Distance Sampling:
Survey Design and Simulation

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Outline

- Introduction
- Importance of Survey Design
  - Available software
- Concepts
  - Design versus Survey
  - Coverage Scores
- Designs for marine mammal surveys
- Stratification
- Simulation

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Introduction

- Good survey design and field methods are essential for obtaining the most accurate and precise estimates of density / abundance from distance sampling.
- Software:
  - Survey design engine in Distance 7.3 for Windows
  - Distance sampling survey design R library (dssd)
  - Distance sampling simulation R library
    - *DSsim – can be accessed from within Distance for Windows or directly in R (using pre-generated survey transects or very simple designs).*
    - *dsims – soon to be released new simulation library which uses survey design R library dssd.*

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Importance of Survey Design

- Good design is essential as distance sampling uses design based estimates
- It is extremely hard and often impossible to compensate for poor design at the analysis stage
- Good design (and field methods) makes analysis more straightforward
- Survey design must have a random component

Survey design – things to consider

- What are your objectives?
- What precision do you need?
- What resources are required?
- Are sufficient resources available?
- Include training in the costings.
- Cost for statistical advice!!
- Conduct a pilot survey.

Design versus Survey

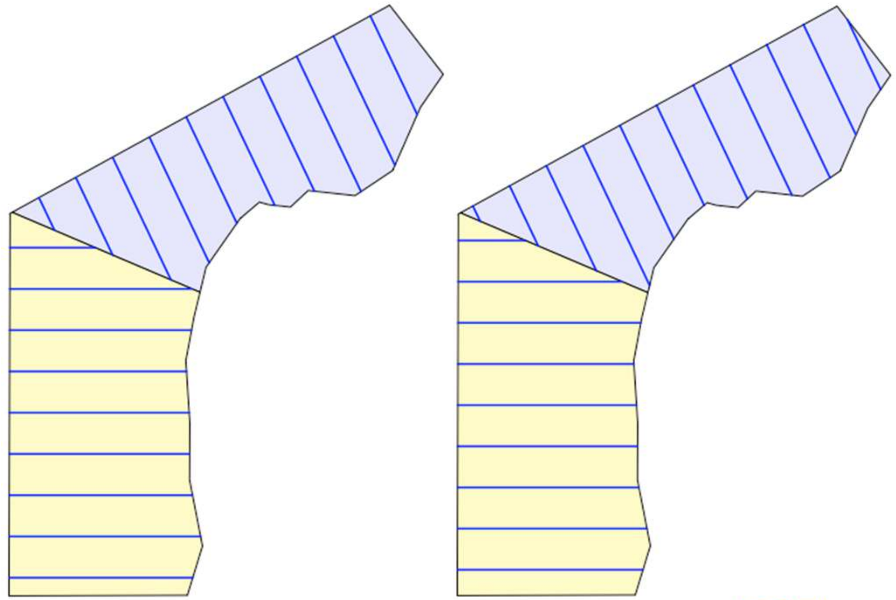
Design

A description of how the transects will be laid out.

E.g. 20 systematically spaced parallel lines with a random start point. The lines will have an angle of 155 degrees in the North stratum and 90 degrees in the South stratum.

Survey

A single set of transects randomly generated from the design.

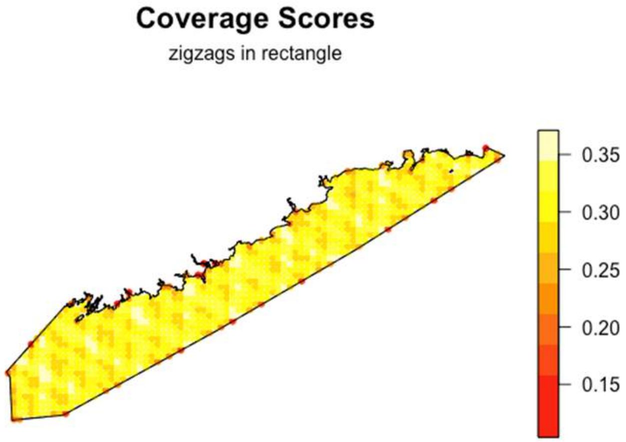


CREEM Centre for Research into Ecological and Environmental Modelling

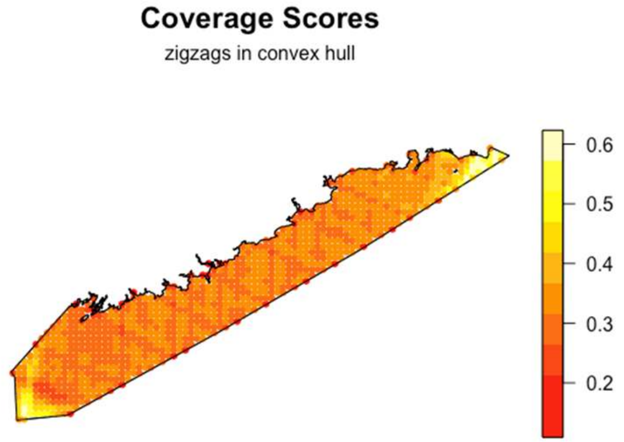
iversity of St Andrews

Coverage Scores

- The design must produce surveys which are representative of the study region.
- Want to be equally likely to sample any area
- If we are more likely to sample an area of high/low density then we will see positive/negative bias in our abundance estimates.
- Our software can help assess coverage scores.



Good



Not so good!

