Survey design

- Introduction
- Some concepts
  - Coverage
  - Plus/minus sampling and edge effects
- Line transect designs
- Effort calculations for design
- Stratification
- Example surveys

See

- *Chapter 7 of Buckland et al. (2001) Introduction to Distance Sampling*
- *Chapter 7 of Buckland et al. (2004) Advanced Distance Sampling*
- *Chapter 2 of Buckland et al. (2015) Distance Sampling: Methods and Applications*

CREEM
Centre for Research into Ecological and Environmental Modelling
University of St Andrews

Study design components

What question are you trying to answer?

What precision is required to answer the question defensibly?

What effort is needed to produce the required level of precision?

Where is that effort deployed within the study area?

Did population change in the past x years?

Because of
- Habitat alternation
- Poaching prevention
- Changes to harvest regulations

$$
\begin{aligned}
\hat{\sigma}^2_{\mathrm{res}} &= \frac{1}{n}\sum^{n} \mathrm{Var}(\hat{A}_i) \\
&= \frac{1}{n}\sum^{n} \mathrm{cv}_1^2 A_1 A_i \\
&= \frac{\mathrm{cv}_1^2 A_1^2}{n}\sum^{n} [1 + r(i-1)] \\
&= (\mathrm{cv}_1 A_1)^2\left[1 + \frac{r}{2}(n-1)\right].
\end{aligned}
$$

$$L = \left(\frac{b}{\left[cv_t(\widehat{D})\right]^2}\right) \times \frac{L_0}{n_0}$$

Are adequate resources available to:
- fund this amount of effort
- fund a pilot study
- and (if necessary) pay for statistical advice?

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

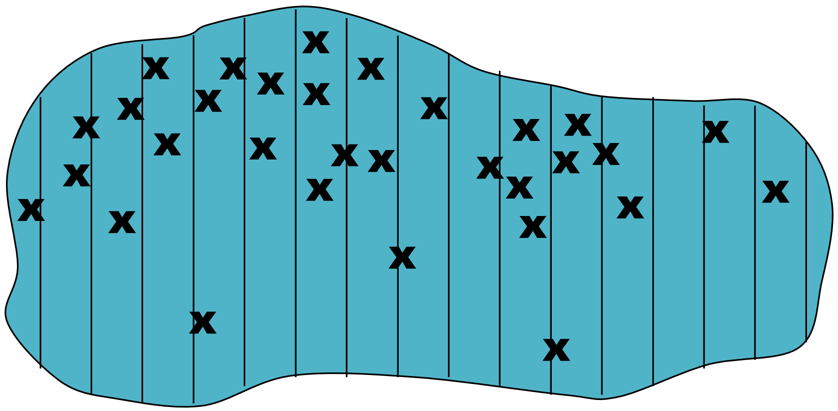
Survey design

Why is design (and good field methods) so important for distance sampling surveys?

- Distance sampling uses design-based estimates
- It is extremely hard and often impossible to compensate for poor design at the analysis stage
- Good design makes analysis less complicated

What do we need from our design?

- Surveyed area needs to be a representative sample of the study area
  - Uniform coverage
  - Use random allocation of transect locations
  - Do not use roads, tracks etc.
- Maximise the number of transects
  - Many short lines are better than a few long lines
- Minimise variability between transects
  - Try to orientate lines perpendicular to density contours or to linear features (e.g. woodland edge or coastline)
- Lines are generally preferable to points



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Design comparison metric

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

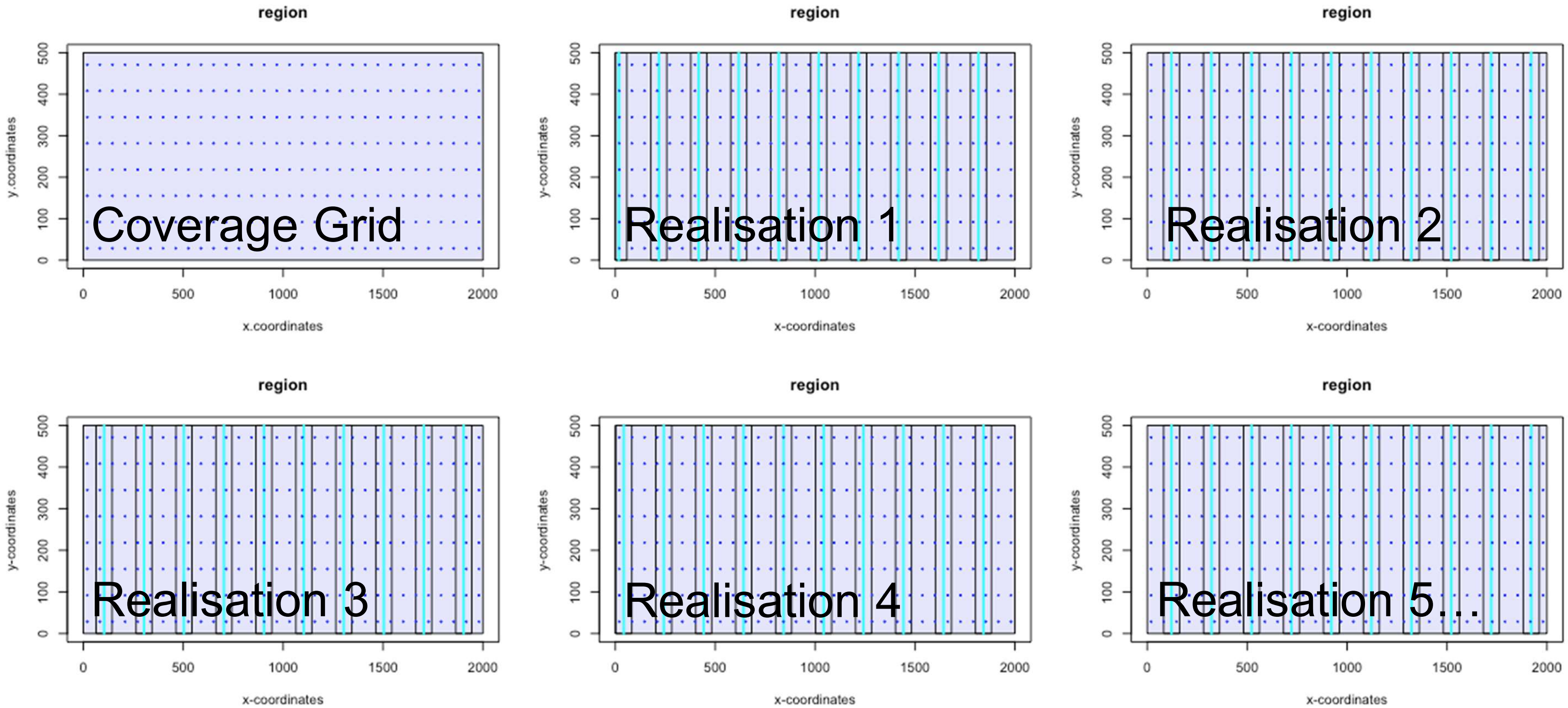
Terminology

Design – a description of how the transects are laid out throughout the survey region

Survey – a single realisation of a design

Coverage score – the average number of times a particular point in the study region will be within a simulated "covered area"

