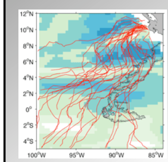
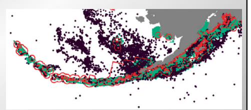
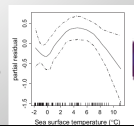
Ecological modelling – the issues

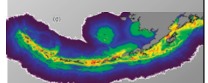
Examining patterns of marine mammal distributions to insights into ecological processes



Edward Gregr
Ecological modelling workshop
Quebec City
October 11, 2009





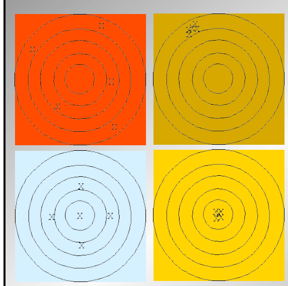
Why model?

- Models are how we understand the world
  - We see our world through implicit cognitive models
  - We learn about our world using formal descriptions
- For marine mammals
  - What?
    - Distribution, abundance, ecosystem role, stock structure
  - Why?
    - Ecology, conservation, competition studies, critical habitat, EBM, MPAs, climate change

Challenges to effective modelling

- Complexity
  - Accuracy, precision & uncertainty
- Data sources and limitations
- Space and scale
- Model evaluation & testing

Accuracy, precision, & uncertainty



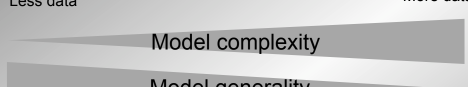
Uncertainty:
- Parameter estimation
- Observational
- Design
- Stochasticity

Complexity
Data
Scaling
Evaluation

Complexity

Few variables
Less data

More variables
More data



Biological linkages assumed | Biological linkages modelled

Increased accuracy | Increased precision

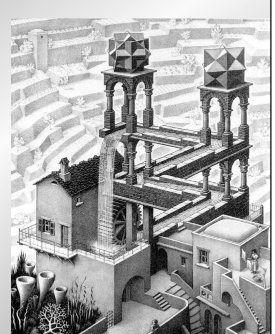
Low uncertainty | High uncertainty

Complexity can increase in a number of ways, including biologically, spatially, and temporally

Complexity
Data
Scaling
Evaluation

Does model complexity guarantee a better model?

Not necessarily …

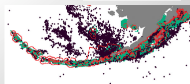


Complexity
Data
Scaling
Evaluation

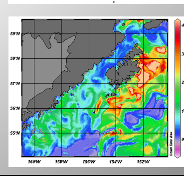
Complexity
Data
Scaling
Evaluation

Data sources

Biological
- Systematic surveys, tagging
- Platforms of opportunity, historic observations, catch data



Physical
- Field sampling, remote sensing, circulation models, floats



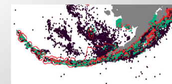
For most applications, need continuous predictions over large spatial extents

---

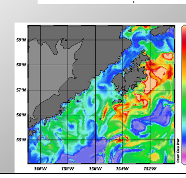
Complexity
Data
Scaling
Evaluation

Data considerations

- Field data
  - Limited in time and space
  - Effort bias?
- Opportunistic data
  - Always effort biased
  - Absences rarely recorded



- Defining a sampling unit
  - Mixed resolution data sets
  - Gridded data
  - Autocorrelation
  - Colinearity

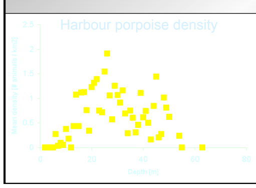


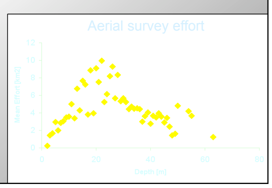
Complexity
Data
Scaling
Evaluation

Effort bias

- Work concentrated on-shelf
- Species are often cosmopolitan in distribution

➢ Does sampling cover a species' range?