CREEM
Centre for Research into Ecological and Environmental Modelling

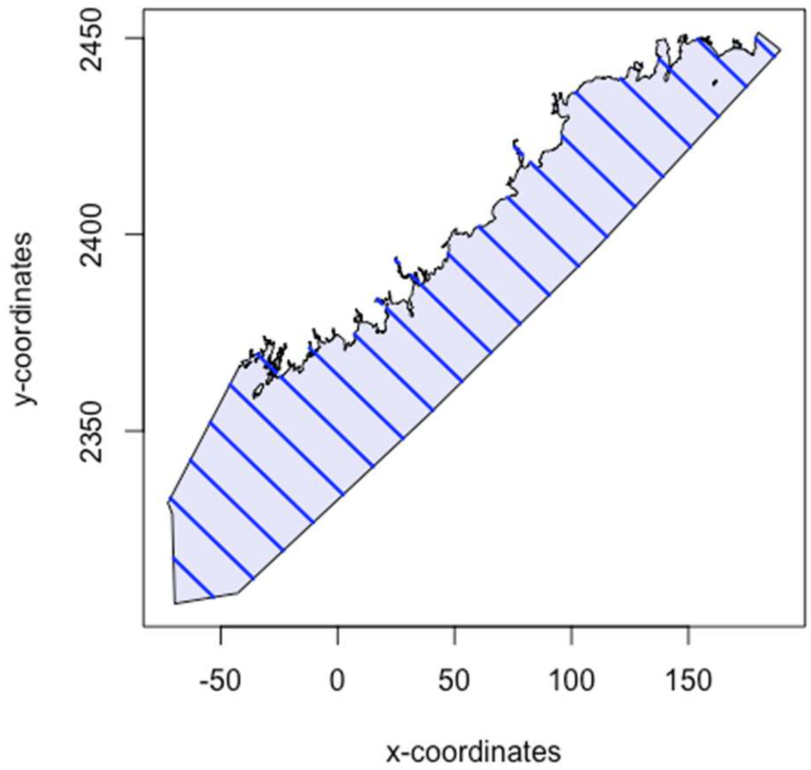
University of St Andrews

Survey Design: Marine Mammal Surveys

- Aerial surveys – often use systematic parallel lines with a random start point

- Shipboard surveys – often use zigzag design. We also offer a complementary zigzag design which can be more efficient when vessels need to start and finish in the same location.