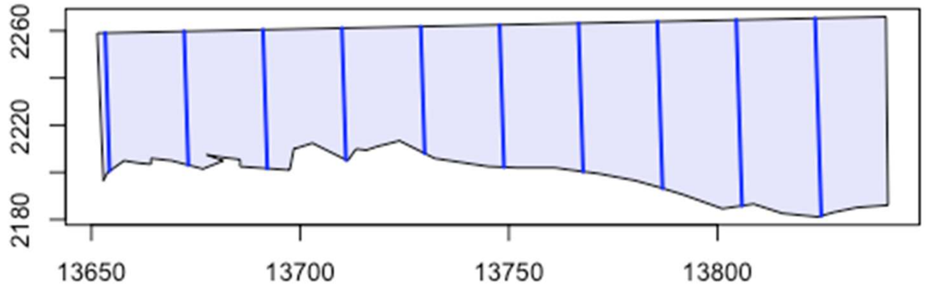
Coverage

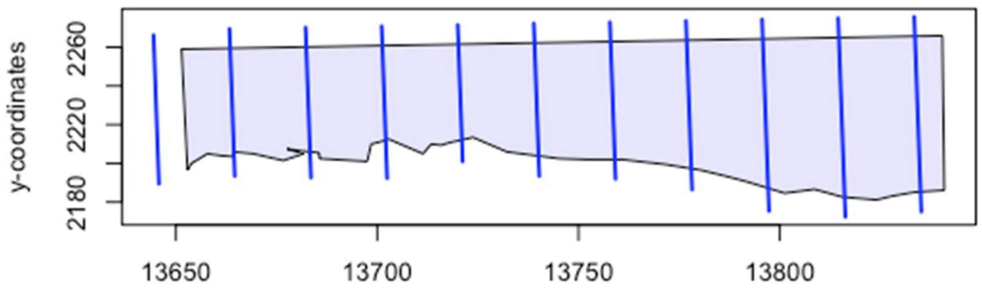


CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

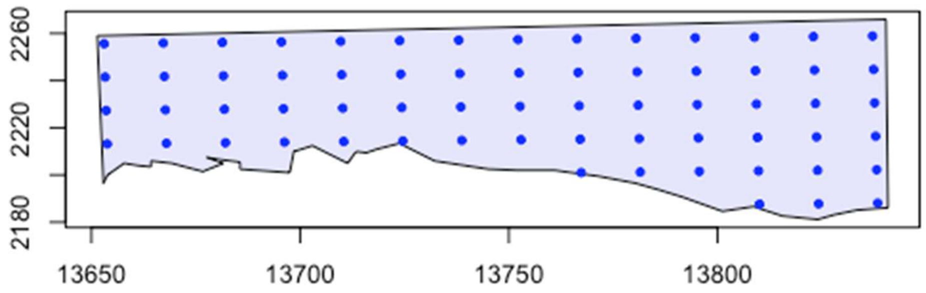
Minus v Plus Sampling

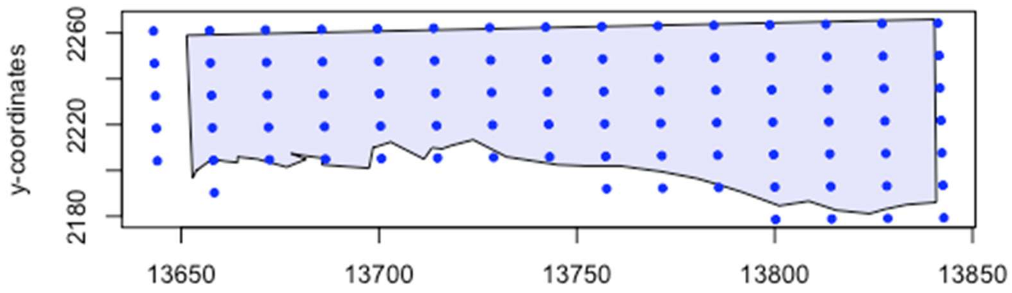




MINUS SAMPLING

PLUS SAMPLING

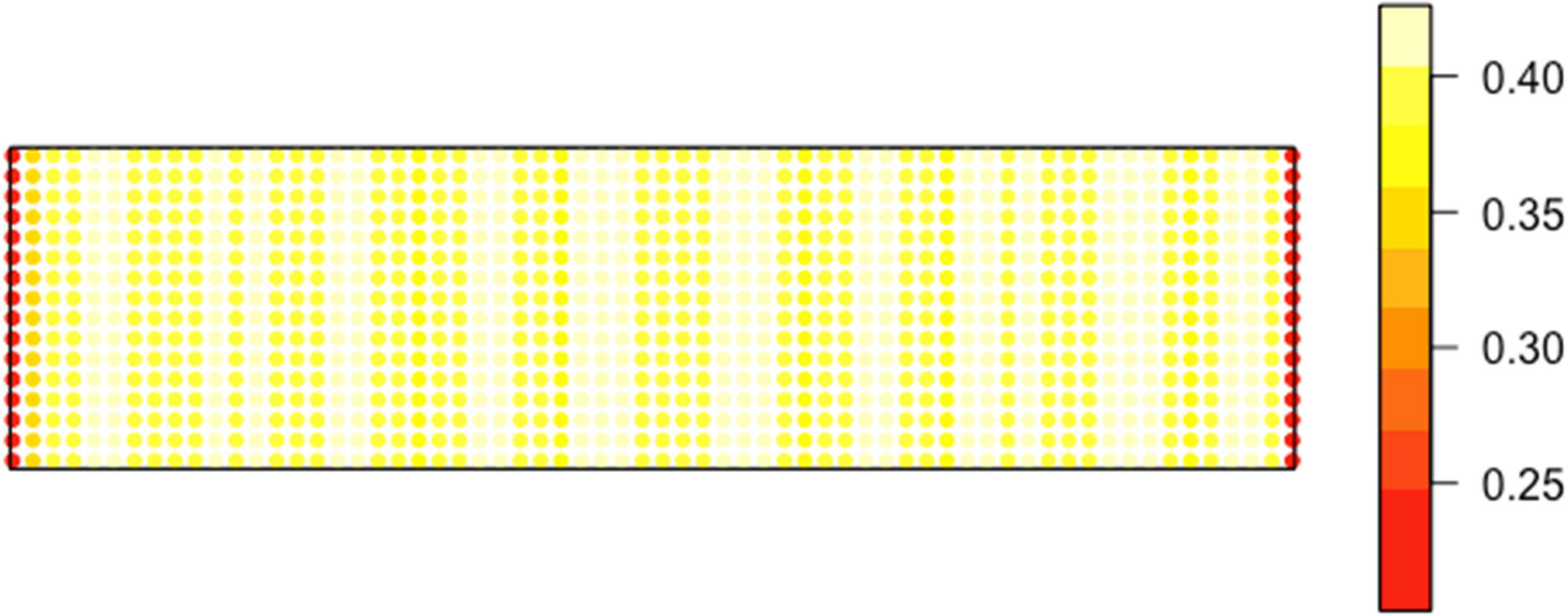




CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

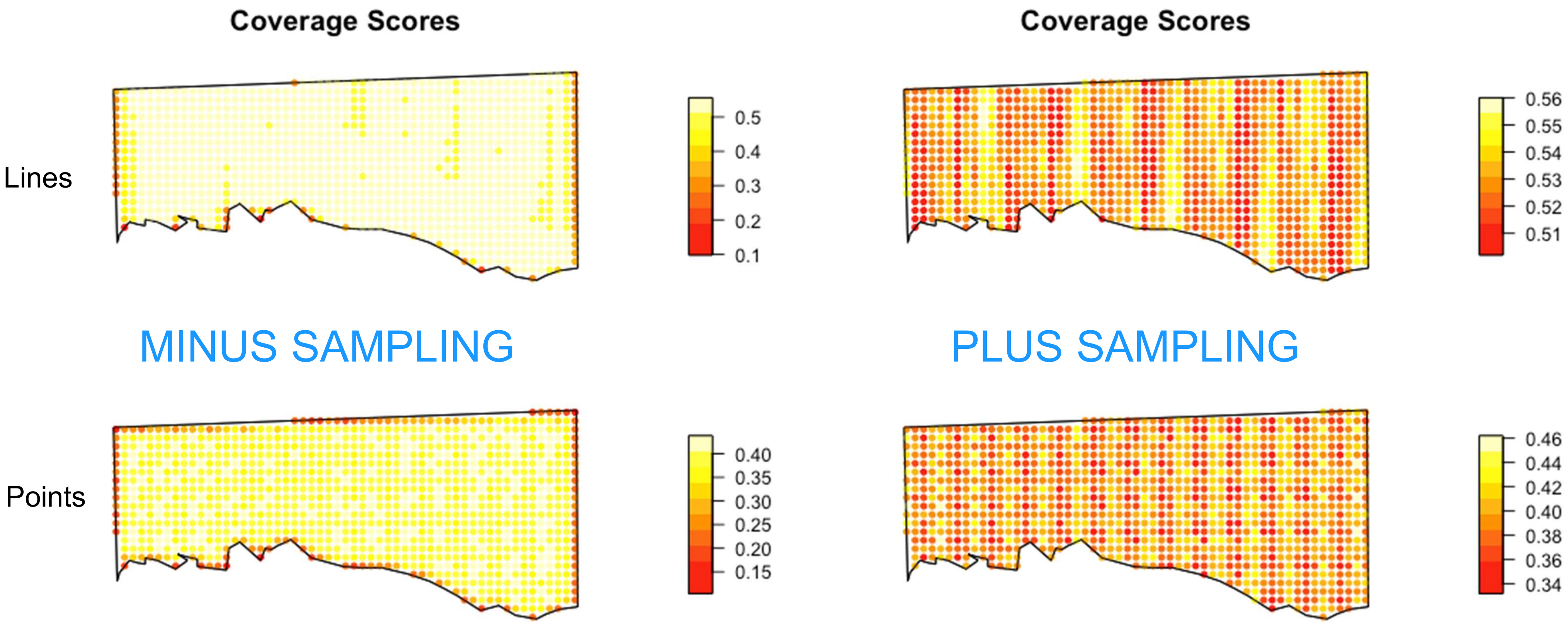