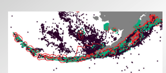




Complexity
Data
Scaling
Evaluation

Getting physical with biology

- Systematic data
  - Presence-absence
  - Standard correlative methods (e.g., regression)
- Opportunistic data
  - Presence-only
    - Envelope models (e.g. BioClim)
    - Machine learning (e.g., MaxEnt)
    - Niche breadth (e.g., ENFA)



Complexity
Data
Scaling
Evaluation

Defining a sampling unit

- Spatio-temporal resolution
- Depends on data
- Depends on the question
  - Pattern or process?

Complexity
Data
Scaling
Evaluation

Gridding your data
(raster aspects of scale)

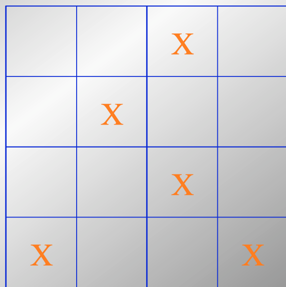
- Extents
  - Influence of study area size on analysis
    - Prediction → extrapolation?
    - Prevalence → presence-only analyses
- Resolution
  - Autocorrelation
  - Influences summary results

Effects of grid cell size …

Complexity
Data
Scaling
Evaluation

Pr(x) = 5/16 = 0.3125

Pr(x) = 4/4 = 1.0



Sinclair (2007): Area impacted by trawl:
30,000 km^2 at 5x5 km^2
9,000 km^2 at 1x1 km^2

Evaluation and testing

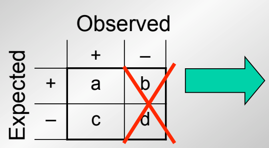
Complexity
Data
Scaling
Evaluation

- Most important
- Frequently ignored
- Recognise that not all models need same level of validation (Rykiel 1996)
- Sensitivity analysis can increase confidence in model accuracy

Model evaluation

Complexity
Data
Scaling
Evaluation

- Correlation studies
  - Statistical assumptions regularly violated
- Performance based on contingency table



- Chi square
- ROC plot
- Kappa statistic

Presence-only data contain no true (i.e., observed) absences.

Possible solutions

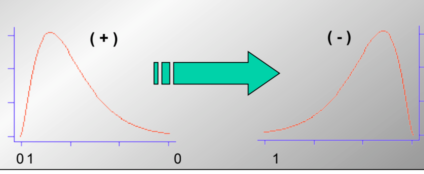
Complexity
Data
Scaling
Evaluation

- Pseudo-absence data
  - Assumes no bias in presence sampling
  - Influenced by extent of study
- Null model comparisons
- Skewness test
  - Let the presence data tell you what is best

Skewness

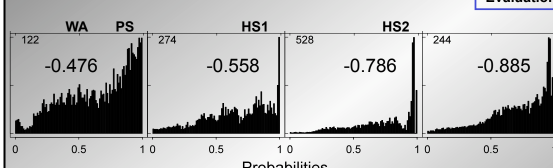
Complexity
Data
Scaling
Evaluation

Assumption:
- A better model gives higher probabilities at "presence" locations
- i.e., the distribution of probabilities at observations will be more negatively skewed



Model comparison

Complexity
Data
Scaling
Evaluation



WA – Winter accessibility
PS – Population-based suitability
HS1 – Partial habitat suitability model
HS2 – Full habitat suitability model

Gregr and Trites 2008. *Marine Ecology Progress Series*

Two warnings …

- Spatial models are pattern descriptions. Describing patterns is potentially risky (just ask stock assessment).
- Sample unit definition requires data pooling. Pooling creates biases in data that can lead to unexpected results.

Modelling tips

- Ask a clear question
- Add complexity only where necessary
- Judiciously ignore available data
- Ensure transparency
  - in purpose
  - in relationships between inputs and outputs
- Document assumptions and limitations
- Pay attention to sensitivity and validation
- Remember that all models are wrong
- Terrestrial literature is informative, but not always useful