Transects from a systematic parallel design

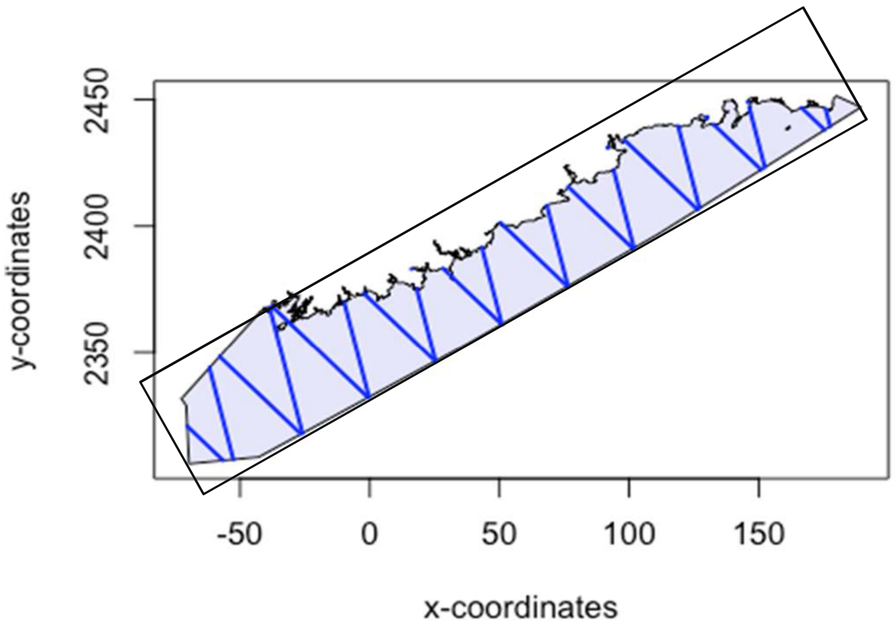


CREEM
Centre for Research into Ecological and Environmental Modelling

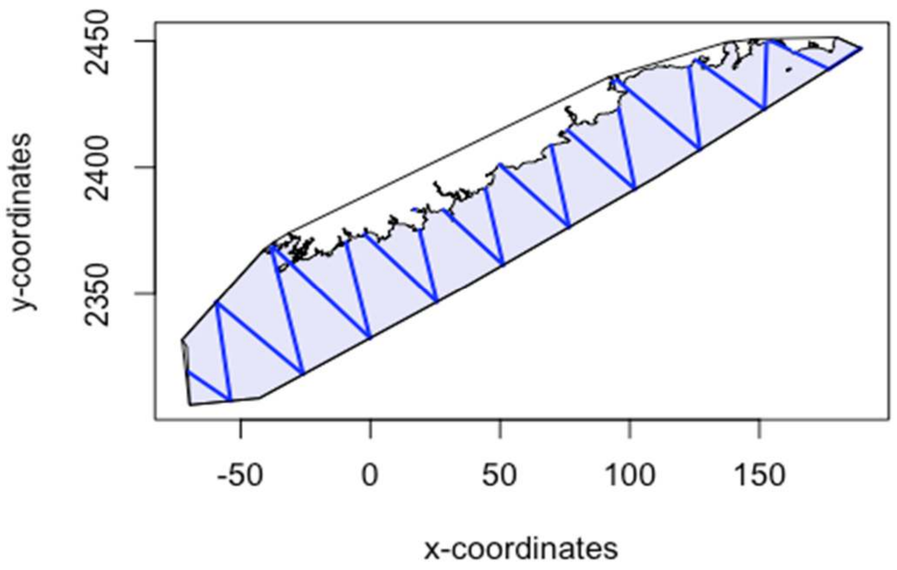
University of St Andrews

Equal Spaced Zigzag Designs

Generated inside a minimum bounding rectangle



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



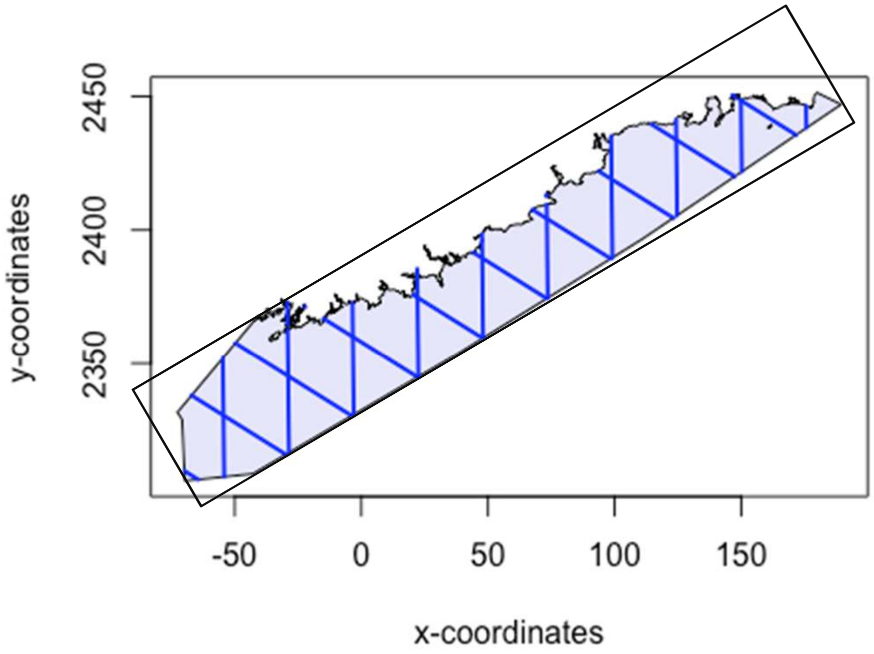
CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

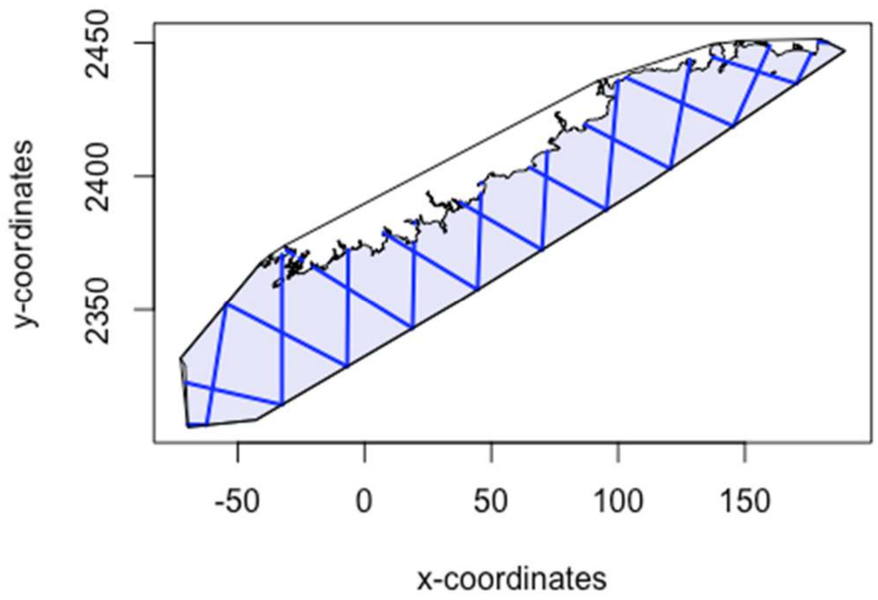
**New Design only available in dssd R library!*

Complementary Equal Spaced Zigzags

Generated inside a minimum bounding rectangle



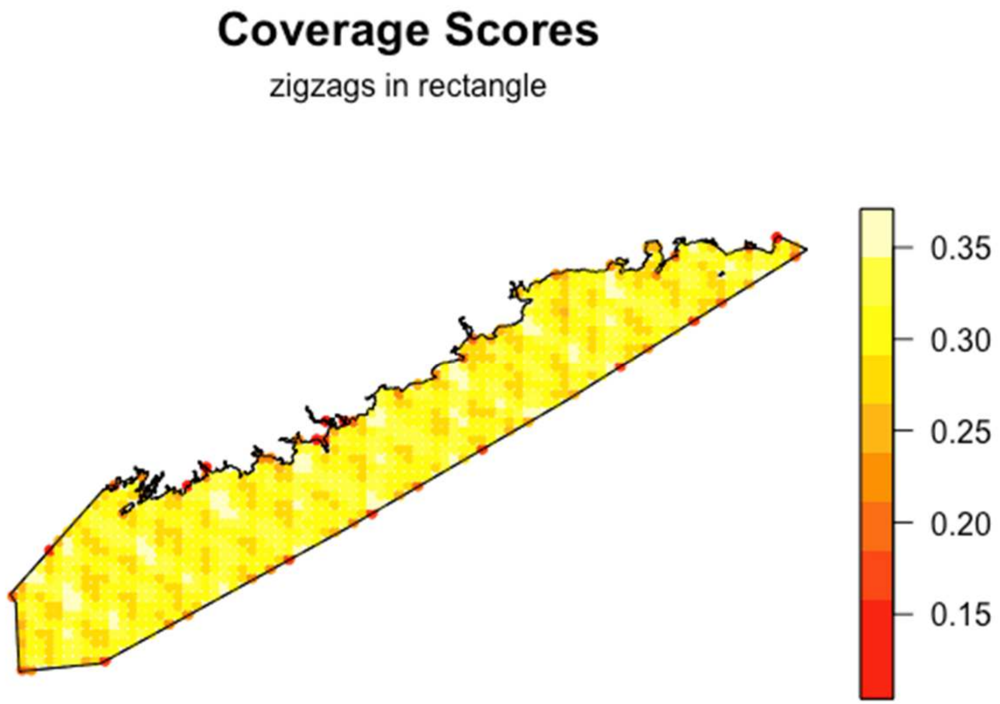
Generated inside a convex hull – like stretching an elastic band around the study region