Coverage



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Coverage for 500 repetitions



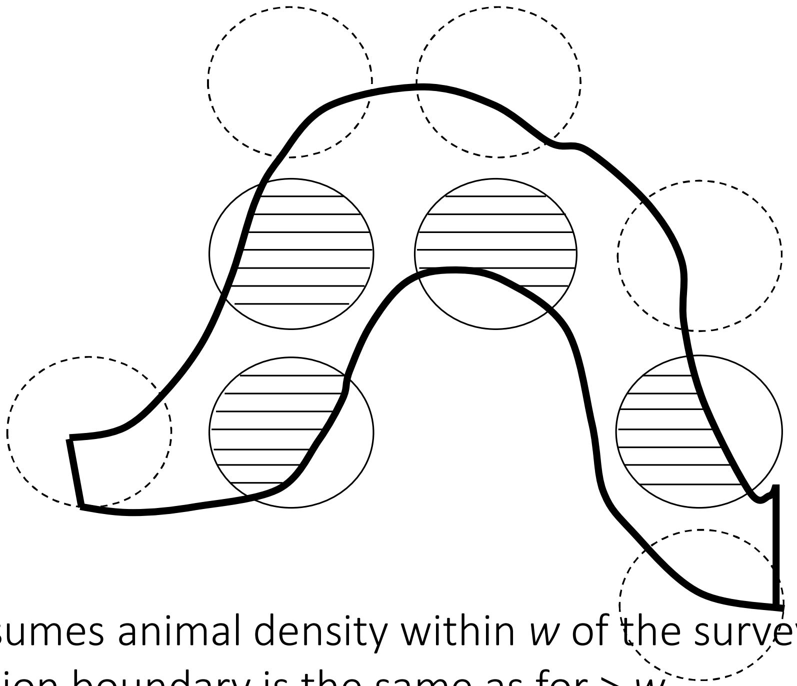
CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Point transect edge effects

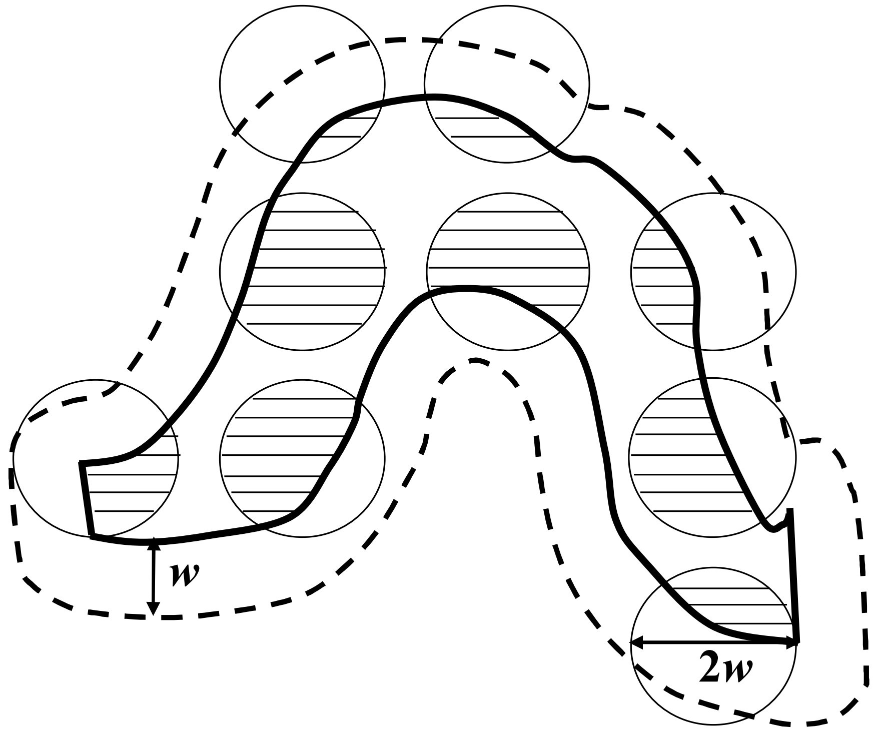
Minus sampling

•Only a problem if study area is very small or narrow relative to *w*



Assumes animal density within *w* of the survey region boundary is the same as for > *w*

