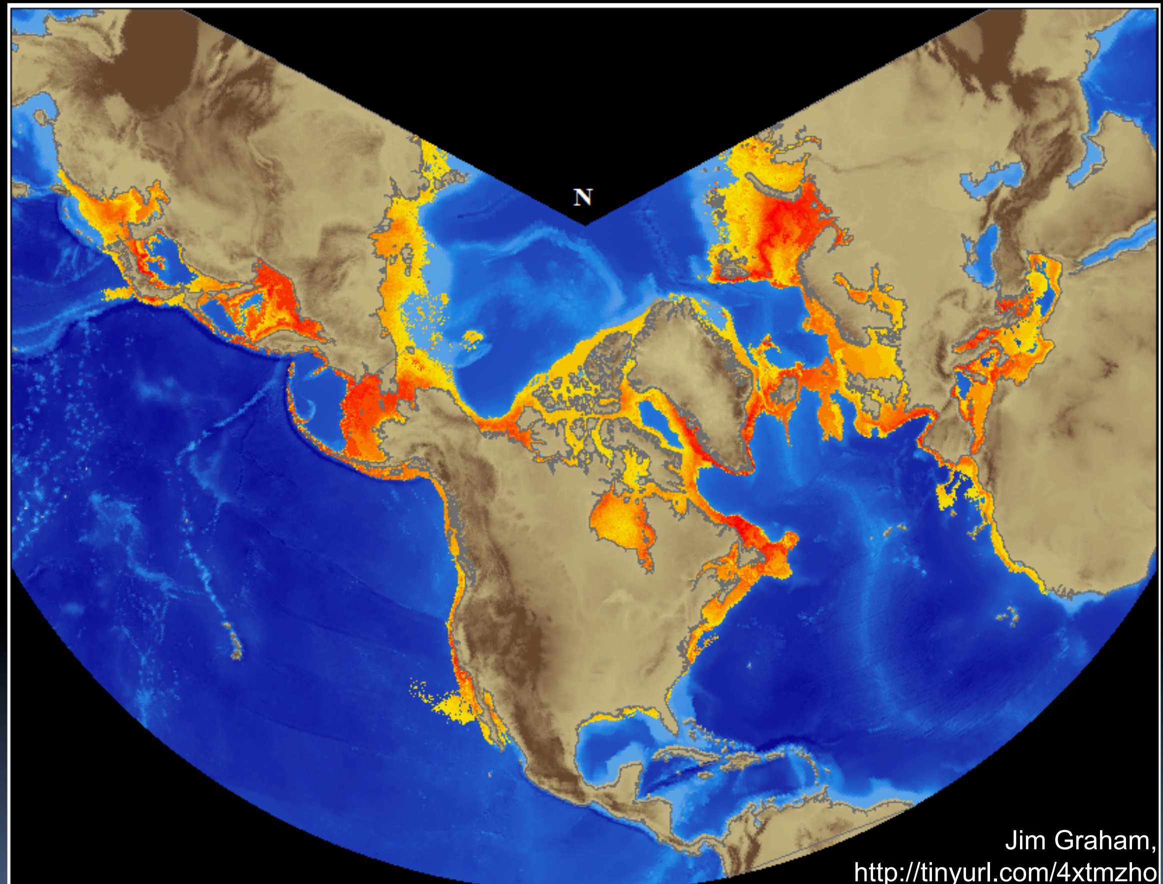
Introduction to Ecological Modeling

Dori Dick
Oregon State University, USA
Elliott Hazen
NOAA Pacific Fisheries
Environmental Laboratory, USA



Jim Graham,
http://tinyurl.com/4xtmzho

Ecological Modeling Workshop
19th Biennial Conference on the Biology of Marine Mammals
Tampa, FL 26 November 2011

Why model?

- Abstractions of real world system or process
- Help to define problems more precisely and concepts more clearly
- Provide approach for analyzing data, providing statistical inference, communicating results
- Allow for predictions

Models and Marine Mammals

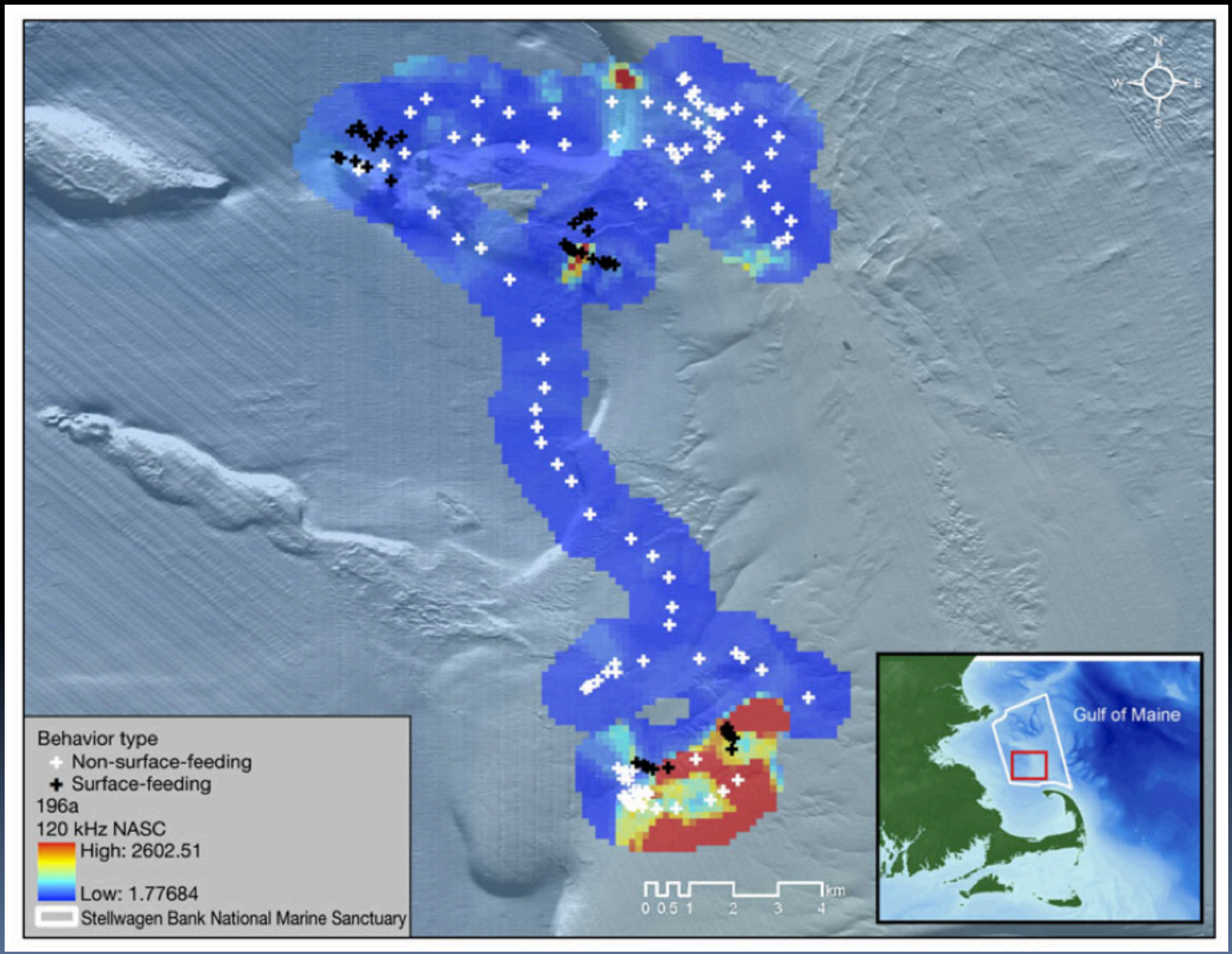
What being modeled?

- Distribution
- Abundance
- Habitat requirements
- Ecosystem role
- Stock Structure
- Genetic population structure

WHY?