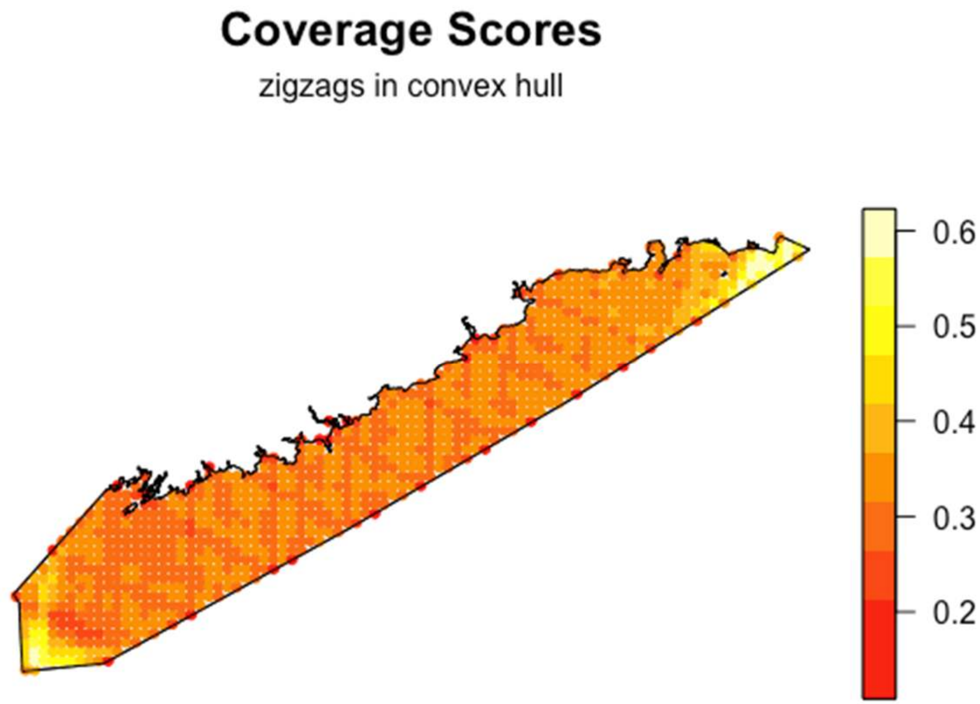


CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Complementary Equal Spaced Zigzags





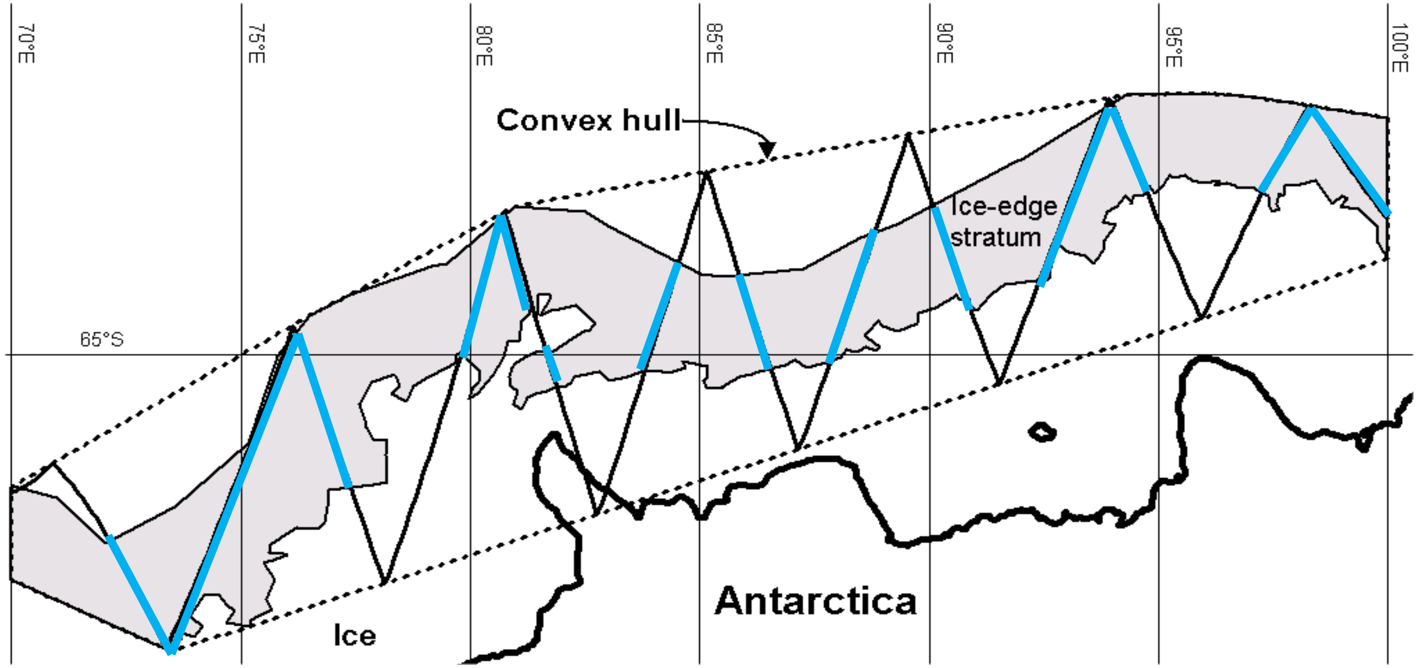
Coverage scores calculated by generating 1000 surveys