Plus sampling

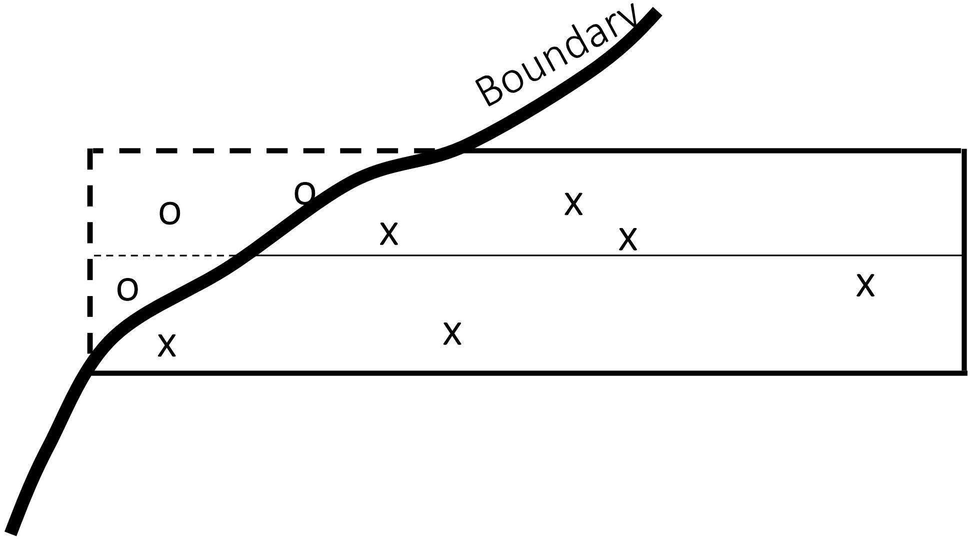


CREEM Centre for Research into Ecological and Environmental Modelling

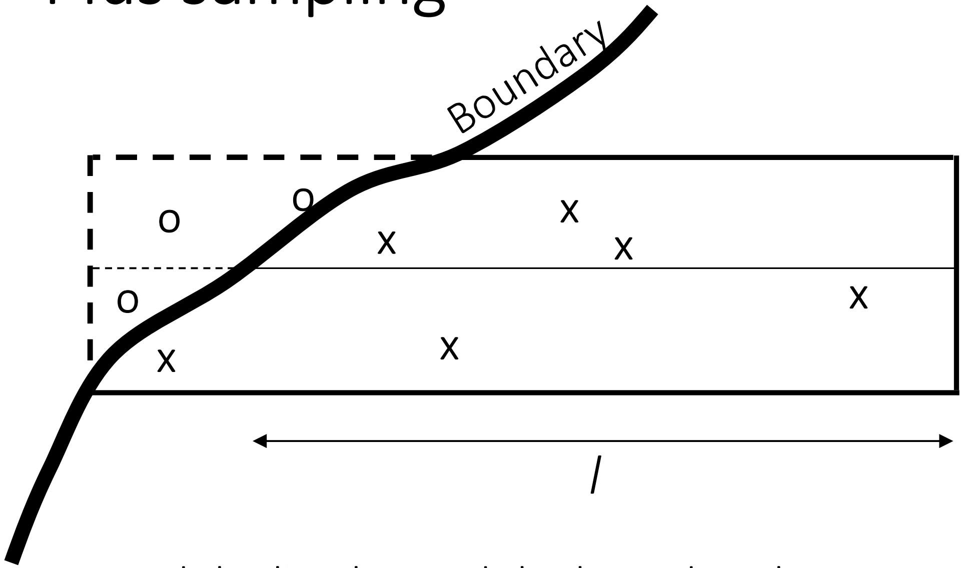
University of St Andrews

Line Transect edge effects

Minus sampling



Plus sampling



Extend the line beyond the boundary, but **don't include the associated effort**, and don't record animals detected outside the region (o)

