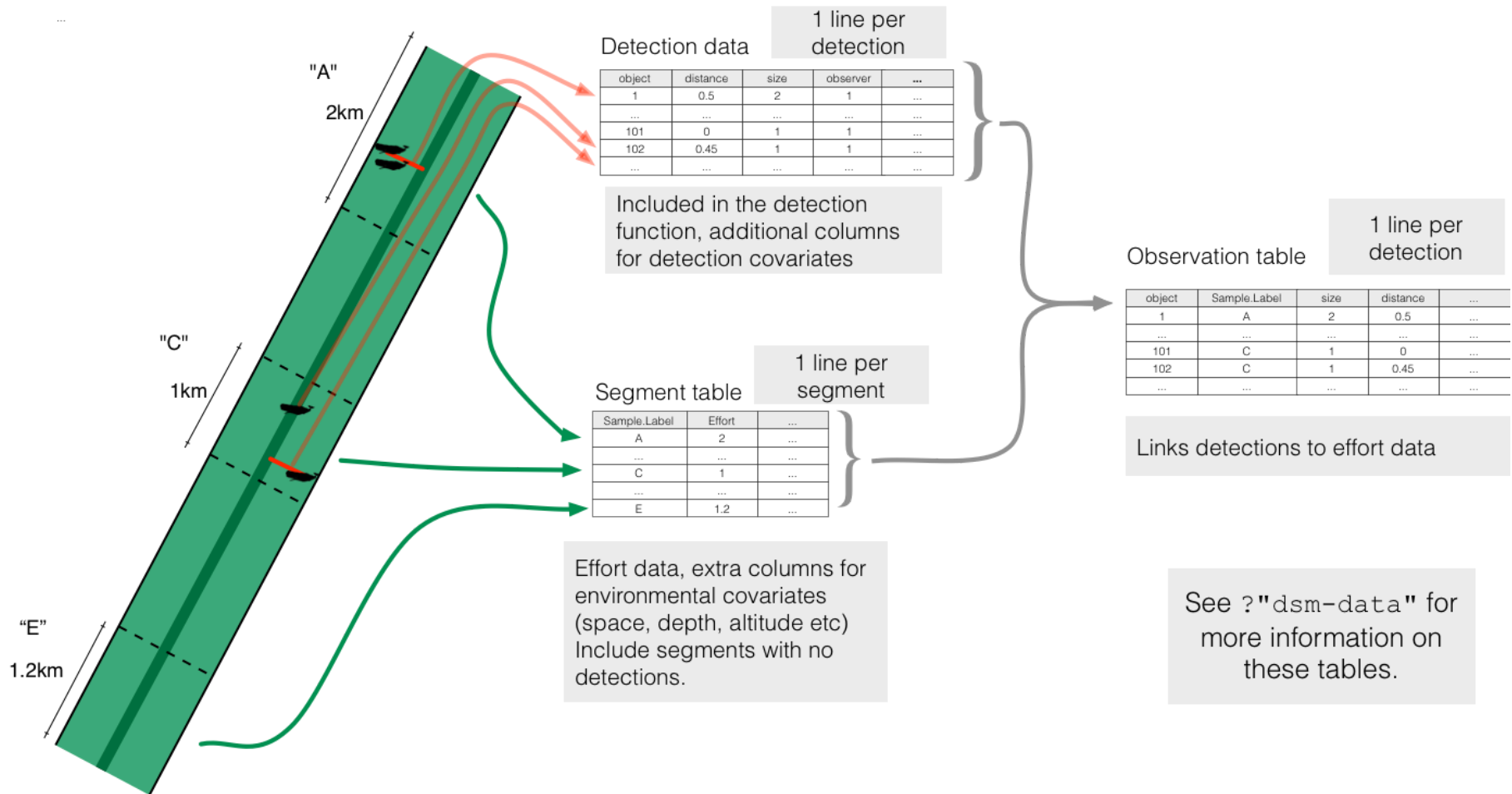
# Data requirements

# What do we need?

- Need to "link" data
  -  Distance data/detection function
  -  Segment data
  -  Observation data (segments  detections)