CREEM Centre for Research into Ecological and Environmental Modelling

University of St Andrews

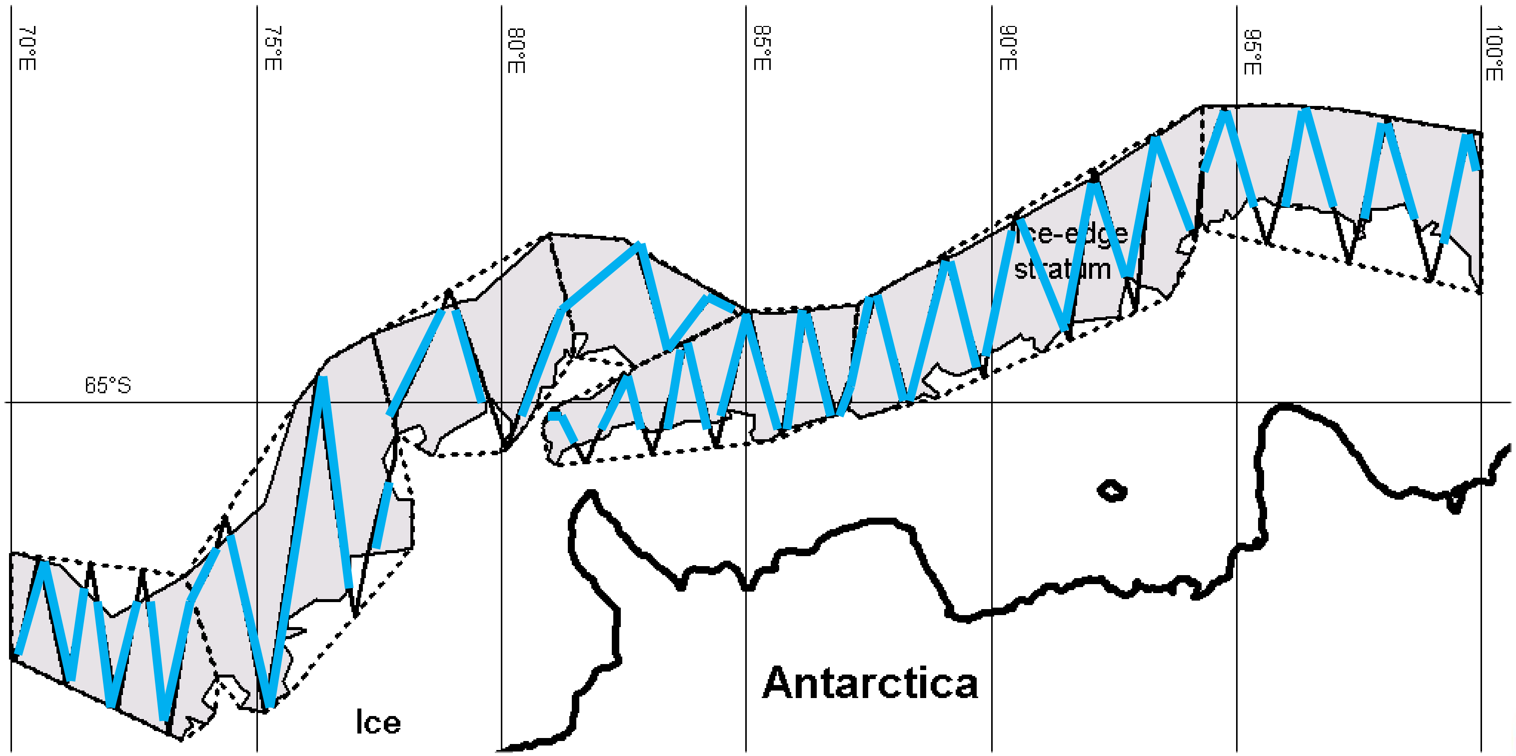
Complex Regions – poor efficiency

Antarctic Minke whale shipboard survey: shaded area is study region, transects have been generated inside a convex hull and then they will be clipped to the study region – blue lines.



Complex Regions – improved efficiency

Study region divided into a set of more convex shapes to increase efficiency



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Stratification

Why stratify?

- For efficiency / logistic reasons
- We want abundance / density 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.*

Stratification

What to stratify?

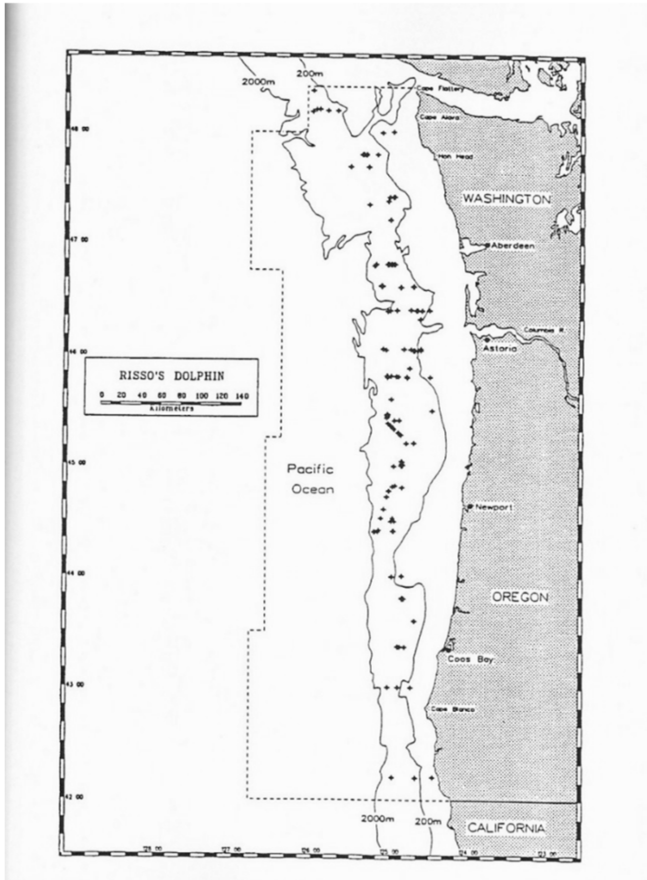
- Encounter rate: Density often varies spatially.
- Detection function: May vary spatially. There are often sample size limitations on stratified estimation (too few detections in some strata).
- Mean cluster size: May vary spatially. There may be sample size limitations on stratified estimation.