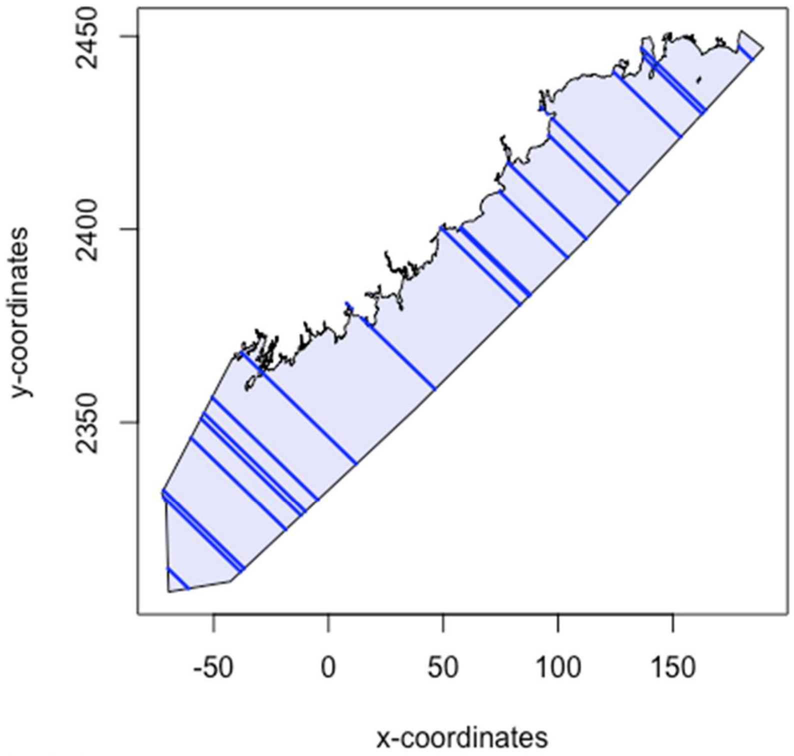
Line Transect Designs

CREEM
Centre for Research into Ecological and Environmental Modelling

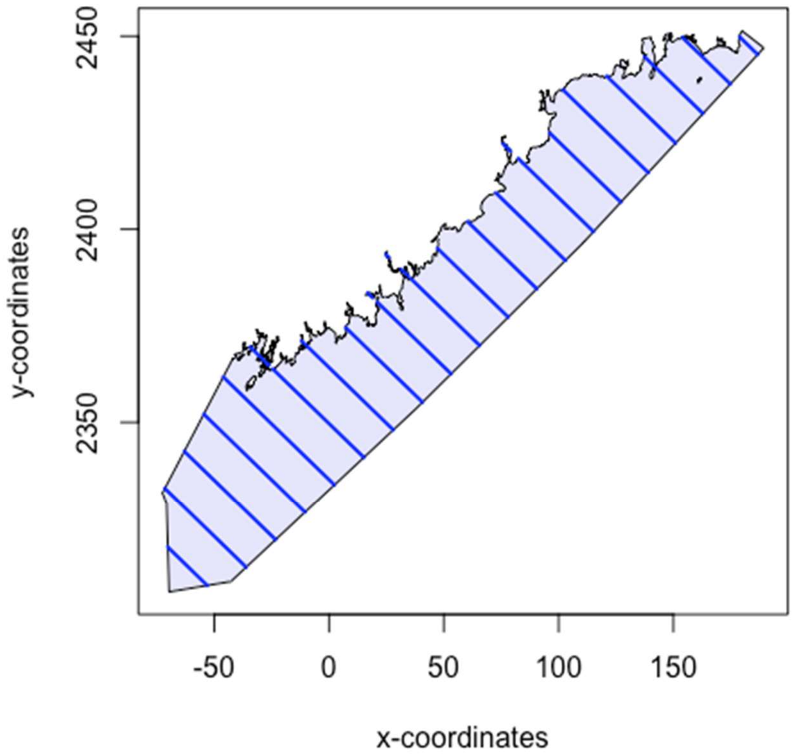
University of St Andrews

Parallel Line Transect Designs

Random Parallel Design



Systematic Parallel Design

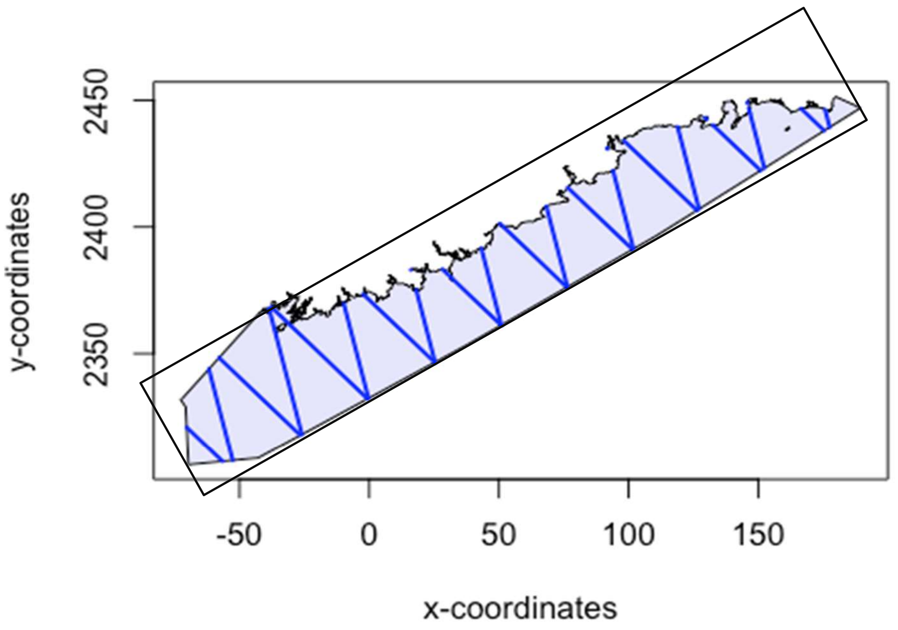


CREEM Centre for Research into Ecological and Environmental Modelling

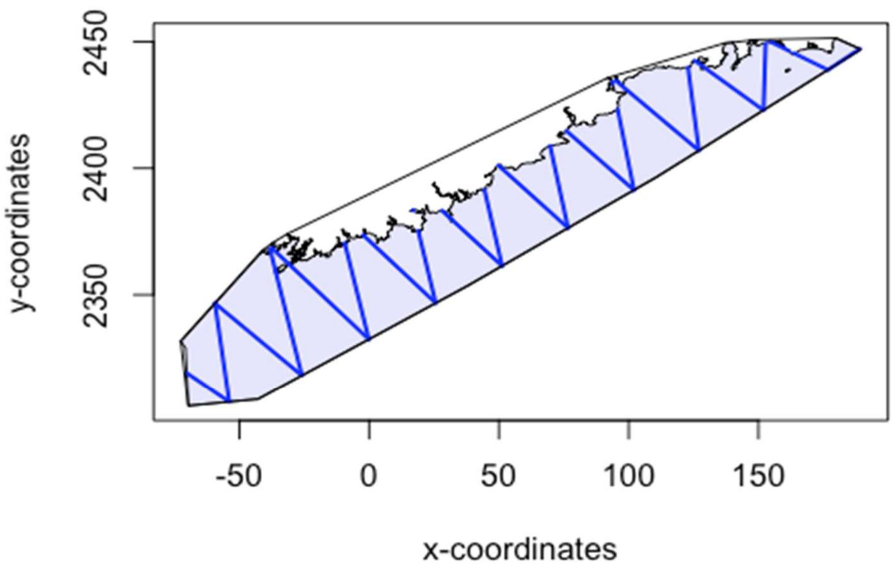
University of St Andrews

Equal Spaced Zigzag Designs

Generated inside a minimum bounding rectangle



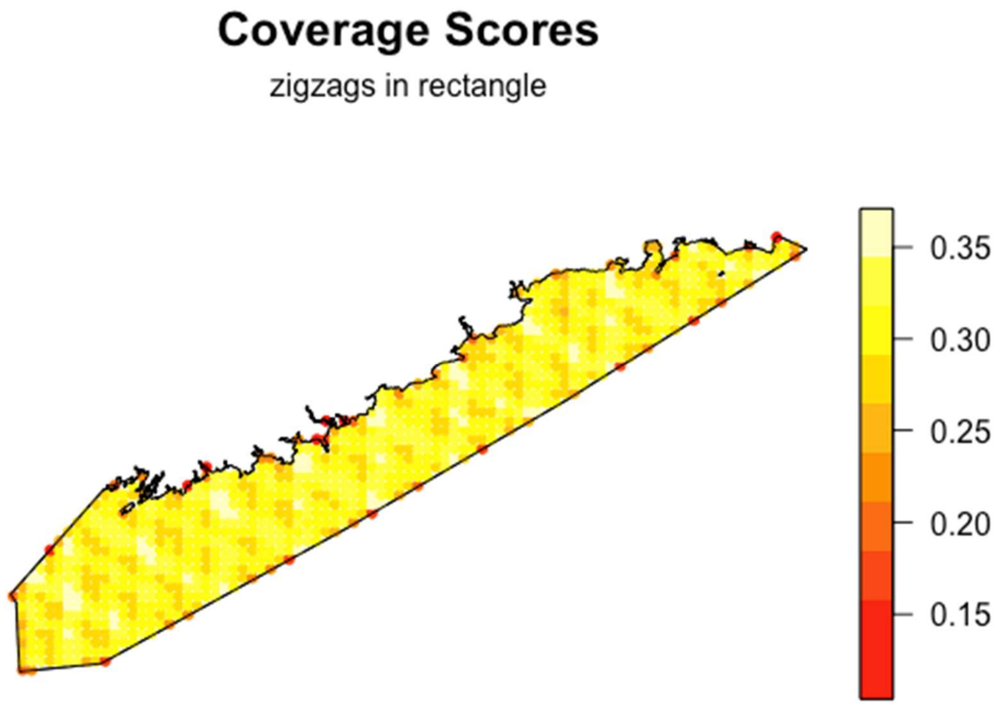
Generated inside a convex hull – like stretching an elastic band around the study region