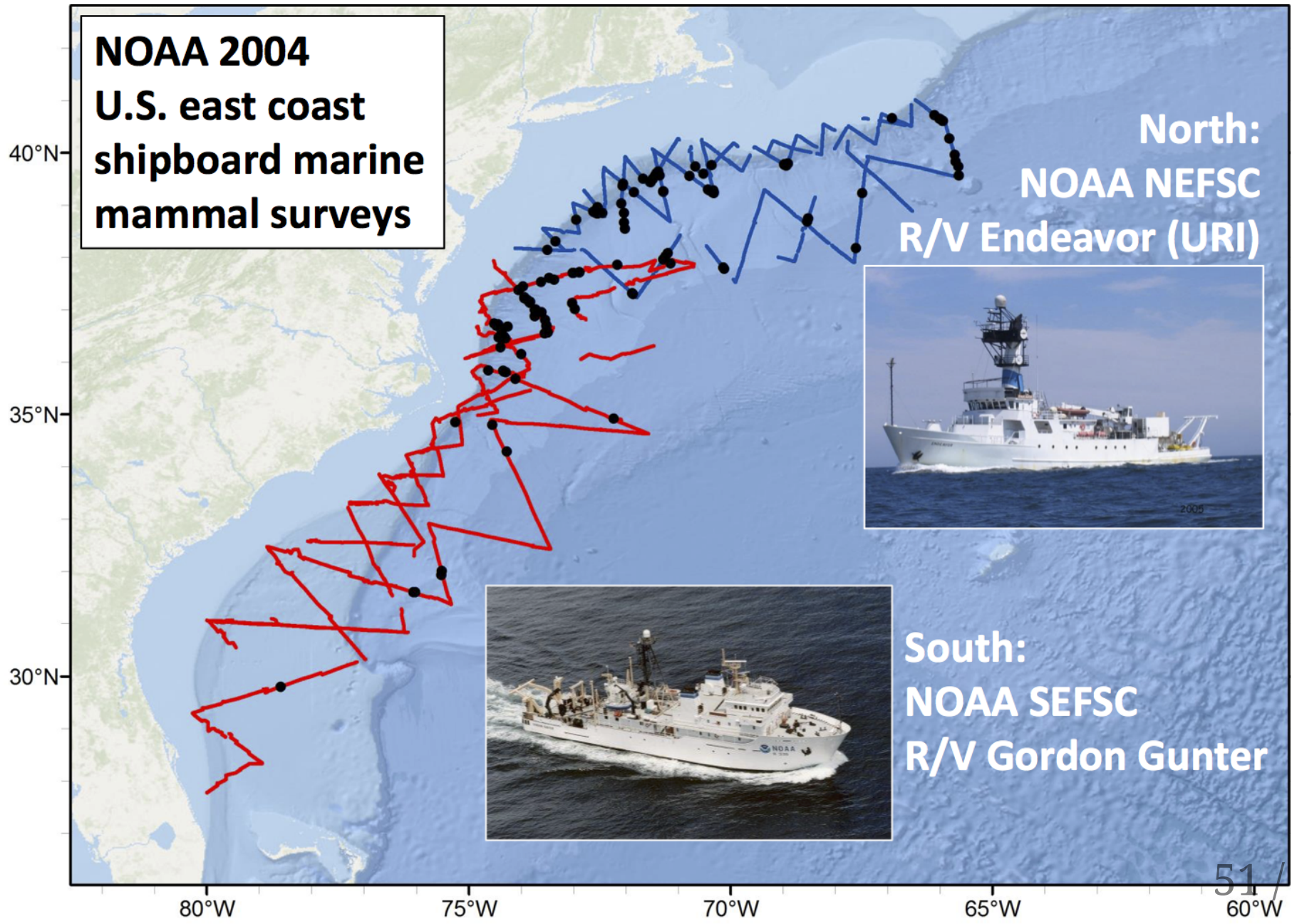
More info on course website.

# Density surface model data setup for package dsm



# Example data

# Example data



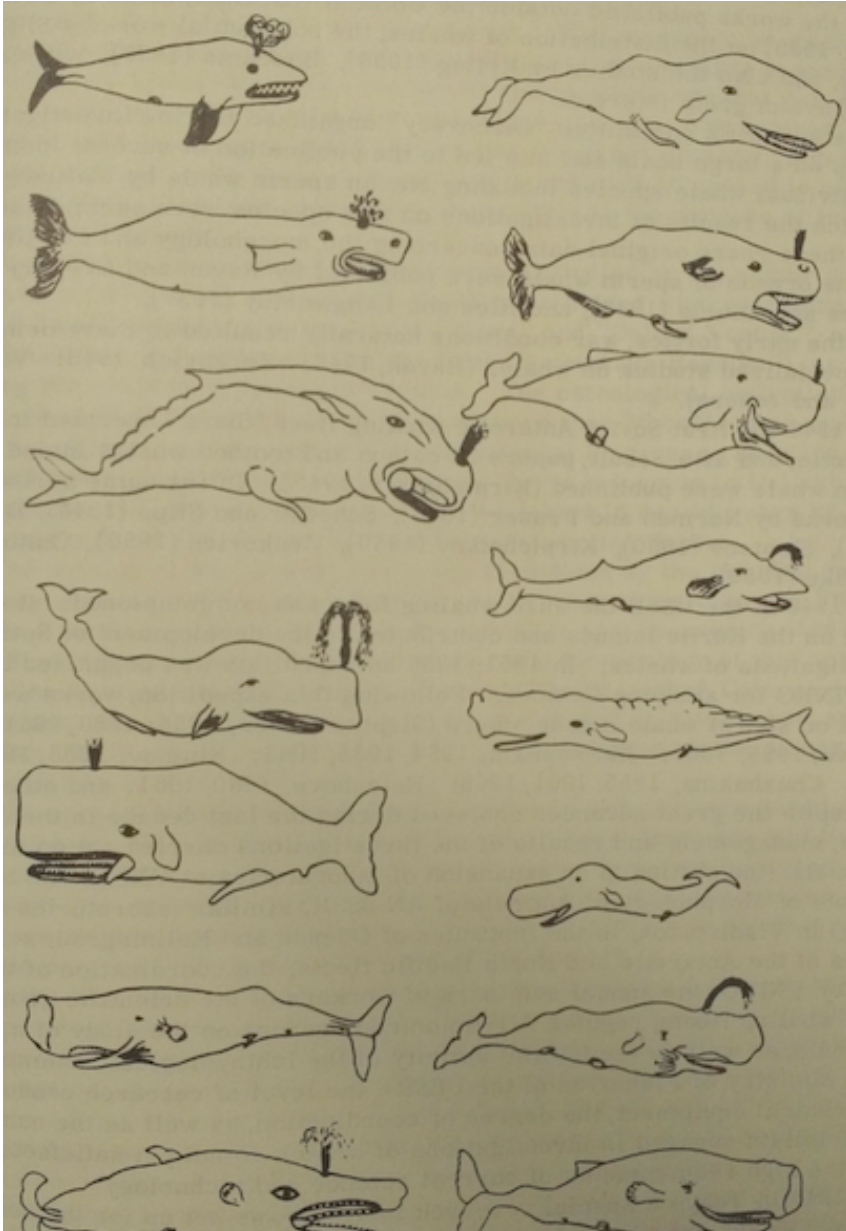


# Example data





# Sperm whales



- Hang out near canyons, eat squid
- Surveys in 2004, US east coast
- Thanks to Debi Palka (NOAA NEFSC), Lance Garrison (NOAA SEFSC) for data. Jason Roberts (Duke University) for data prep.

# Recap

- Model counts or estimated abundance
- The effort is accounted for differently
- Flexible models are good
- Incorporate detectability
- 2 tables + detection function needed