NB: Careful! If any of the above are estimated by pooling across strata, when in reality they differ between strata, within-stratum estimates will be biased.

CREEM
Centre for Research into Ecological and Environmental Modelling

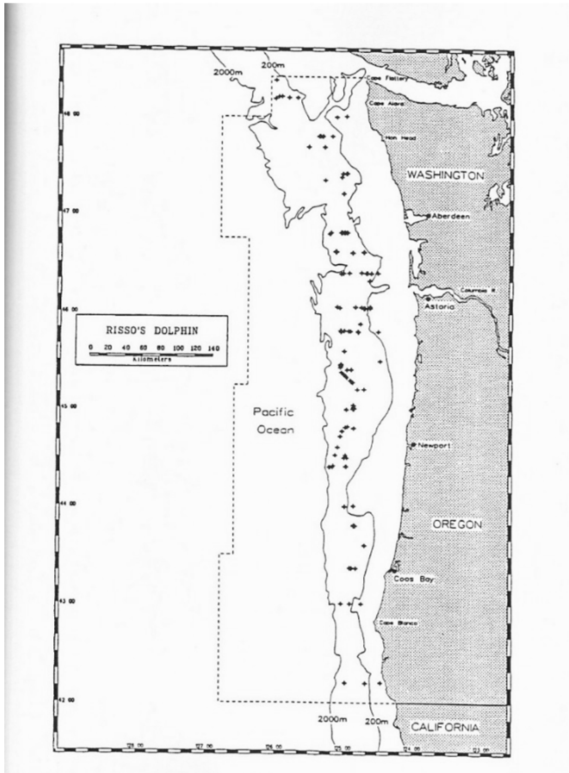
University of St Andrews

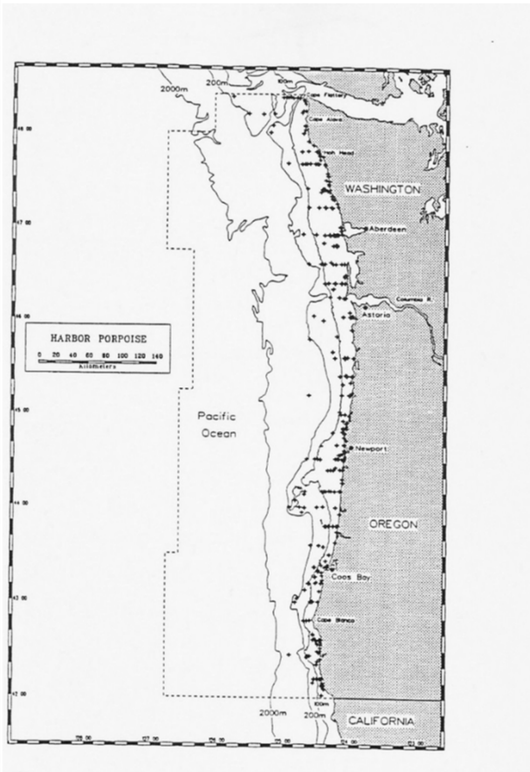
Spatial Stratification



- Most animals between 200m and 2000m contours, so put more effort into a shelf-edge stratum?
- What if our sample size is too low in some strata?
  - With unequal coverage between strata our overall sample might not be representative of the entire study area so need to be careful when pooling detections!
- Other species?

Spatial Stratification





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

- Uniform effort across strata is often a good design for multi-species surveys.