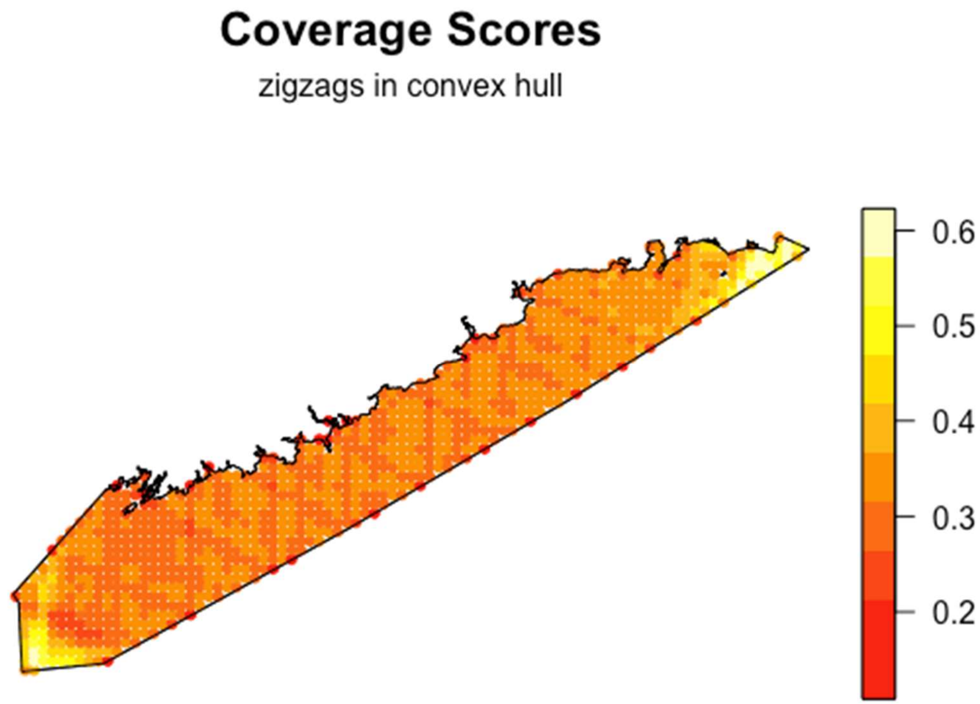


CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Complementary Equal Spaced Zigzags

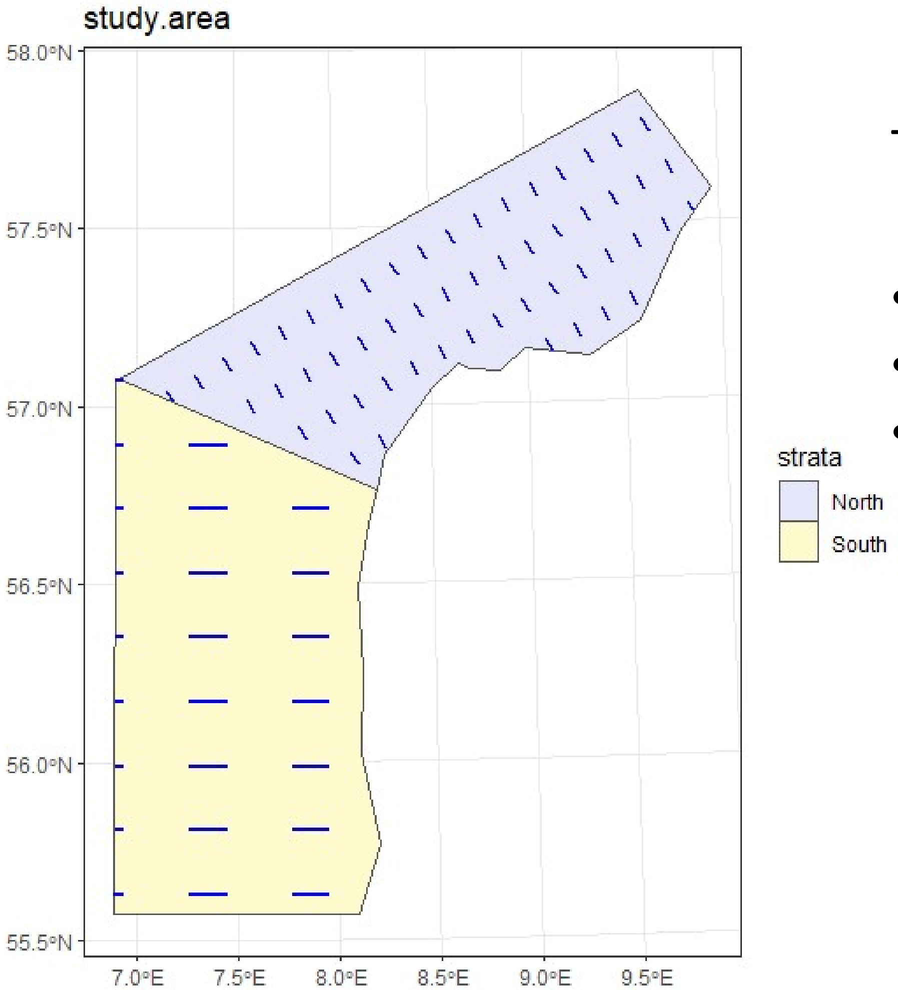




Simulation based on generating 1000 surveys

Comparison of Line Transect Designs

- **Uniform coverage** – parallel line designs and zigzags generated inside a rectangle have uniform coverage (excluding edge effects)
  - *Zigzags inside a convex hull can have non-uniform coverage*
- Systematic designs (systematic parallel and zigzag) have more **even coverage** for any given realisation
- Zigzags generated inside a convex hull are usually more **efficient** (less off-effort transit between transects) and complementary zigzags can further improve efficiency
- Can have **overlap** of samplers in the parallel random design. Also some overlap in zigzag designs.



Transects not full width of study area

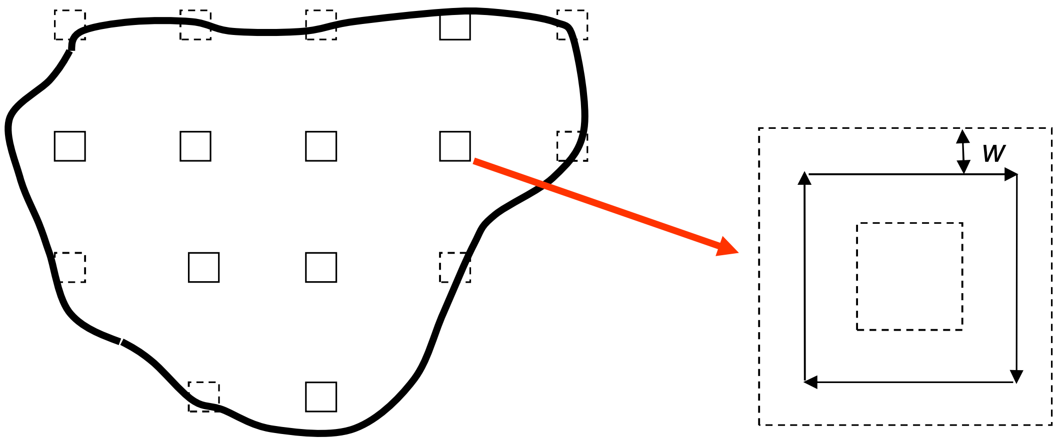
- Grid placed upon study area
- Transects centred upon grid points
- *"many short transects preferrable to few long transects"*

CREEM Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Segmented Designs

- Systematic segmented grid gives even coverage
- The between segment spacing should be the same in the x and y directions to maximise the number of sampling units
- Consider random orientation of lines, may give more uniform coverage
- Other designs (such as circuit samplers or segmented zigzags) could be used



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Sample size guidance

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Sample size

- Aim for at least 60-80 sightings for fitting the detection function
- and at least 20 lines or points for estimating encounter rate *n/L* or *n/k*
- Whether reliable estimates can be obtained from smaller samples is a matter of luck
  - depends on the data
  - do not employ a design that relies upon luck

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Sample size

More observations are required:

- if detection function is spiked
- if population is highly aggregated
- for point transect sampling

Increasing sample size using repeat counts

If a line is sampled three times,

- pool the distance data from the three visits
- enter survey effort as three times the line length.

If a point is sampled three times,

- enter survey effort as 3.

Determining total line length

Pilot study: n_0 animals (or clusters) counted from lines totalling L_0 in length.

Total line length required in main survey is

$$L = \left(\frac{b}{\left[cv_t(\hat{D}) \right]^2} \right) \times \frac{L_0}{n_0}$$

Where $cv_t(\hat{D})$ is the target cv (e.g. 10% is 0.1) and...

Determining line length

b is approximately $\dfrac{Var(n)}{n} + \dfrac{n \cdot Var[\hat{f}(0)]}{[\hat{f}(0)]^2}$

Pilot studies are typically too small to estimate *b*. If past similar data sets are not available, assume *b* = 3.

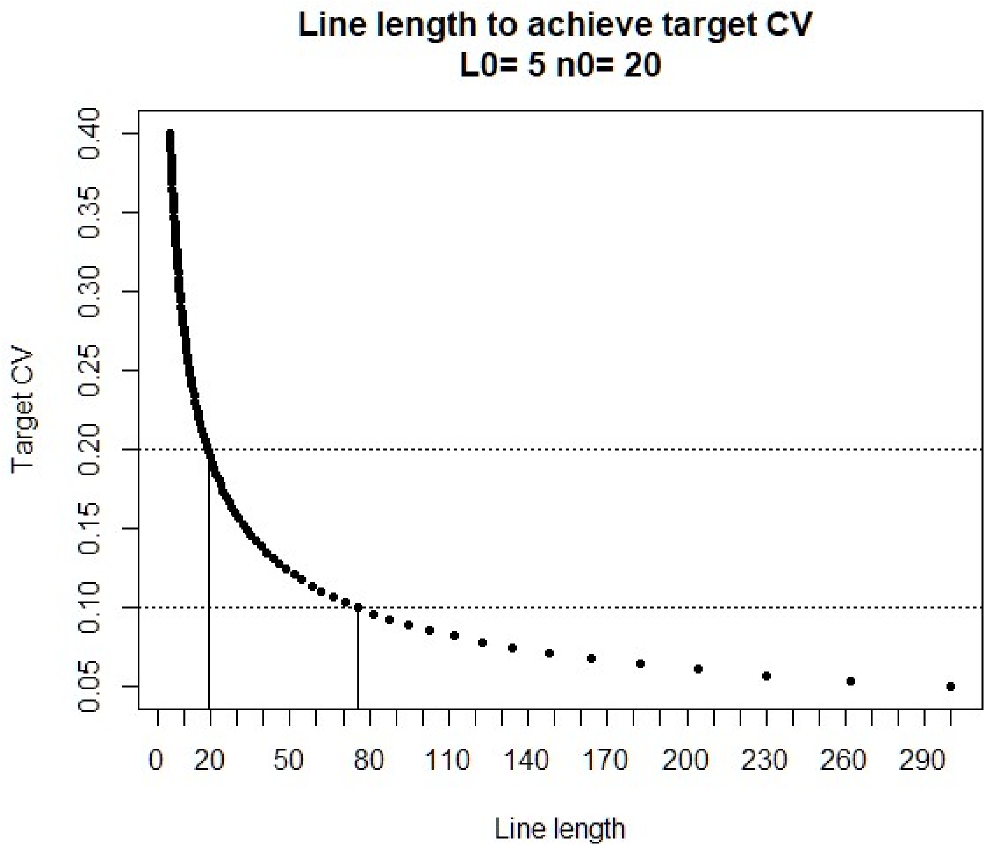
Line length example

A pilot study yields n_0 = 20 observations from lines of total length 5km. We require a CV of 10% and assume b = 3.

$$L = \frac{3}{0.1^2} \times \frac{5}{20} = 75\text{km}$$

Estimated sample size is

$$n = L \times \frac{n_0}{L_0} = 75 \times \frac{20}{5} = 300$$



Graph at right made using function `calculate.effort()` in `dssd` package

CREEM Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Determining line length

If pilot survey is sufficiently large, calculate line length for main survey as

$$L = \frac{L_0[cv(\widehat{D}_0)]^2}{[cv_t(\widehat{D})]^2}$$

where

$cv(\hat{D}_0)$ is the cv of estimated density obtained from the pilot survey, and L is total line length in the main survey

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Point transects: number of points

$$k = \left(\frac{b}{[cv_t(\hat{D})]^2}\right) \times \frac{k_0}{n_0}$$

or

$$k = \frac{k_0[cv(\hat{D}_0)]^2}{[cv_t(\hat{D})]^2}$$

where k_0 points in the pilot survey yielded n_0 detections, or estimated density of $\hat{D}_0$

CREEM Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Stratification

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Stratification (Geographic)

Why stratify?

- We might want estimates by sub-region/stratum
- To improve precision.
  - *Estimate inter-stratum differences rather than have them contribute to variance.*
  - *Reduce overall variance by increasing effort in strata which contribute most to variance.*
- For logistic reasons

Stratification (Geographic)

What to stratify?

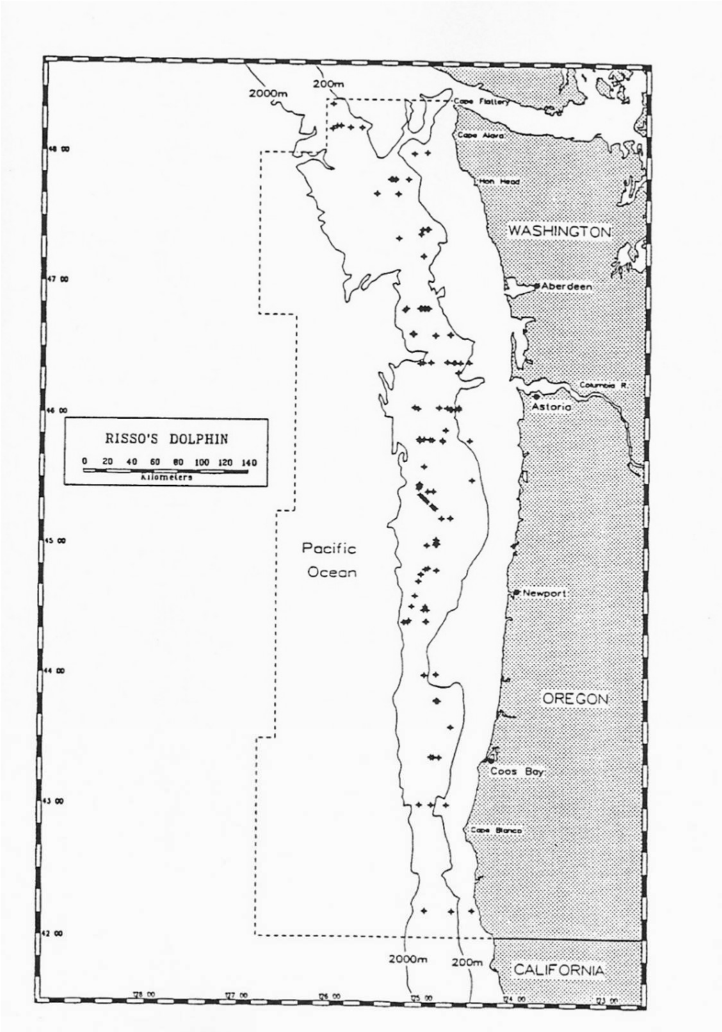
- Encounter rate: Density often varies spatially.
- Detection function: May vary spatially. There are often sample size limitations on stratified estimation (too few detections in some strata).