Ecology

Humpback foraging in Gulf of Maine

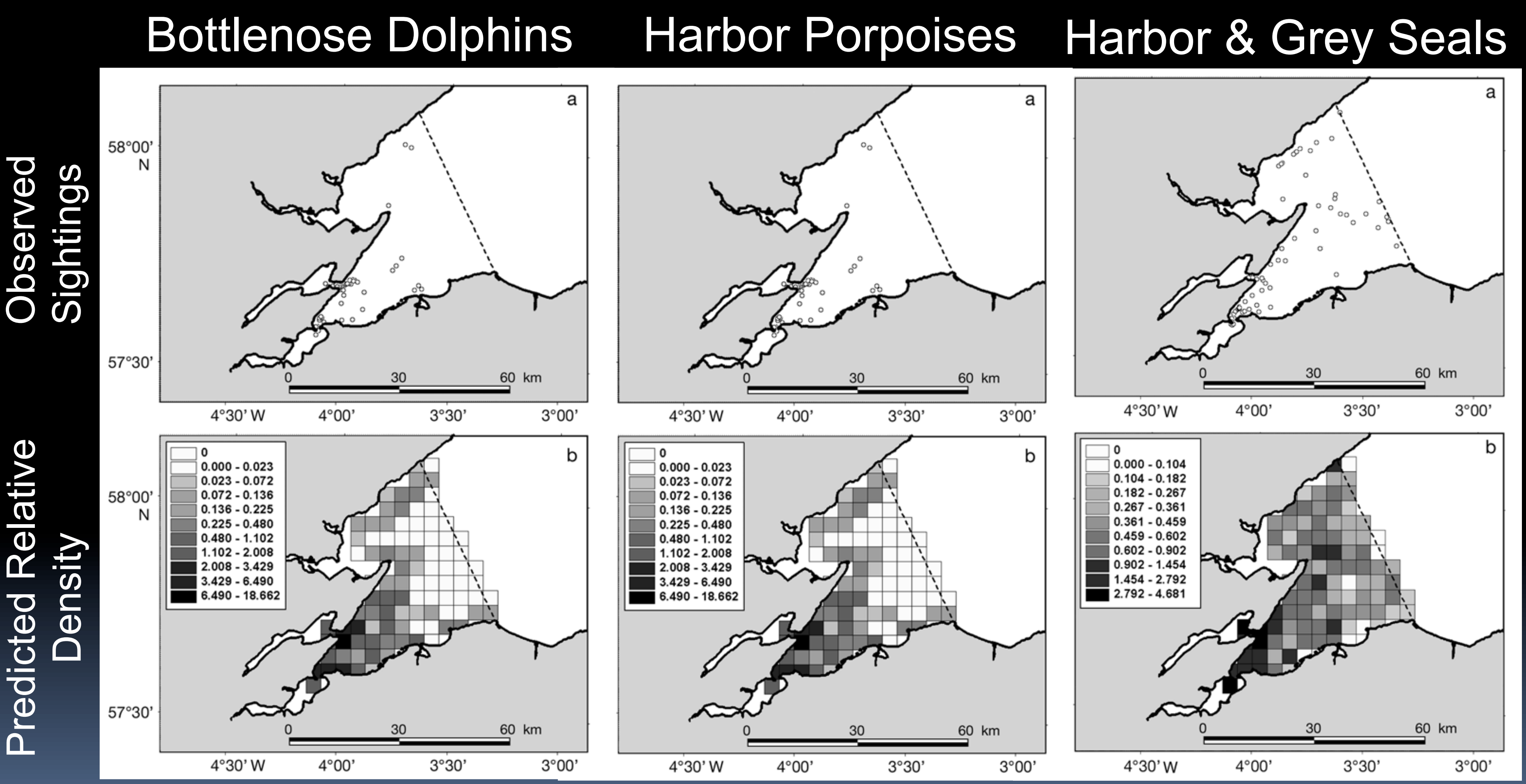


- Generalized Additive Mixed Models
- Classification regression trees
- Generalized linear models
- Modeled whale movement and feeding in relation to their prey

Hazen, E. L., Friedlander, A. S., Thompson, M. A., Ware, C. R., Weinrich, M. T., Halpin, P. N., and Wiley, D. N. 2009. Fine-scale prey aggregations for foraging ecology of humpback whales *Megaptera novaeangliae*. Marine Ecology Progress Series. 395:75-89.

Conservation

Habitat modeling to id conservation zones within MPAs



Bailey, H. and Thompson, P. M. 2009. Using marine mammal habitat modelling to identify priority conservation zones within a marine protected area. Marine Ecology Progress Series. 378:279-287.

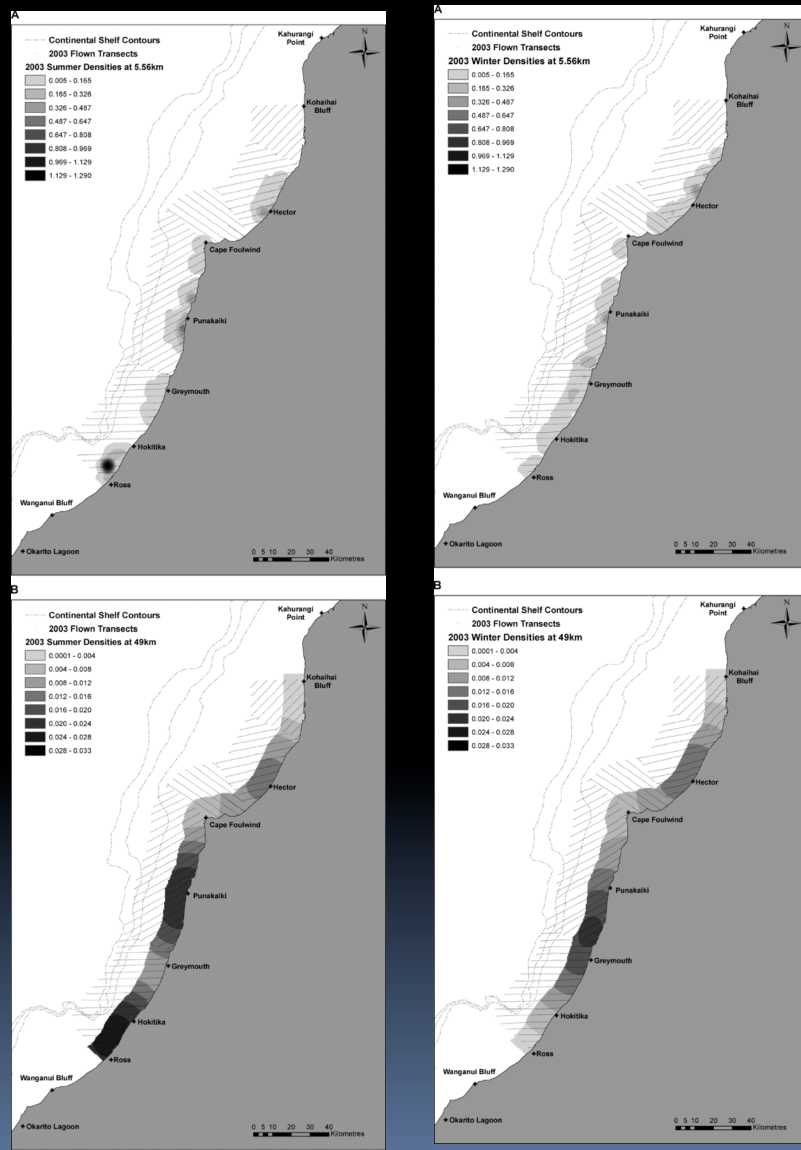
Conservation

Distribution of Hector's dolphins, implications for conservation

2003, Summer | 2003, Winter

Search Radius 5.56 km

Search Radius 49 km

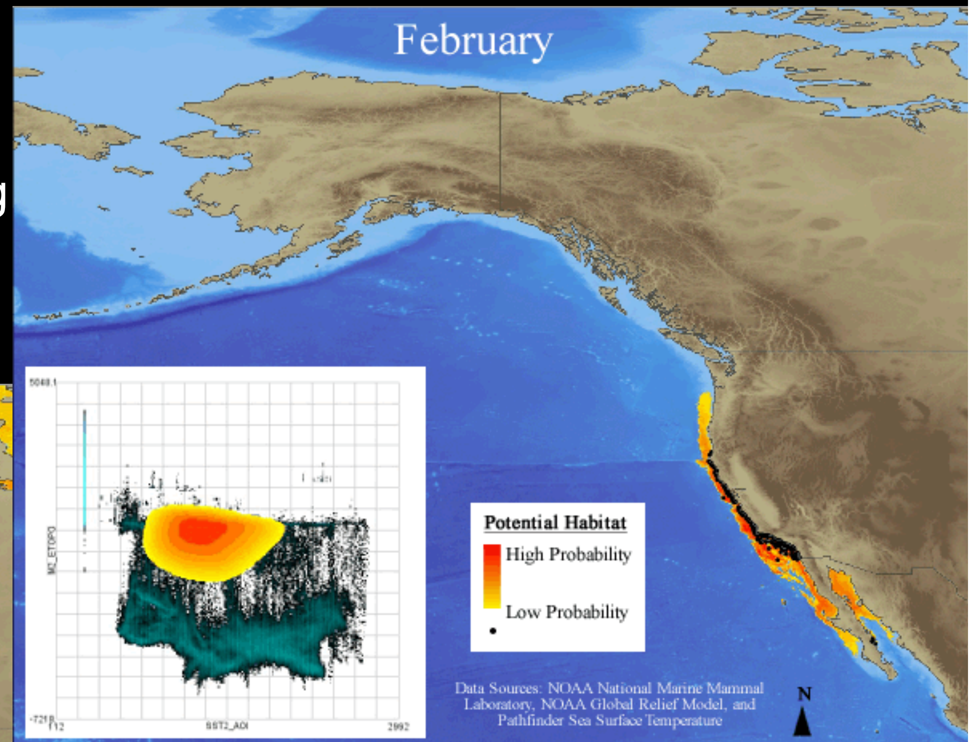


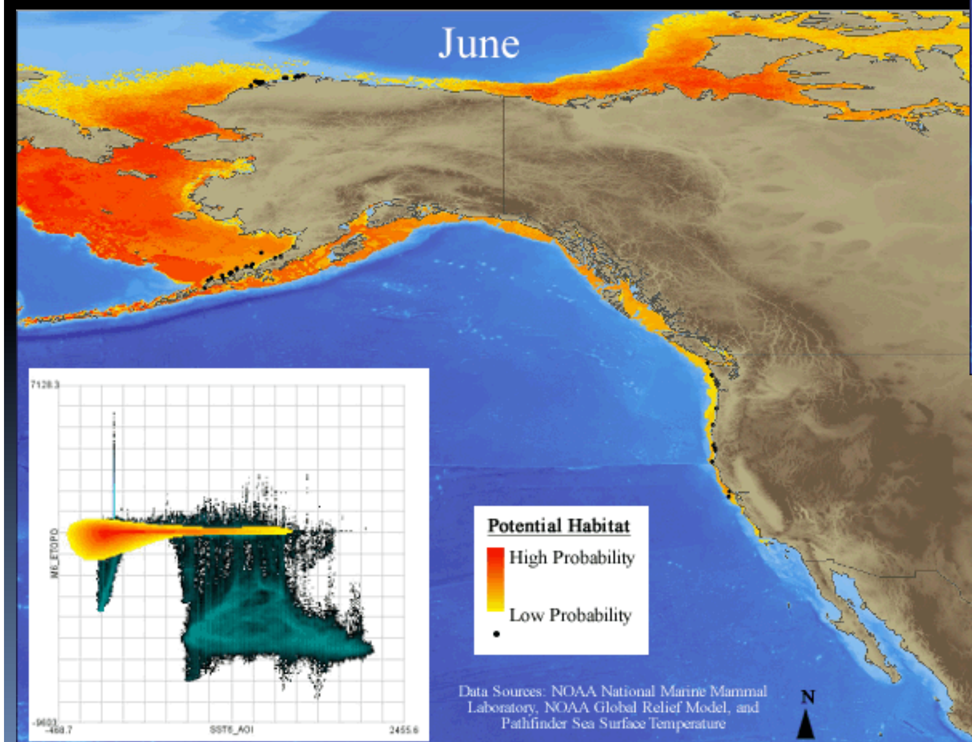
Rayment, W., Clement, D., Dawson, S., Slooten, E. and Secchi, E. 2011. Distribution of Hector's dolphin (*Cephalorhynchus hectori*) off the west coast, South Island, New Zealand, with implications for the management of bycatch. Marine Mammal Science, 27: 398–420.

Habitat Suitability/Critical Habitat

Potential habitat of gray whales throughout the Northern Hemisphere

- Developed new Hyper Envelope Modeling Interface (HEMI)





- Incorporates ecological niche theory into habitat suitability modeling using Bezier functions, creates niche envelopes
- Takes into account biological variables, allows explicit visualization/ modification of species' environmental niche

Jim Graham, http://tinyurl.com/4xtmzho

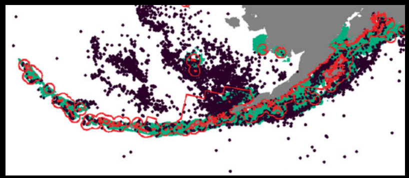
Modeling Challenges

- Data sources
  - → Biological and physical
- Accuracy, precision, uncertainty, sample size
- Scale
  - → Grain, extent
- Complexity
  - → Increases with more variables, amount of data, more general model
- Spatial vs. temporal
- Evaluation and testing
  - → assumptions, transformations, autocorrelation

Data Types

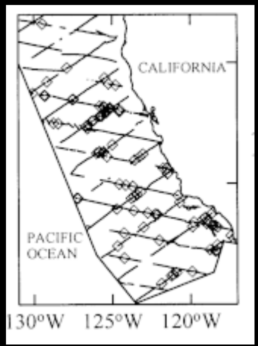
Points

animal locations, sample location



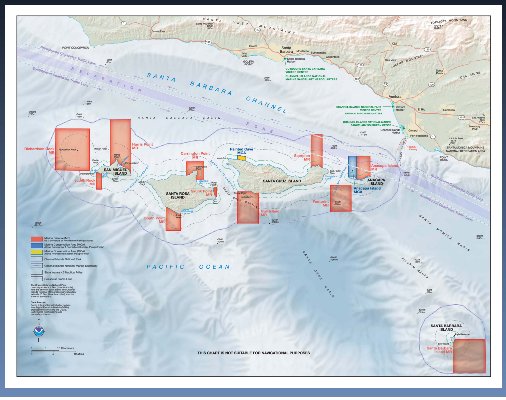
Lines

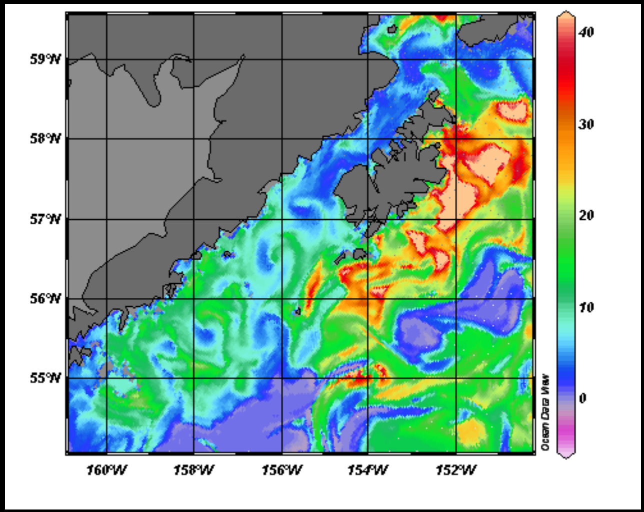
ship track, shoreline



Polygons

MPAs, oil spill, harmful algal bloom





Surfaces

bathymetry, remotely sensed SST, chlorophyll concentration

Time (or location) series

weather buoy, satellite telemetry

Data Sources

Biological

Surveys – systematic line transects (e.g. distance sampling)

Tagging - movement models (e.g. particle filters, state space models)

Platforms of opportunity (e.g. generalized additive models)

Genetic, Historical observations, Catch data

Physical

in situ variables (depth, salinity), remote sensing (SST, chlorophyll concentration, derived frontal features), circulation models

For most applications, need continuous predictions over large spatial extents

Biological Field Data

Systematic

→ Get presence-absence data

→ Correlation methods (regression), abundance/density estimates (Distance sampling), Mantel tests

Examples

Rayment, W. et al. 2011. Distribution of Hector's dolphin (C*ephalorhynchus hectori*) off the west coast, South Island, New Zealand, with implications for the management of bycatch. *Marine Mammal Science*. 27: 398–420.

Herr, H. et al. 2009. Seals at sea: modelling seal distribution in the German bight based on aerial survey data. *Marine Biology*. 156:811-820.

Biological Field Data

Opportunistic

→ Presence only

→Always effort biased

→ Remember that just because animal is absent does not mean it was not there, could have missed it

Examples

Williams, R. et al. 2006. Modeling distribution and abundance of Antarctic baleen whales using ships of opportunity. *Ecology and Society* **11**(1): 1. [online] URL: http://www.ecologyandsociety.org/vol11/iss1/art1/

Cotté, C. et al. 2009. Scale-dependent habitat use by a large free-ranging predator, the Mediterranean fin whale, Deep Sea Research Part I: Oceanographic Research Papers. 56(5):801-811.

- MaxEnt (http://www.cs.princeton.edu/~schapire/maxent/)
- Envelope models (BioClim (http://ecobas.org/www-server/rem/mdb/bioclim.html)
- AquaMaps (http://www.aquamaps.org/)

Field Data

- Limited in time/space – often due to cost or project time frame

- Be aware of:

  → Effort bias

  What about species with cosmopolitan distribution? Will sampling cover species range?

  → Spatial or temporal autocorrelation

  Look and account for it (if present), otherwise limits type of statistics available, can influence results

Physical Data

Often used as a proxy for prey

- *In situ*
- Remotely sensed
- Climatology
- Ocean Models
- Static

In situ Data

<u>Primary Measurements</u>

Water Depth

Temperature

Salinity

Currents

Fluorescence

Zooplankton

Acoustic Backscatter

In situ Data Sources

Along-track

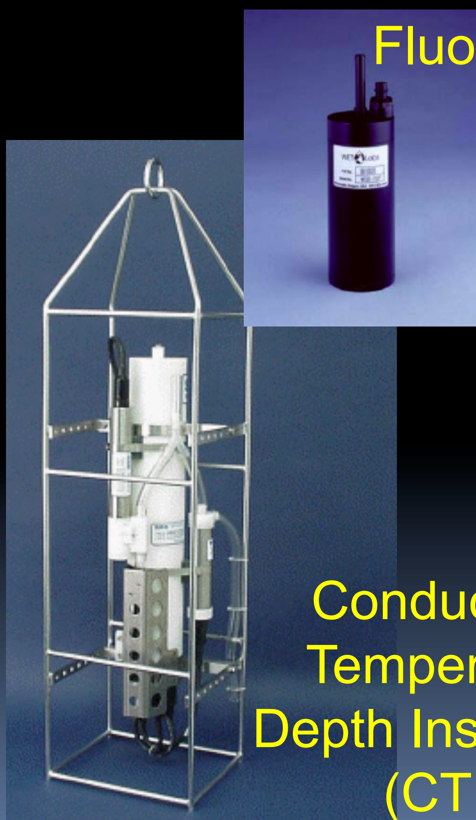
Thermosalinograph

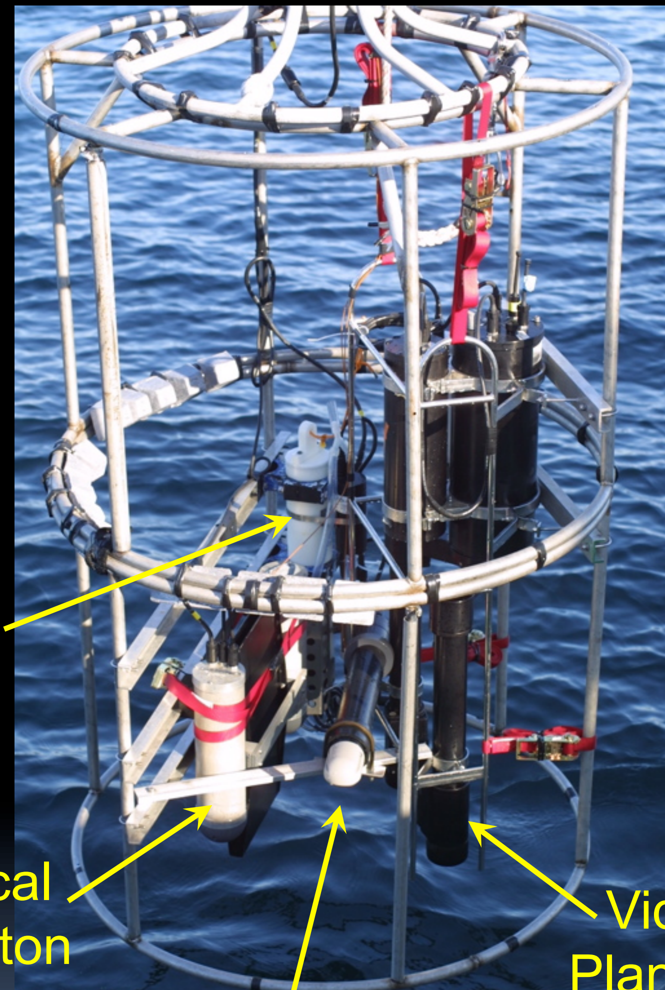
Scientific Echosounder

Acoustic Doppler
Current Profiler

In situ Data Sources

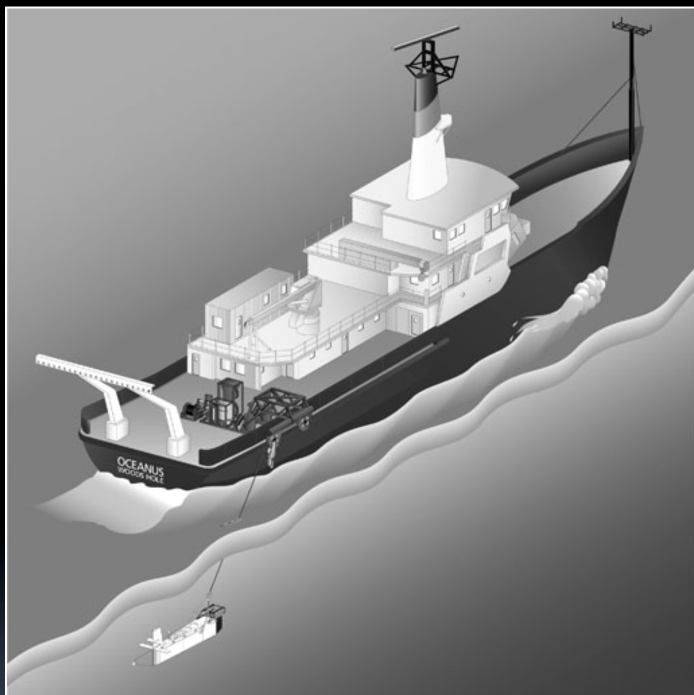
Profile

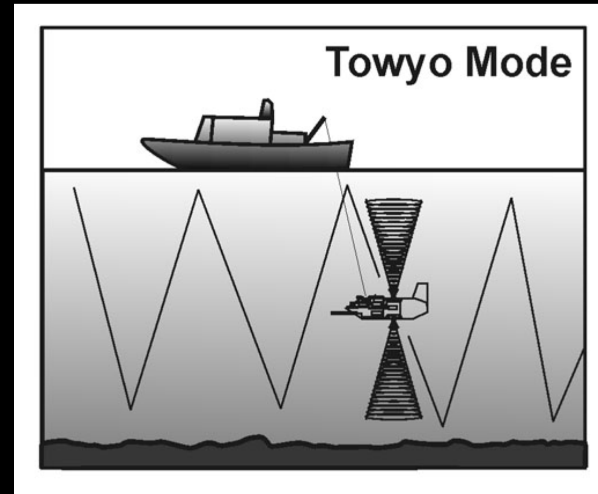


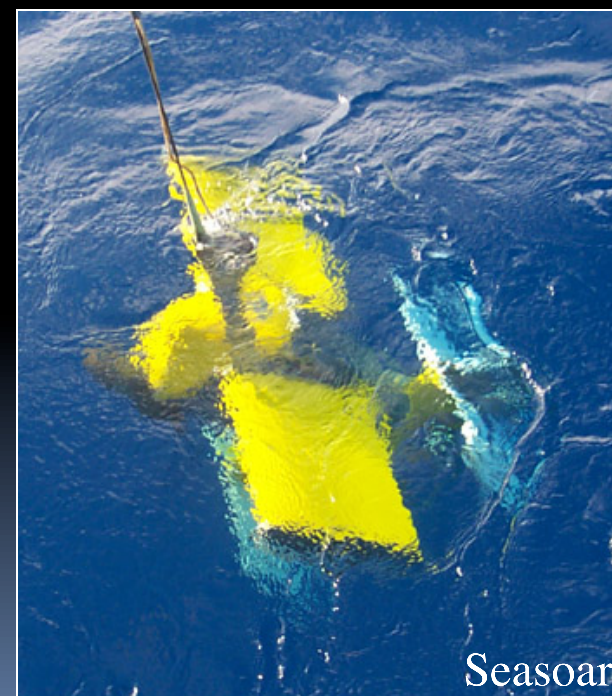


In situ Data Sources

Tow-yo'ed

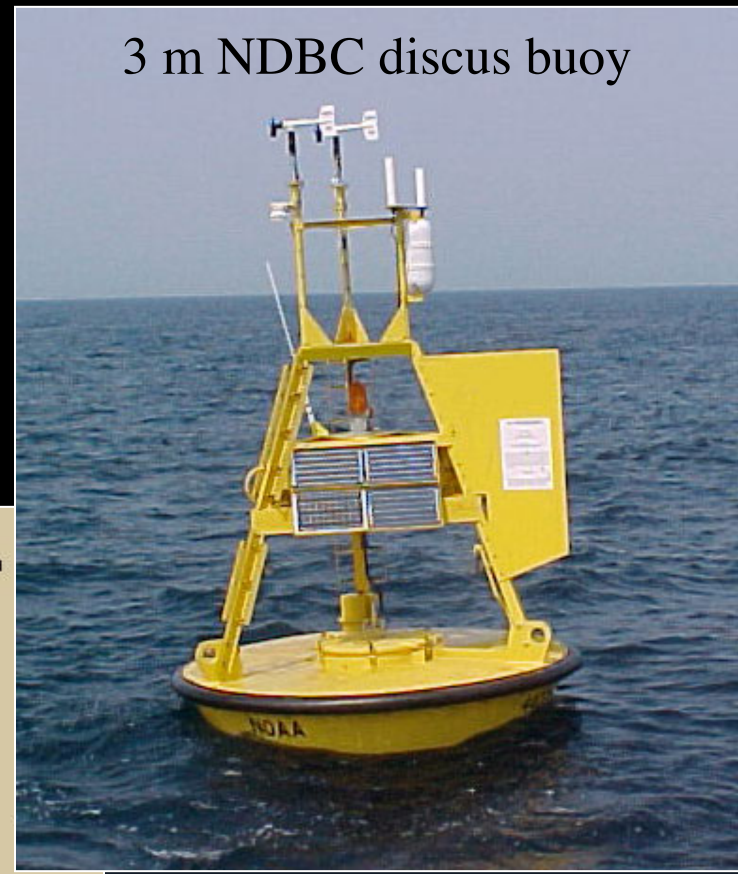






In situ Data Sources

Moored



3 m NDBC discus buoy

In situ Data Resources

Instrumentation

- → Oceanographers
- → Instrument Manufacturers

Data Archives

- → National Oceanographic Data Center (www.nodc.noaa.gov)
- → National Data Buoy Center (www.ndbc.noaa.gov)

Remote Sensing Data

SeaWiFS

Passive

→ SST

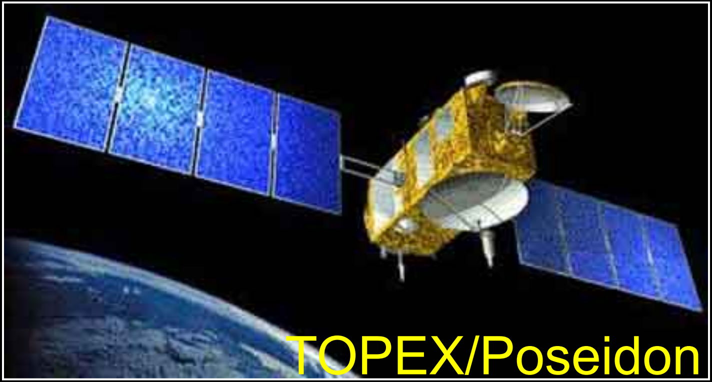
(AVHRR, MODIS, TERRA, AQUA, GHRSST)

→ Ocean color/ 1° production

(SeaWiFS, MODIS)

→ Harmful Algal Blooms

(SeaWiFS)



TOPEX/Poseidon

AVHRR

Active

→ Sea surface height

(GEOS-3, Jason 1, OSTM- Jason 2, TOPEX-POSEIDON)

→ Surface Wind

(METOP-A, QuickSCAT, SSM/I,)

→ Salinity (Aquarius – launched Aug 2011)

Remote Sensing Data Resources

JPL Data Archive -- http://podaac.jpl.nasa.gov/

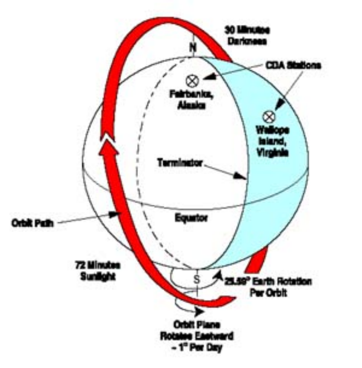
SST, Sea Surface Height, Surface wind, Salinity

NOAA CoastWatch -- http://coastwatch.noaa.gov/

SST, winds, ocean color, Harmful algal blooms

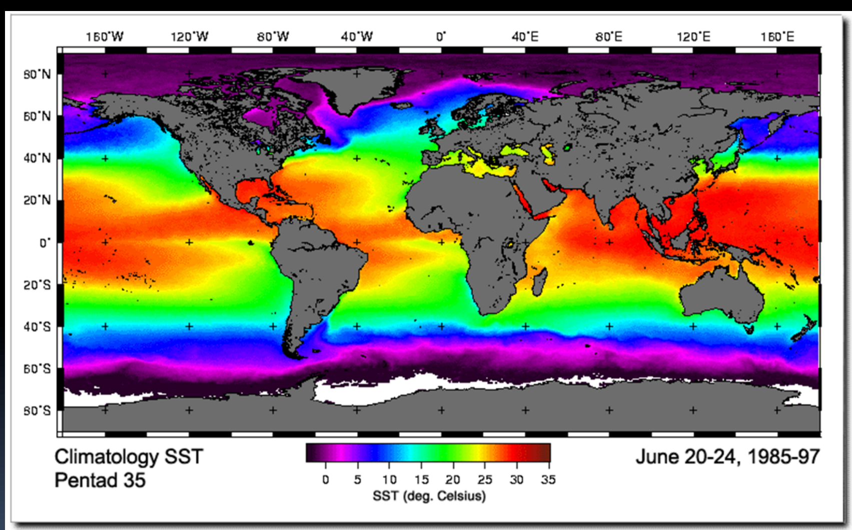
Aviso - http://www.aviso.oceanobs.com

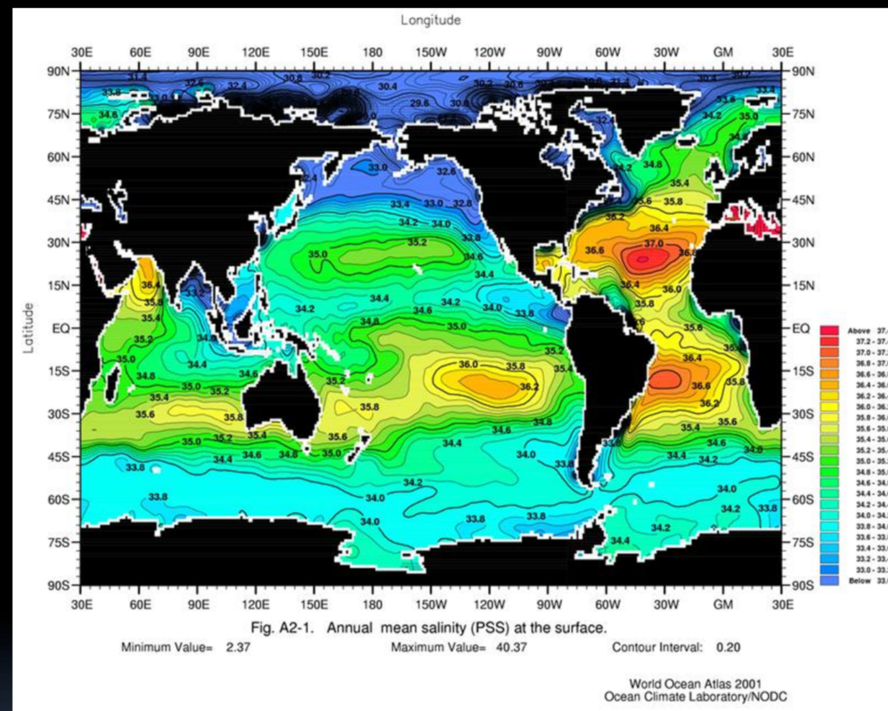
Sea surface height, surface wind, wave height, mean sea level



Climatology Data

When should you use climatology (e.g. what is the appropriate temporal scale for your model)?

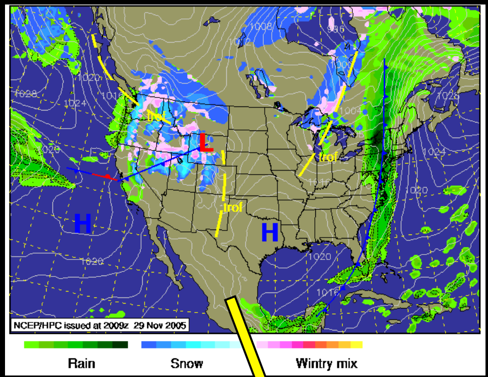


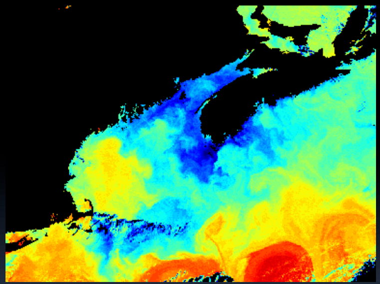


Surface Salinity Climatology
from World Ocean Atlas ("Levitus")

Often used to look at long term average conditions and/or deviations
→ SST, salinity

Ocean Model Data

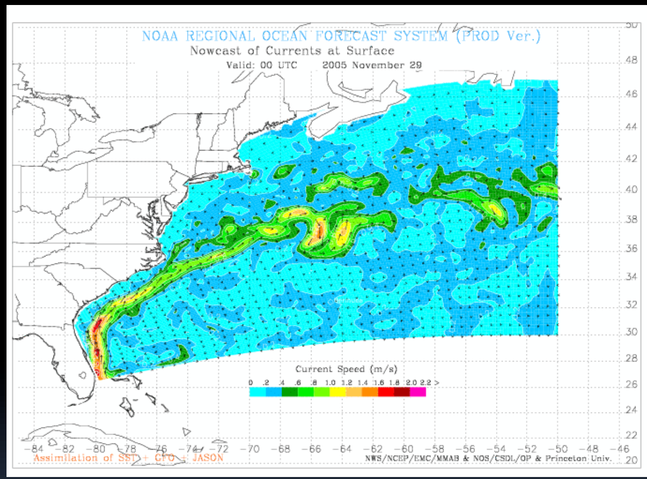




$$\frac{\partial u}{\partial t} - fv + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0$$

$$\frac{\partial v}{\partial t} + fu + \frac{1}{\rho}\frac{\partial p}{\partial y} = 0$$

$$-g\rho + \frac{\partial p}{\partial z} = 0$$



Model Surface Currents in the Northwest Atlantic from the NOAA *Regional Ocean Forecast System*

Climatology and Ocean Model Data Resources

Global

→ World Ocean Database and World Ocean Atlas
www.nodc.noaa.gov/OC5/indprod.html

→ U.S. Navy Models
http://www7320.nrlssc.navy.mil/global_nlom/

Regional

→ EMC Marine Modeling & Analysis Branch
http://polar.ncep.noaa.gov/ofs/

→ Regional Ocean Modeling System (ROMS)
http://www.myroms.org/

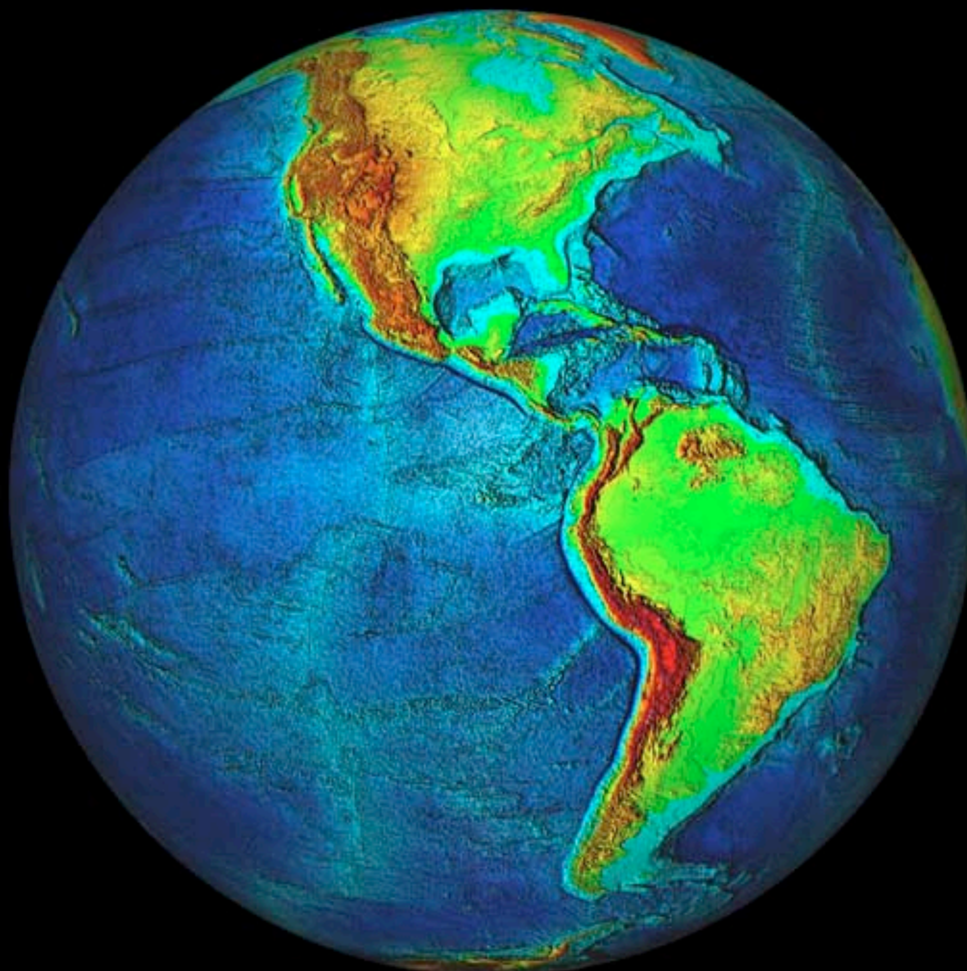
Local

→ Look in literature, request data from authors or local modelers

Static Data

Bathymetry

Coastlines



Static Data Resources

Scripps Institution of Oceanography

http://topex.ucsd.edu/marine_topo/

Bathymetry

U.S. National Geophysical Data Center

http://www.ngdc.noaa.gov/mgg

Bathymetry, coastlines

NOAA Shoreline Website

http://shoreline.noaa.gov/

Coastlines

Natural Earth

http://www.naturalearthdata.com/

Bathymetry, coastlines, oceans, reefs, rivers, Antarctic ice shelves

Data Considerations

Accuracy

Precision

Uncertainty

→ Parameter estimation
→ Observational
→ Design
→ Stochasticity

Sample size

→ How big is your sample?

→ Is it enough to detect a meaningful pattern/ process?

→Do you have data to test & evaluate model?

Defining a Sampling Unit

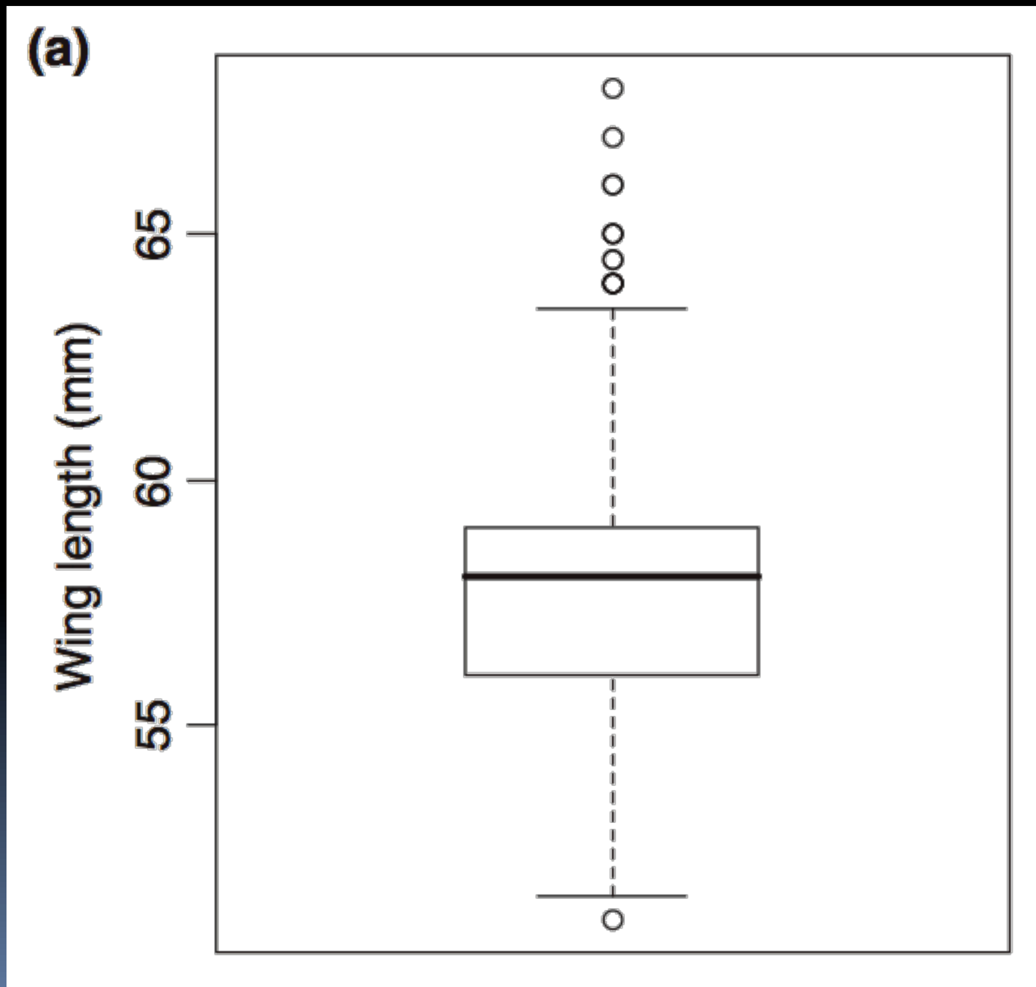
What is research question?

What scale would be applicable to detect pattern or process being studied?

- Spatio-temporal resolution of datasets often mixed
- Depends on data
- Detection of pattern or process?

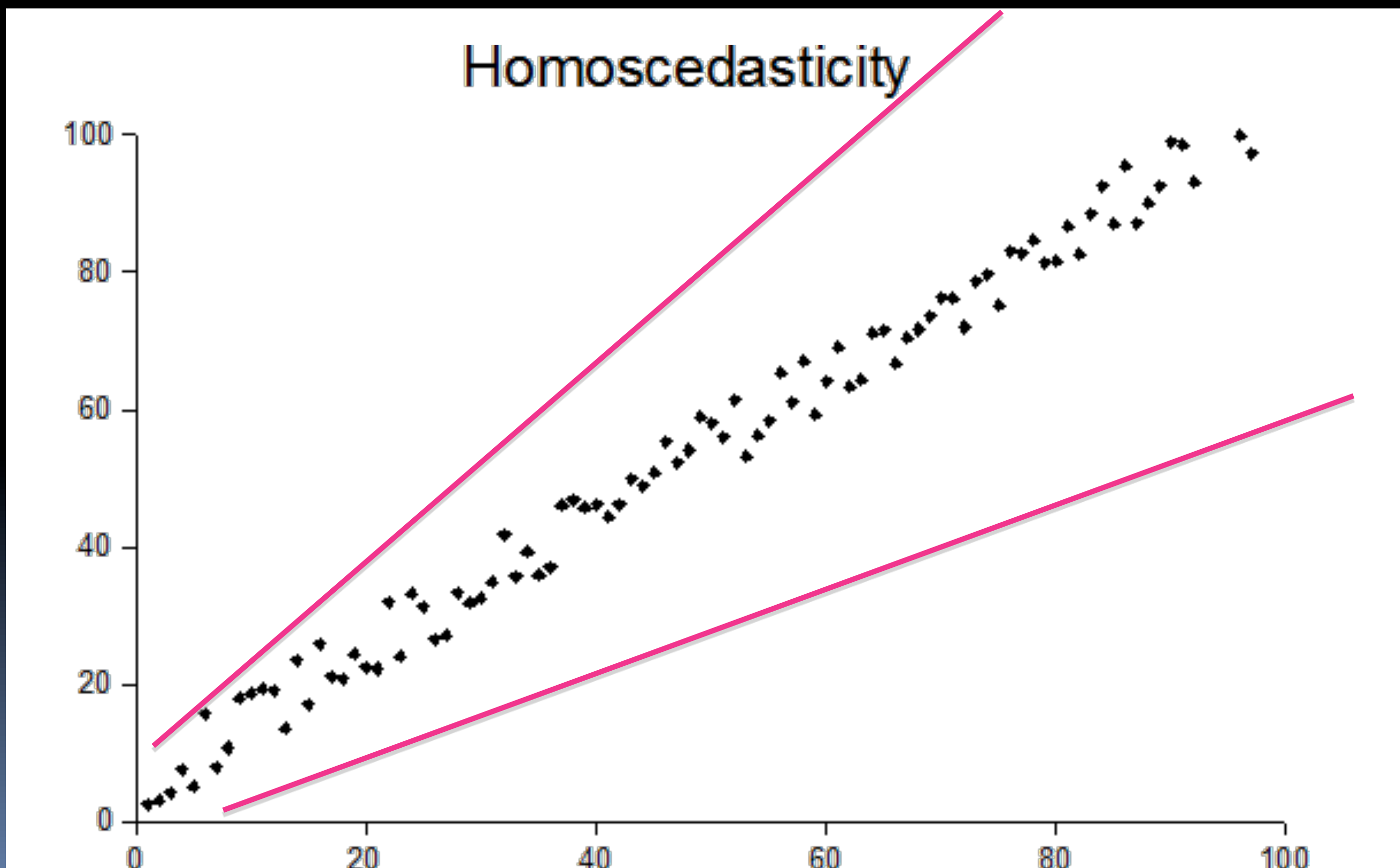
Data exploration (Zuur et al. 2010)

- Step 1: Are there outliers in Y and X?



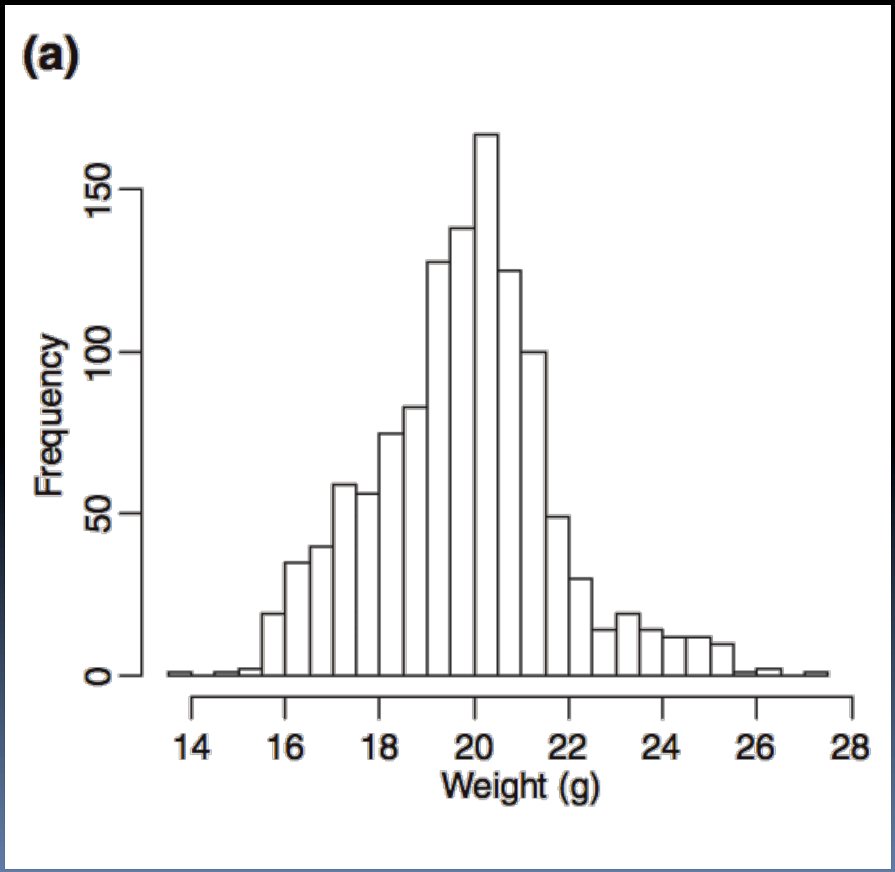
Data exploration (Zuur et al. 2010)

- Step 2: Do we have homogeneity of variance?



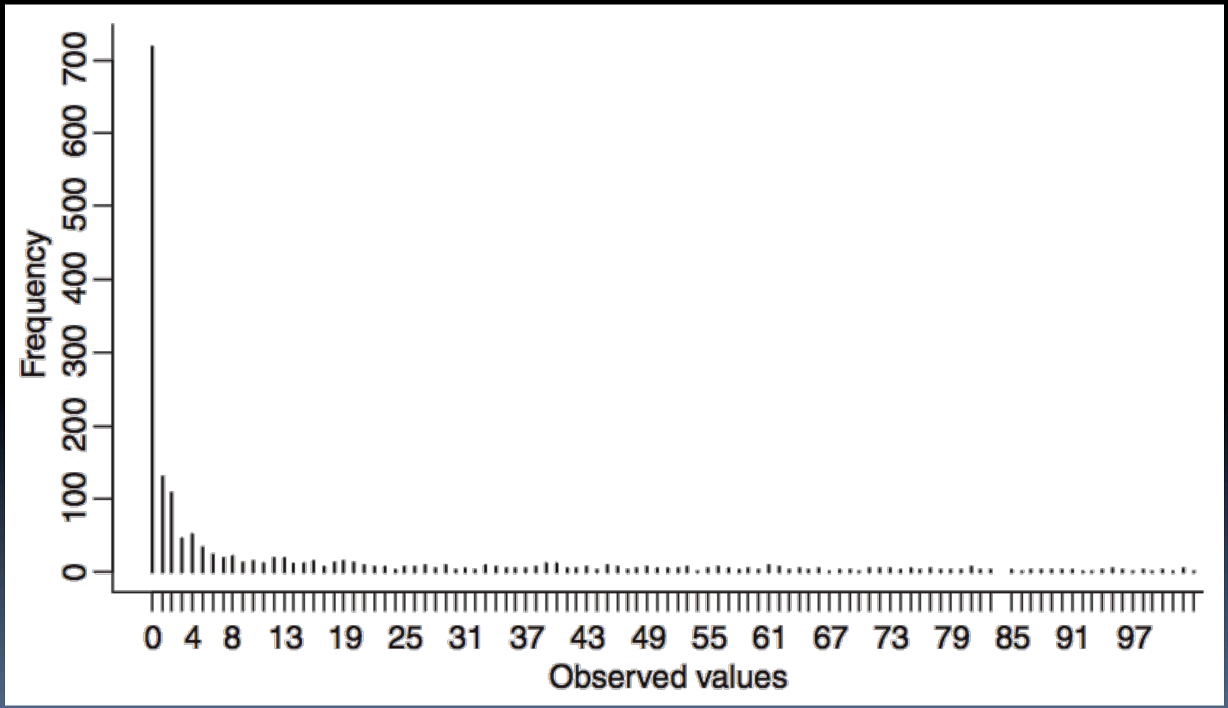
Data exploration (Zuur et al. 2010)

- Step 3: Are the data normally distributed?



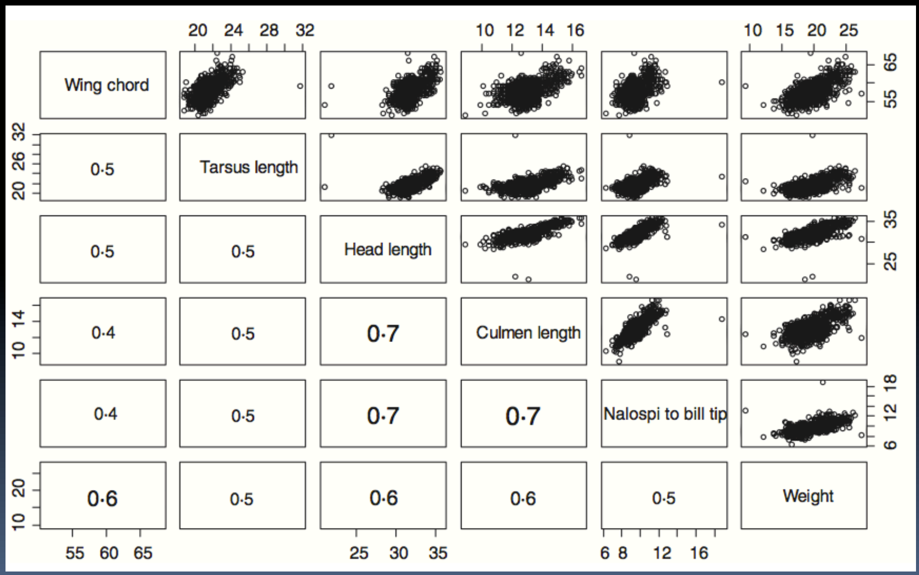
Data exploration (Zuur et al. 2010)

- Step 4: Are there lots of zeros in the data?



Data exploration (Zuur et al. 2010)

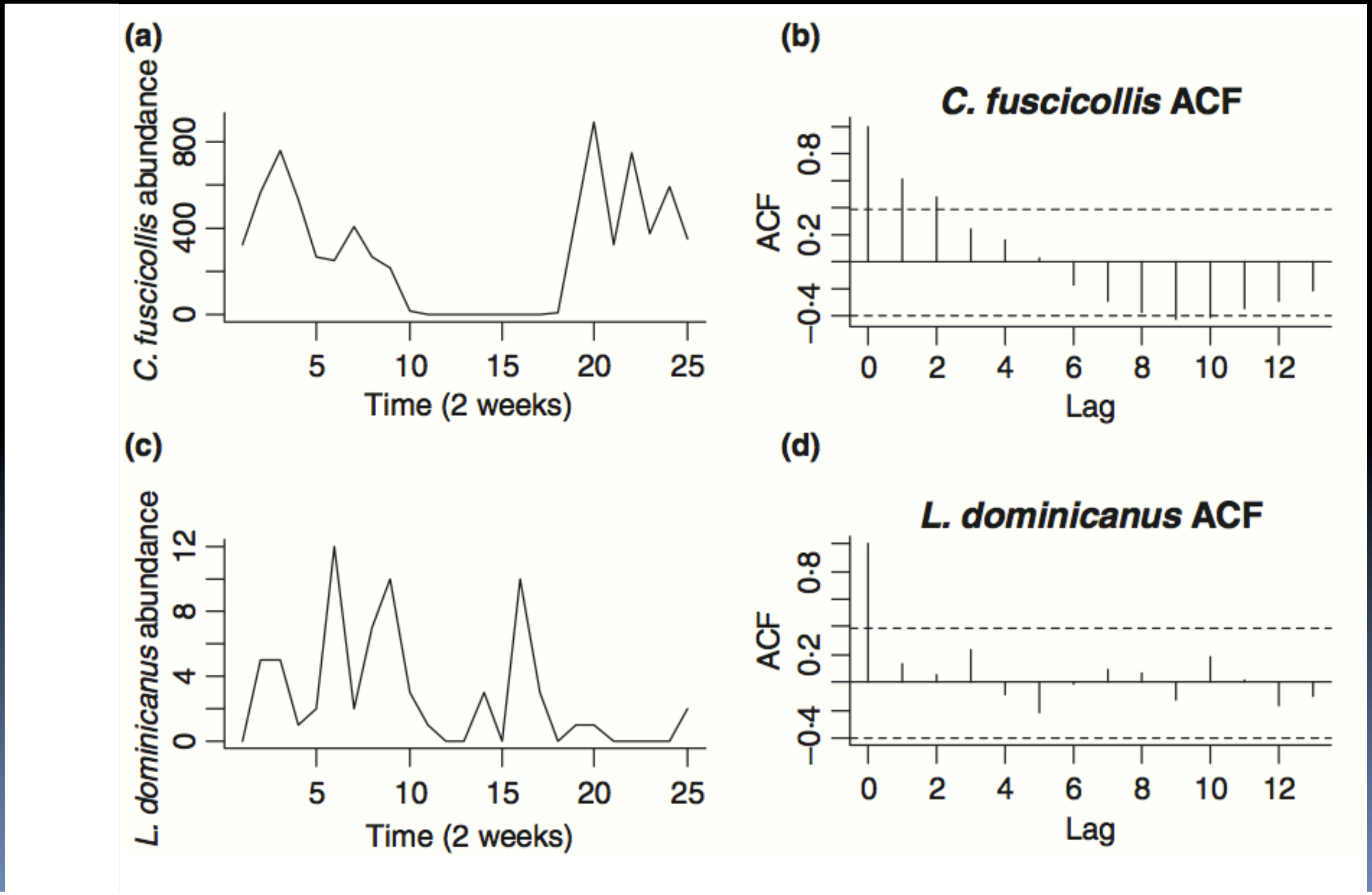
- Step 5: Is there collinearity among the covariates?
- Step 6: What are the relationships between Y and X variables?



Data exploration (Zuur et al. 2010)

- Step 7: Should we consider interactions?

- Step 8: Are observations of the response va
independent?



Evaluation and Testing

- Most important
- Frequently ignored

- Recognize that not all models need same level of validation (Rykiel 1996)

- Sensitivity analysis can increase confidence in model accuracy

Model evaluation

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

Observed

Expected	+	-
+	a	b
-	c	d

- Chi square
- Receiver operator curves
- Kappa statistic

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

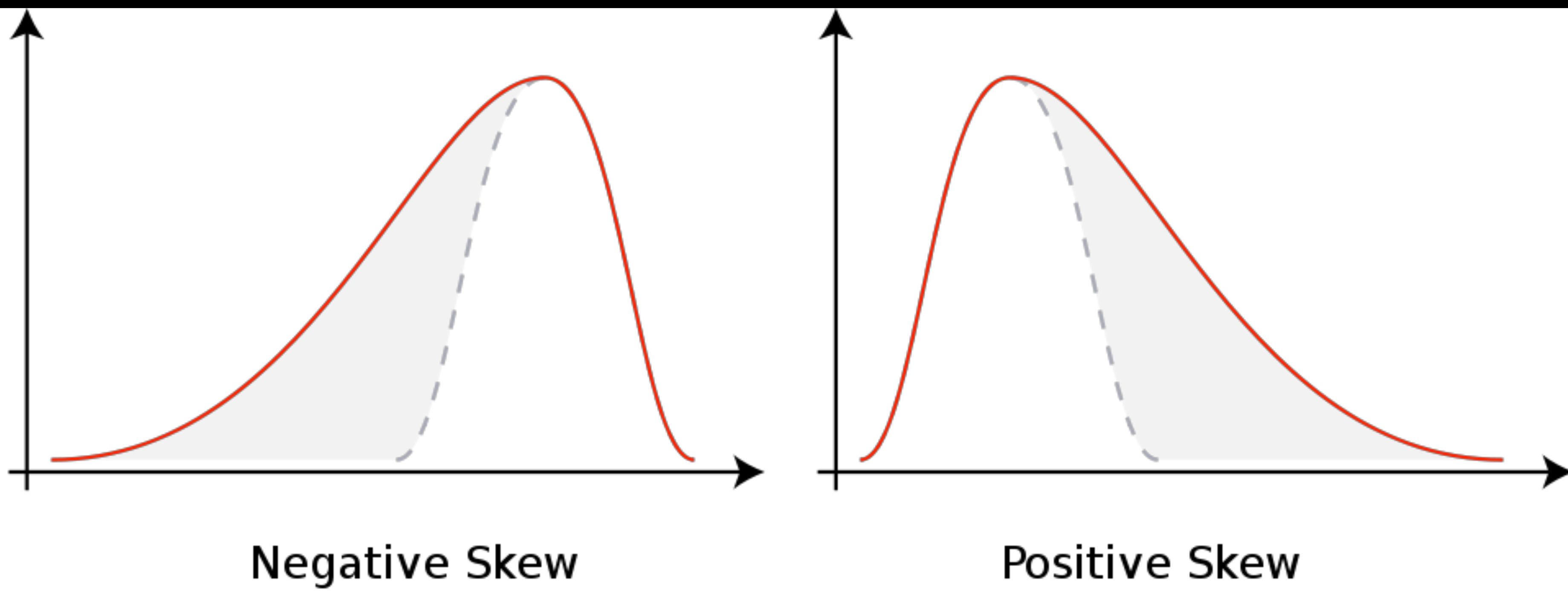
Possible solutions

- Pseudo-absence data
  - → Assumes no bias in presence sampling
  - → Influenced by extent of study
- Null model comparisons
- Presence-only models
  - → E.g. skewness test
  - → Let the presence data tell you what is best

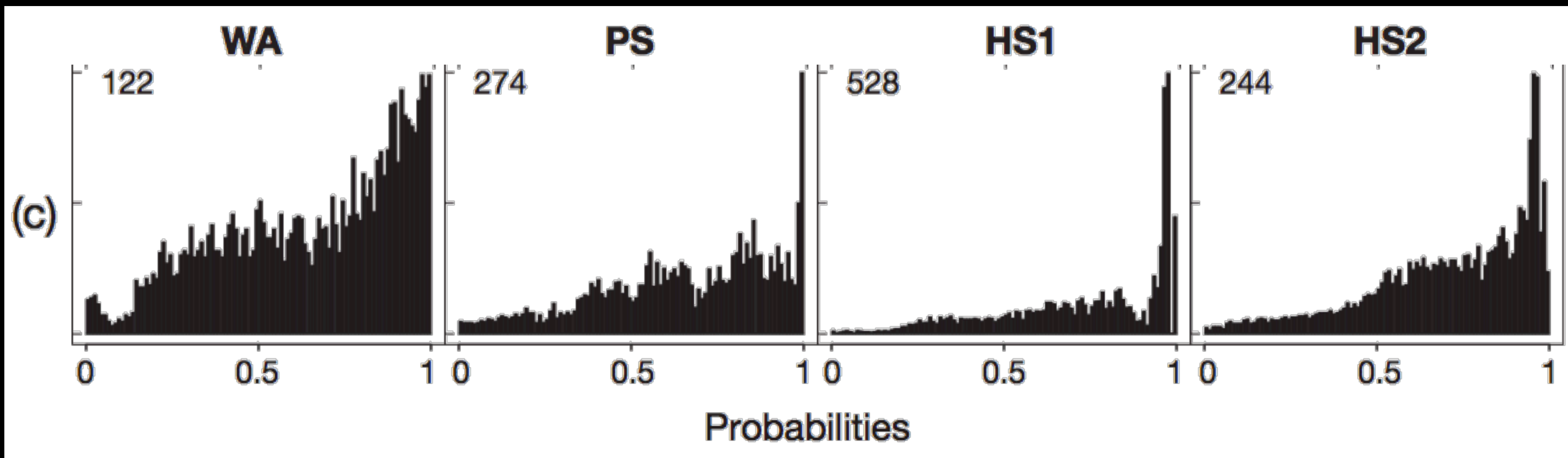
Skewness test

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



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

$$SS_m = \frac{\sum_{N_P}(1-\text{presence})^2 + \sum_{1-N_P}(\text{absence})^2}{N - 2p_m}$$

Gregr and Trites 2008. *Marine Ecology Progress Series*

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.

Modeling tips

- Ask a clear question and let it drive your approach
- Add complexity only where necessary
- Understand your data before you try to model them
  - → Exploratory analyses (Zuur et al. 2010)
- Ensure transparency
  - → in purpose, in methodology, & 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 land does not move (on the same scales as the ocean)