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

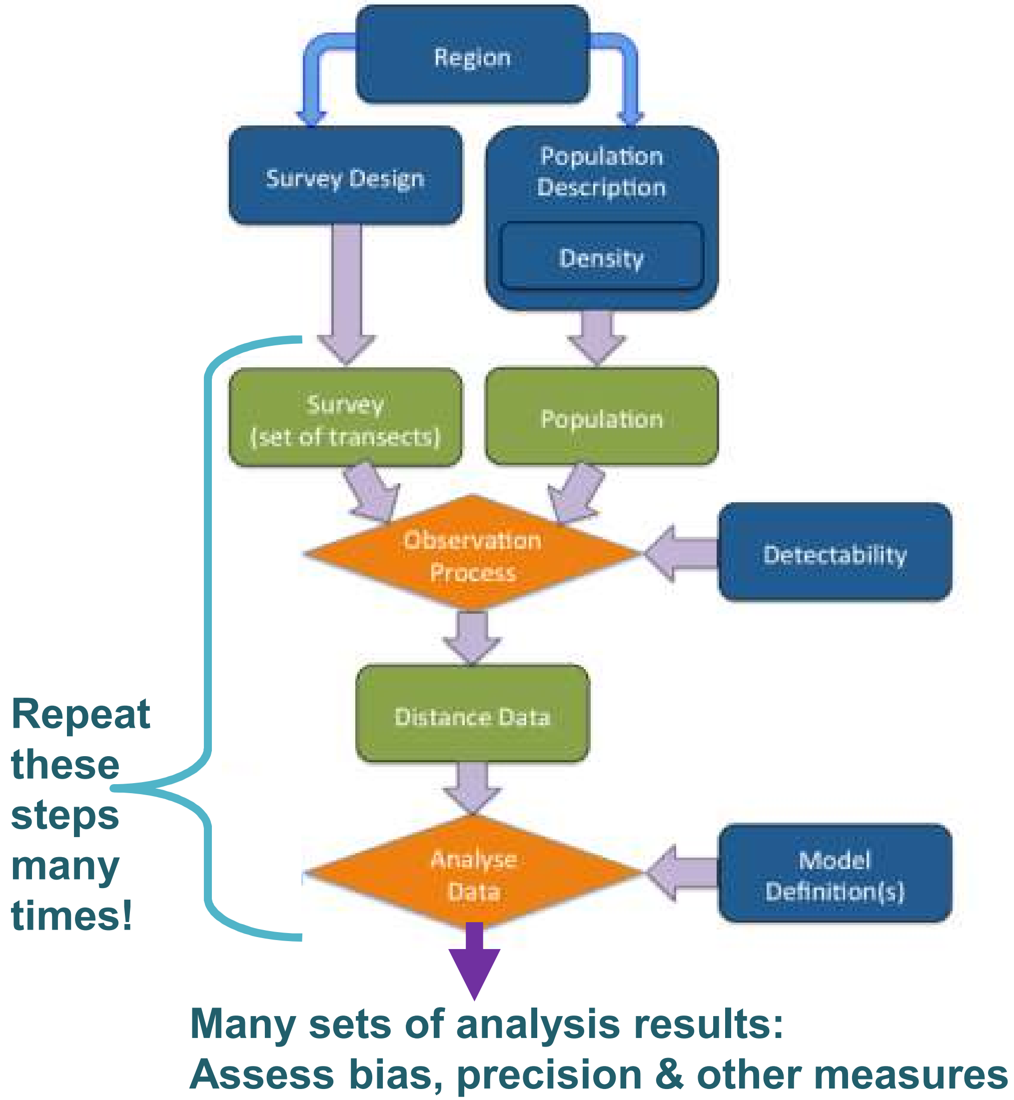
Simulation

- In the past researchers have usually relied on general rules of thumb.

- Simulation allows researchers to tailor survey design to their specific study and therefore help achieve the most accurate and precise estimates.

- Cheap to simulate – expensive to survey!

- Available in Distance 7.4 and in R



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

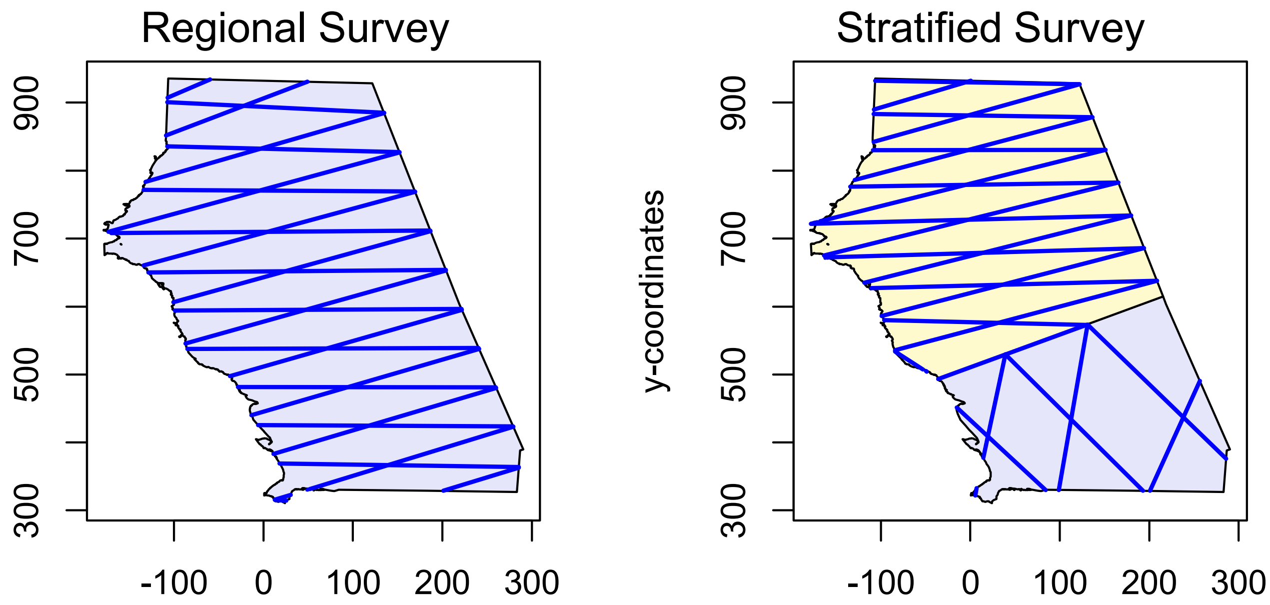
Simulation Example

- Let's plan a Minke whale survey in the North Sea
- Is it sensible to use the Small Cetaceans in the European Atlantic and North Seas (SCANS) survey to allocate more effort where we expect more animals?
  - There are more animals in the north so put more effort here?

Survey region over-layed on 1994 predicted density of Minke whales.

Simulation Example - design

- Complementary zigzag designs which are comparable in cost.
- Regional design has 60 km spacing and 13.3% average coverage
- Stratified design has a spacing of 50 km in the northern strata and 100km in southern strata giving coverage of 15.8% and 8.4%, respectively.