Caution:

- If either of the above are estimated by pooling across strata
    - when in reality they differ between strata
- within-stratum estimates are biased

Rexstad, E., Buckland, S., Marshall, L., & Borchers, D. (2023). Pooling robustness in distance sampling: Avoiding bias when there is unmodelled heterogeneity. *Ecology and Evolution*, 13, e9684. https://doi.org/10.1002/ece3.9684

Stratification (Spatial) – Risks!



- Most animals between 200m and 2000m contours, so put more effort into a shelf-edge stratum?

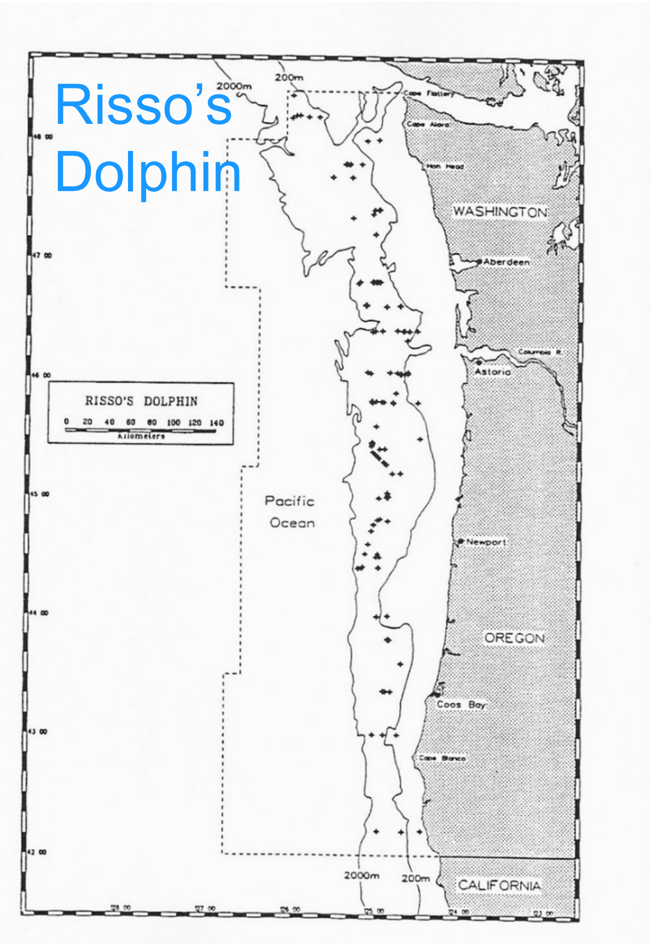
- What if our sample size too low in some strata?
  - With unequal coverage between strata pooling robustness is lost!
  - Our overall sample is no longer representative of the study area as a whole.

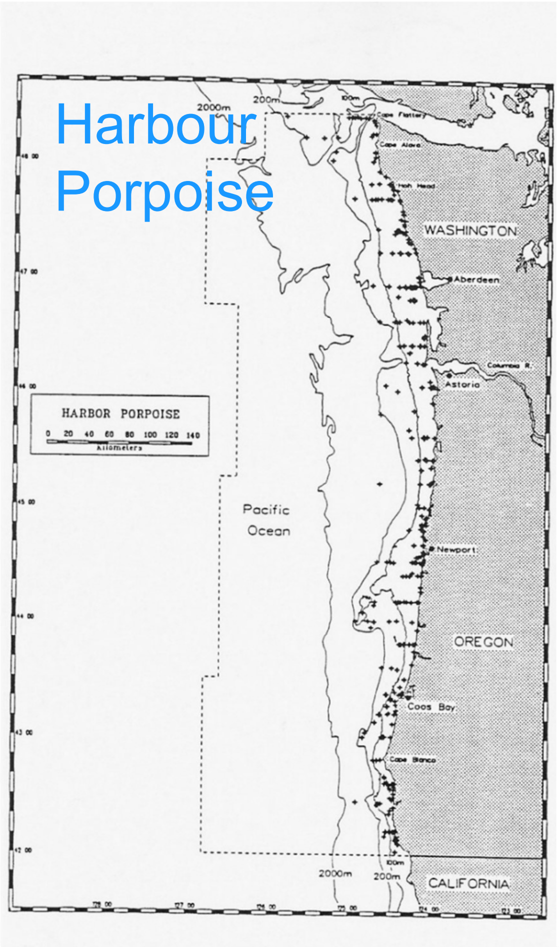
- Other species?

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Stratification (Spatial) – Risks!





Optimal effort location for one species may be poor for another species!

Uniform effort across strata is often the best design for multi-species surveys.

CREEM
Centre for Research into Ecological and Environmental Modelling

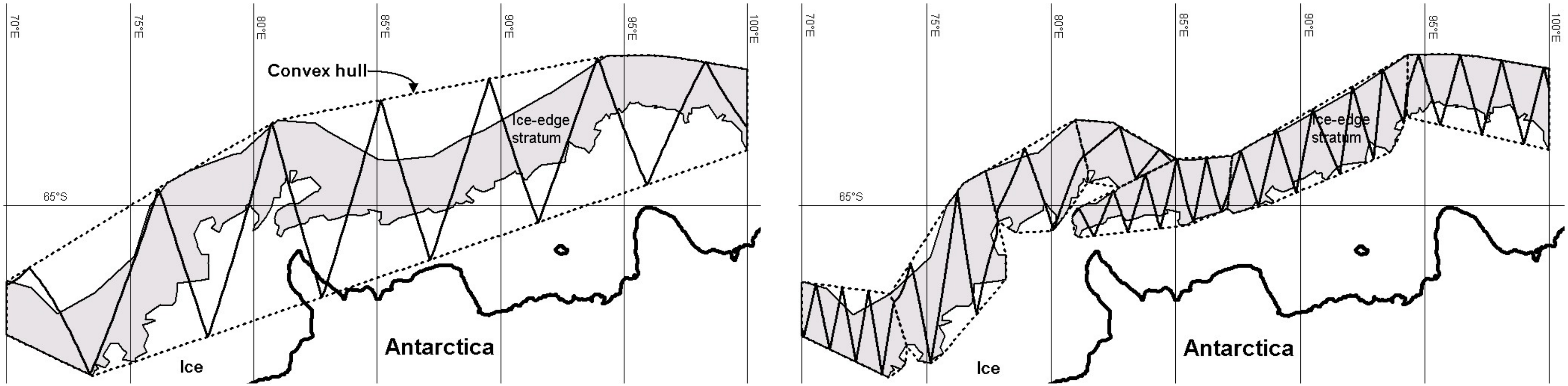
University of St Andrews

Example Surveys

Dealing with Complex Regions

Antarctic Minke whale shipboard survey

Study region divided into suitable strata to increase efficiency



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Main Points

- Line transects are generally preferable to points
- Try to achieve uniform coverage
- Systematic designs give more even coverage for any one survey
- Zig-zag designs often more efficient
- Lines should be placed parallel to density gradient (perpendicular to density contours) or to maximise the number of samplers
- Choose spacing values for points and segments which maximise sampling units
- Take care with unequal coverage stratified designs
  - *If coverage cannot be assumed equal, then it must be measured*
- Abundance/density estimation must involve the computed coverage
  - *Much more complex analysis*