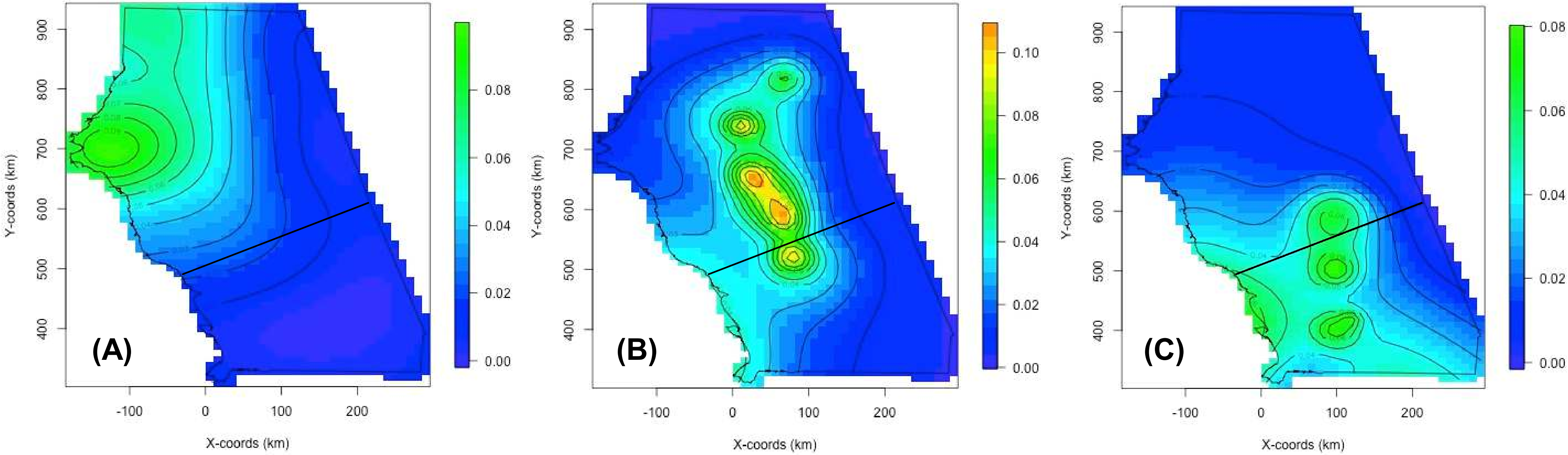


CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Simulation Example – animal density scenarios

3 test scenarios for animal distribution: (A) distribution similar to that in 1994 survey (best case), (B) shifted to a more central distribution as seen in 2005 survey, (C) distribution shifted further south (represents assumed worst case scenario given stratified design)



Simulation Results (Truth: N = 4550)

Scenario A: 1994 density surface

	Regional Design	Stratified Design
Mean $\hat{N}$ (%bias)	4599 (1.07%)	4549 (-0.03%)
CI coverage	99%	96%
Mean $\widehat{se}$	907	659
RMSE	653	609

Scenario B: 2005 density surface

	Regional Design	Stratified Design
Mean $\hat{N}$ (%bias)	4666 (2.56%)	4629 (1.74%)
CI coverage	97%	98%
Mean $\widehat{se}$	759	783
RMSE	665	682

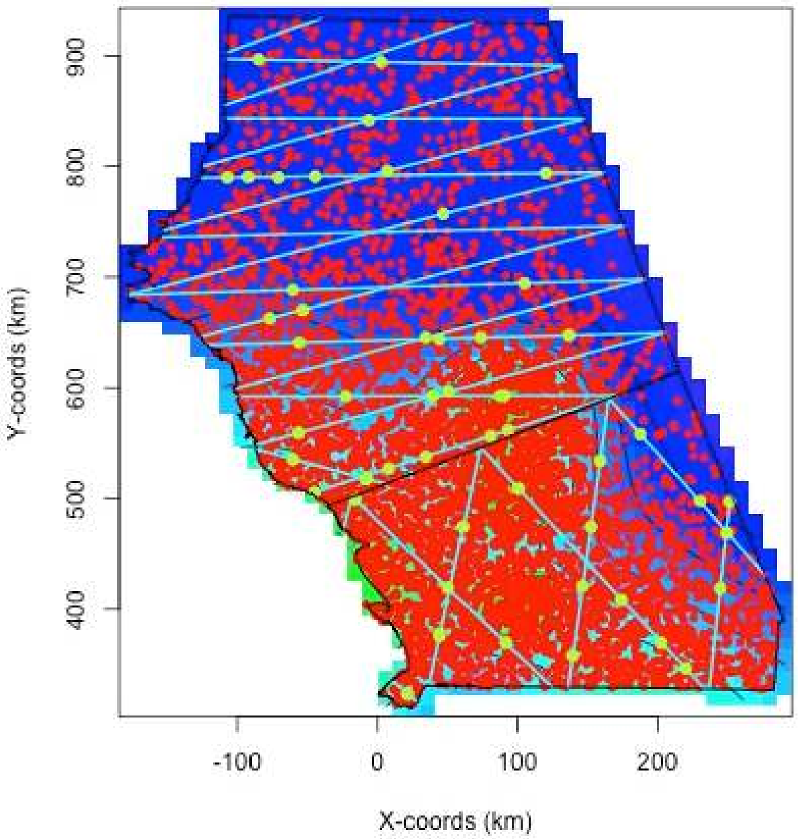
Scenario C: worst case scenario

	Regional Design	Stratified Design
Mean $\hat{N}$ (%bias)	4606 (1.23%)	4637 (1.90%)
CI coverage	99%	97%
Mean $\widehat{se}$	851	862
RMSE	629	763

Discussion

- Seems like there may be small improvements in precision using the stratified design if the distribution of animals corresponds to scenario A (but maybe not enough to justify added complexity).
- Doesn't seem to be much to lose using the stratified design in scenario B or even in our worst case scenario C.

- NOTE: These results are specific to the parameters specified in this survey! Various assumptions were made about the population and detectability.



Example simulated survey from scenario C. Density surface overlaid with population (red dots) transects (blue lines) and sightings (yellow/green dots)

CREEM Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Other Simulation Suggestions

- Use as a tool to justify a change in survey design – simulate from current design and demonstrate how it can be improved.
- Use it to test how badly biased estimates might be when you have non-uniform coverage – for example using zigzag designs in a convex hull or when you suspect edge effects are an issue.
- We have used it to demonstrate that collecting binned data accurately is much better than collecting exact distances inaccurately… even when there are only 3 distance bins!

…

- Let us know what simulations would help you design your surveys!

Further Resources

- Online introductory workshop: https://workshops.distancesampling.org/online-course/index.html

- List of books and other resources: https://workshops.distancesampling.org/online-course/more-resources/

- Vignettes (code examples using the design and simulation packages):
  - DSsim: https://github.com/DistanceDevelopment/DSsim/wiki
  - dssd: inside the R package, type `browseVignettes("dssd")` in R.

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews