

Survey design and statistical analysis for conservation science



Philipp Boersch-Supan
British Trust for Ornithology



with thanks to Alison Johnston

Survey design and statistical analysis for conservation science

SURVEY DESIGN AND SAMPLING

STATISTICAL ANALYSIS

COMMUNICATING RESULTS

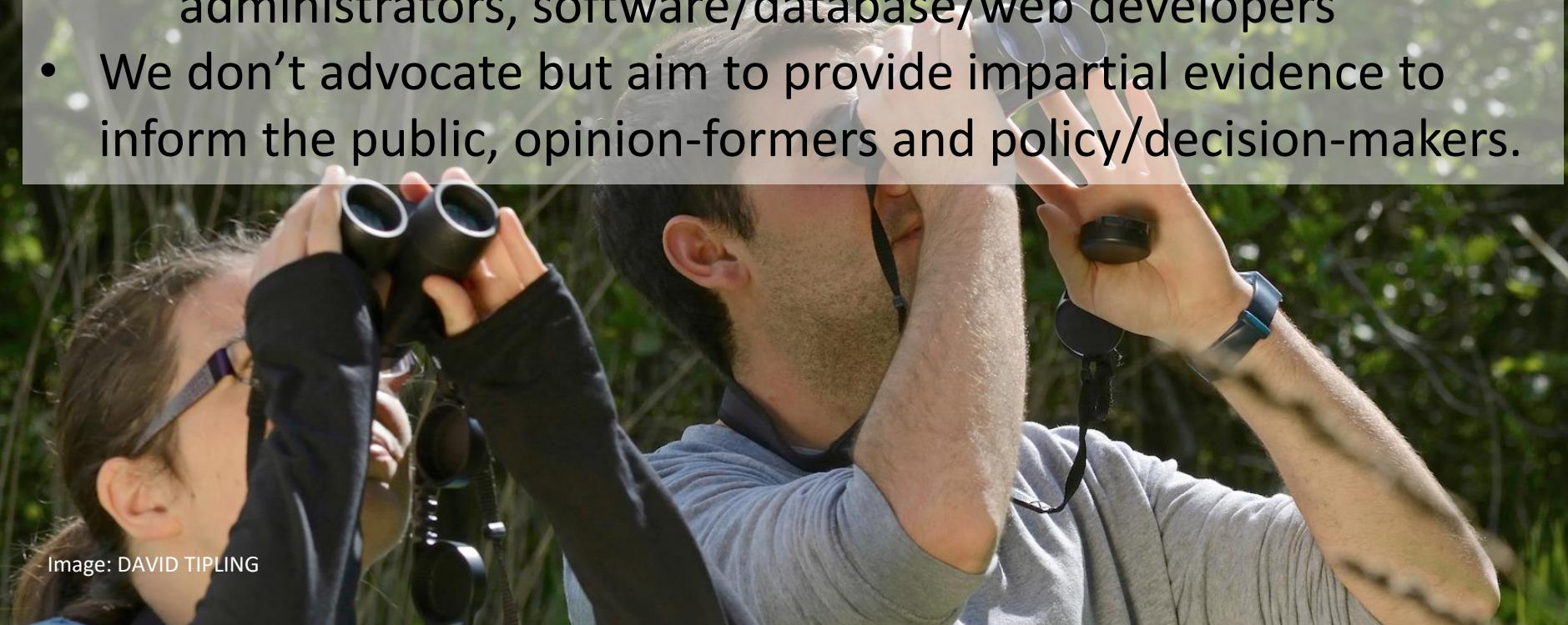
About me

		Field	Lab	Maths	Stats	Coding
2005	Field Ornithologist					
2006-2008	UG in Chemistry					
2008-2009	MRes in Environmental Biology					
2009-2013	PhD in Marine Ecology					
2013-2014	Analyst at British Antarctic Survey					
2014-2018	Postdoc Quantitative Ecology					
2018-	Ecological Statistician					

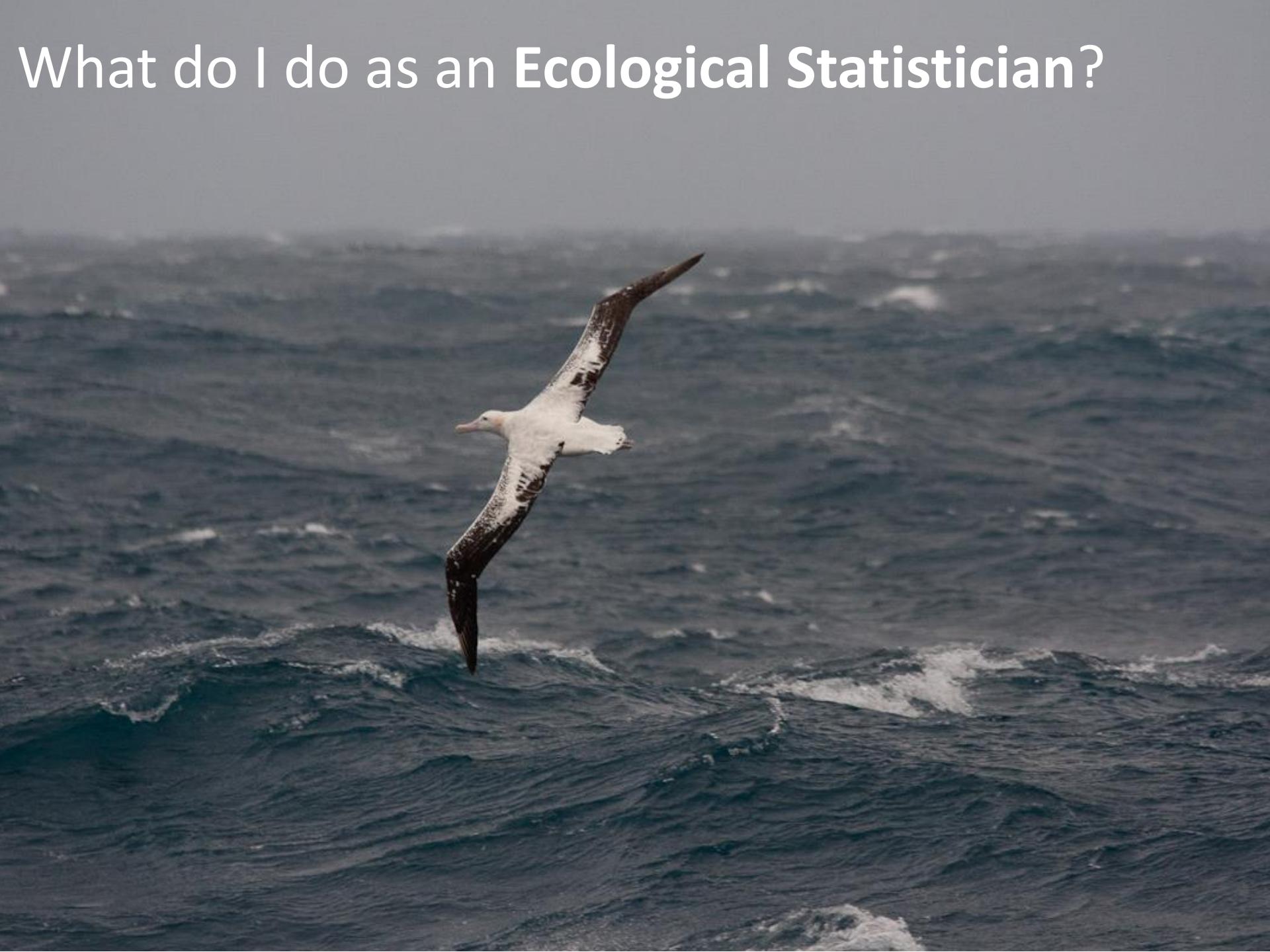


The British Trust for Ornithology

- We are an independent charitable research institute
- We monitor UK birds with the help of over 40,000 volunteers
- We employ c. 100 staff
 - c. 50 scientists
 - c. 50 communicators, fundraisers, volunteer coordinators, administrators, software/database/web developers
- We don't advocate but aim to provide impartial evidence to inform the public, opinion-formers and policy/decision-makers.



What do I do as an Ecological Statistician?



I try to build models that

- turn monitoring data into assessment tools
- lead to a better understanding of **biological processes**
- allow predictions outside the range of our data
- enable informed decisions for conservation actions



- make the best use of limited (and sometimes ‘messy’) data
- are statistically and computationally **tractable**
- use tools/workflows that are easy to reuse and share

Ecological models

Tactical/Phenomenological

Strategic/Mechanistic

Describe patterns
without elucidating
mechanism

Prediction within range of data

Statistical models
(regressions, etc.)

Focus on process/mechanisms

Explanation or understanding
Prediction outside range of data

Math models e.g. ODEs, PDEs
Individual/Agent based Models

Historically largely separate communities of researchers within ecology,
but increasingly integration of data and theory-focussed approaches happens

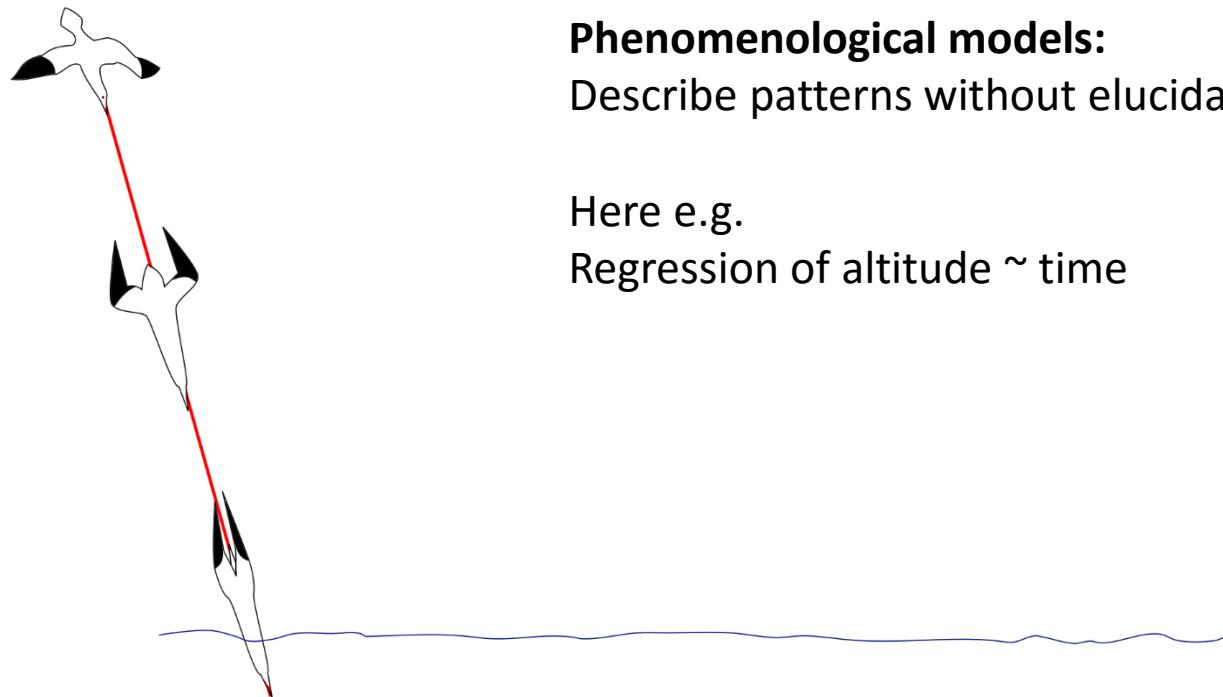
Models in ecology – an analogy



Observations



Models in ecology – an analogy



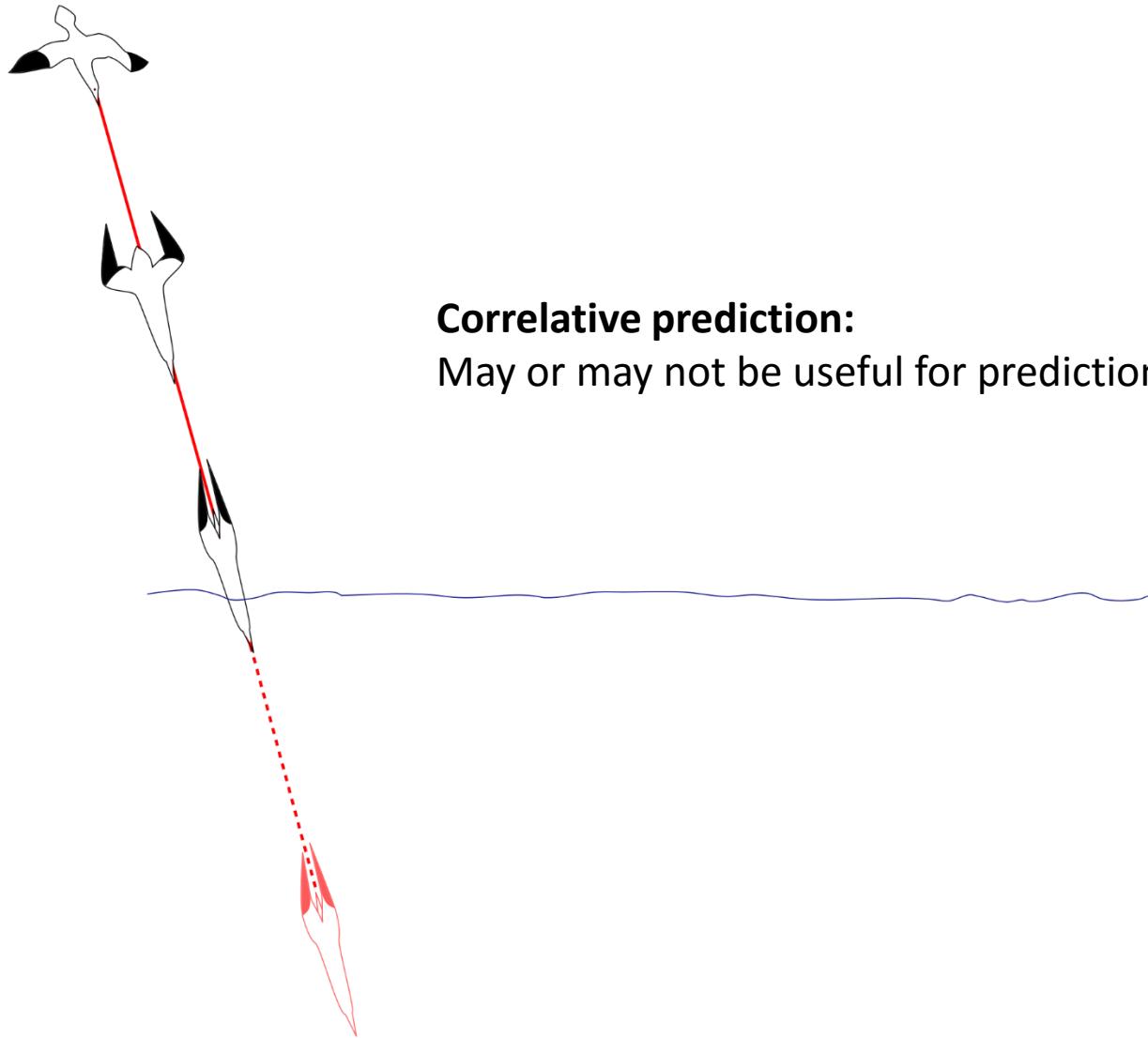
Phenomenological models:

Describe patterns without elucidating mechanism

Here e.g.

Regression of altitude \sim time

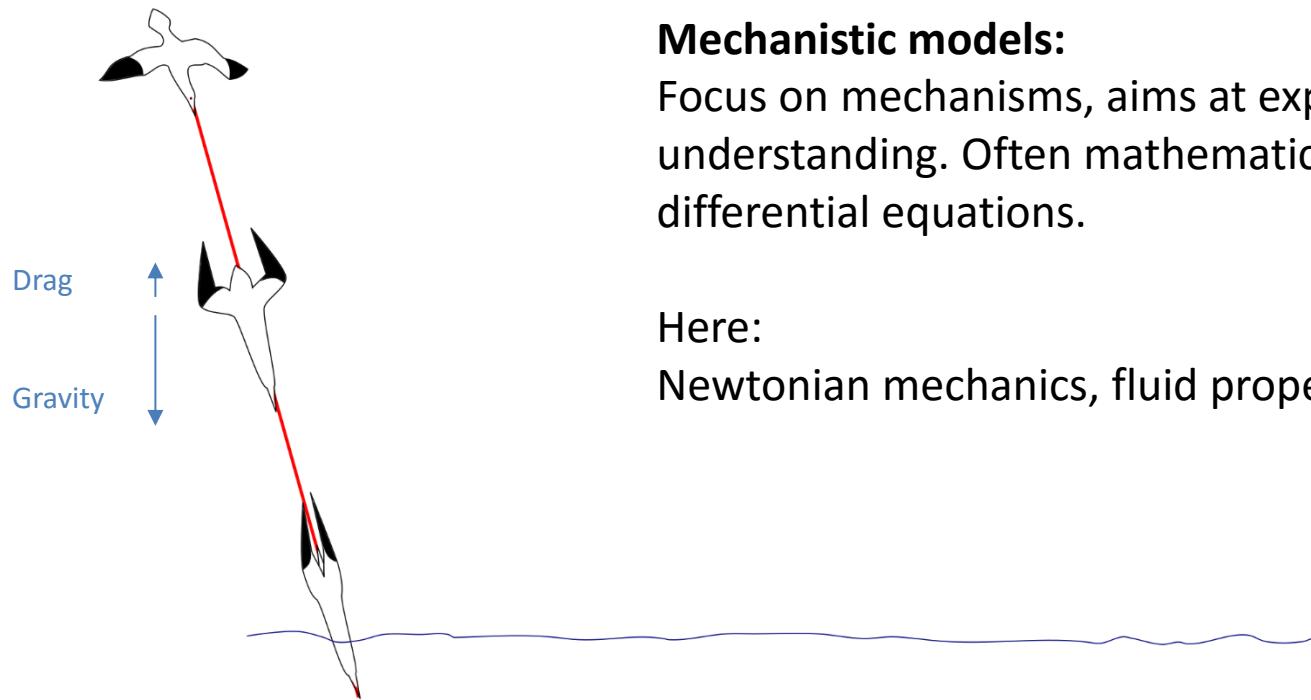
Models in ecology – an analogy



Correlative prediction:

May or may not be useful for prediction beyond data

Models in ecology – an analogy



Mechanistic models:

Focus on mechanisms, aims at explanation or understanding. Often mathematical models like differential equations.

Here:

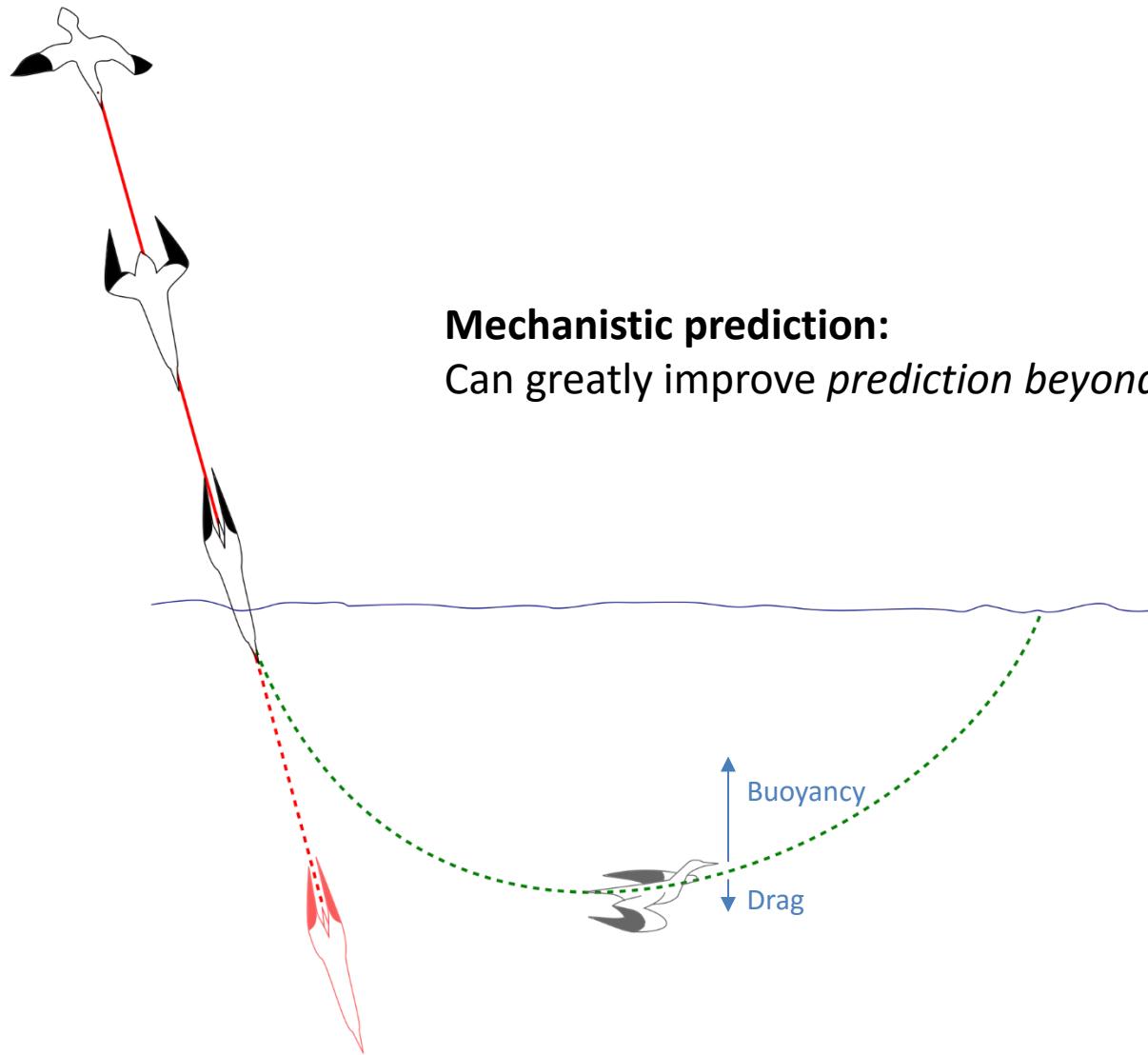
Newtonian mechanics, fluid properties

Other examples:

Lotka-Volterra predator-prey models

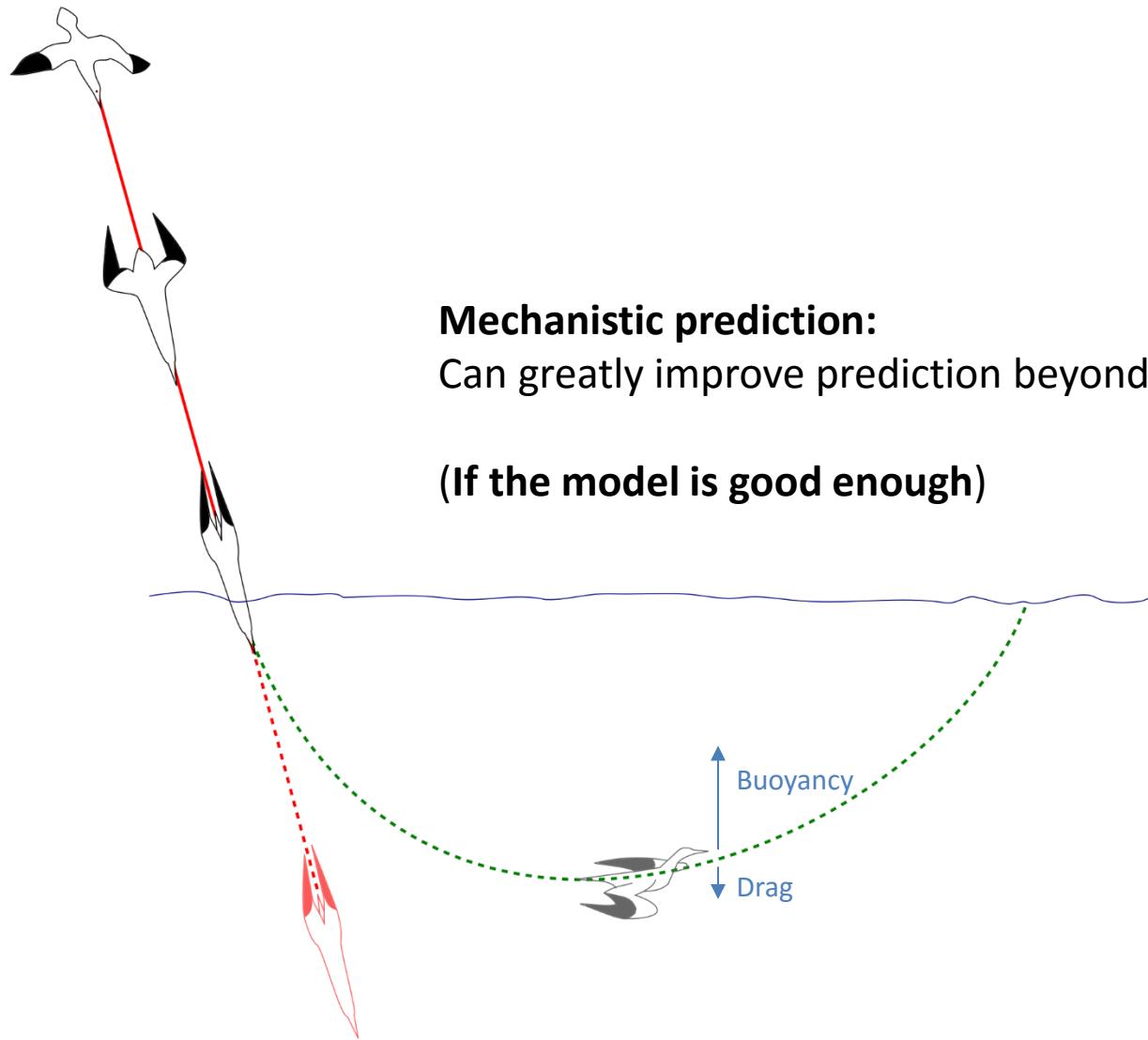
SIR model in epidemiology

Models in ecology – an analogy



Mechanistic prediction:
Can greatly improve *prediction beyond data*

Models in ecology – an analogy



Tactical/Phenomenological

Strategic/Mechanistic

Describe patterns
without elucidating
mechanism

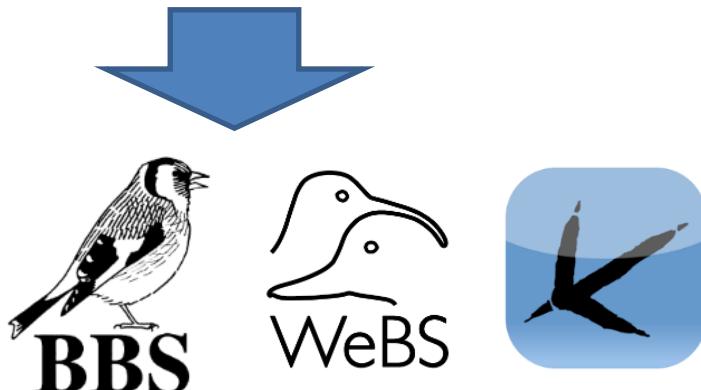
Prediction within range of data

Statistical models
(regressions, etc.)

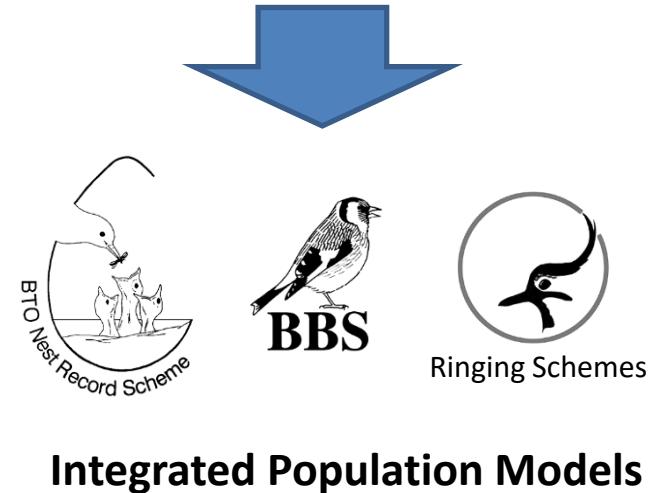
Focus on process/mechanisms

Explanation or understanding
Prediction outside range of data

Math models e.g. ODEs, PDEs
Individual/Agent based Models



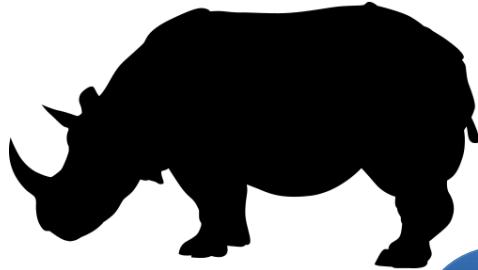
UK Bird Trends



Integrated Population Models

SURVEY DESIGN AND SAMPLING





How many rhino horns are sold each year and where?

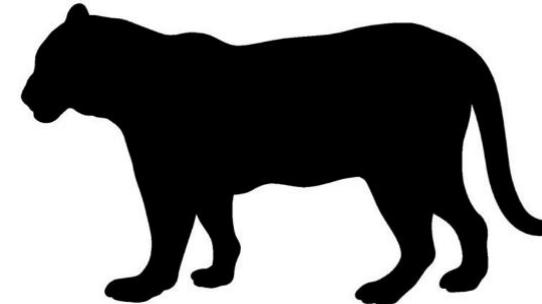
How does habitat affect blue tit nesting success?

How is sea temperature affecting coral reefs?

What's the maximum fragmentation that a tropical frog species can tolerate?

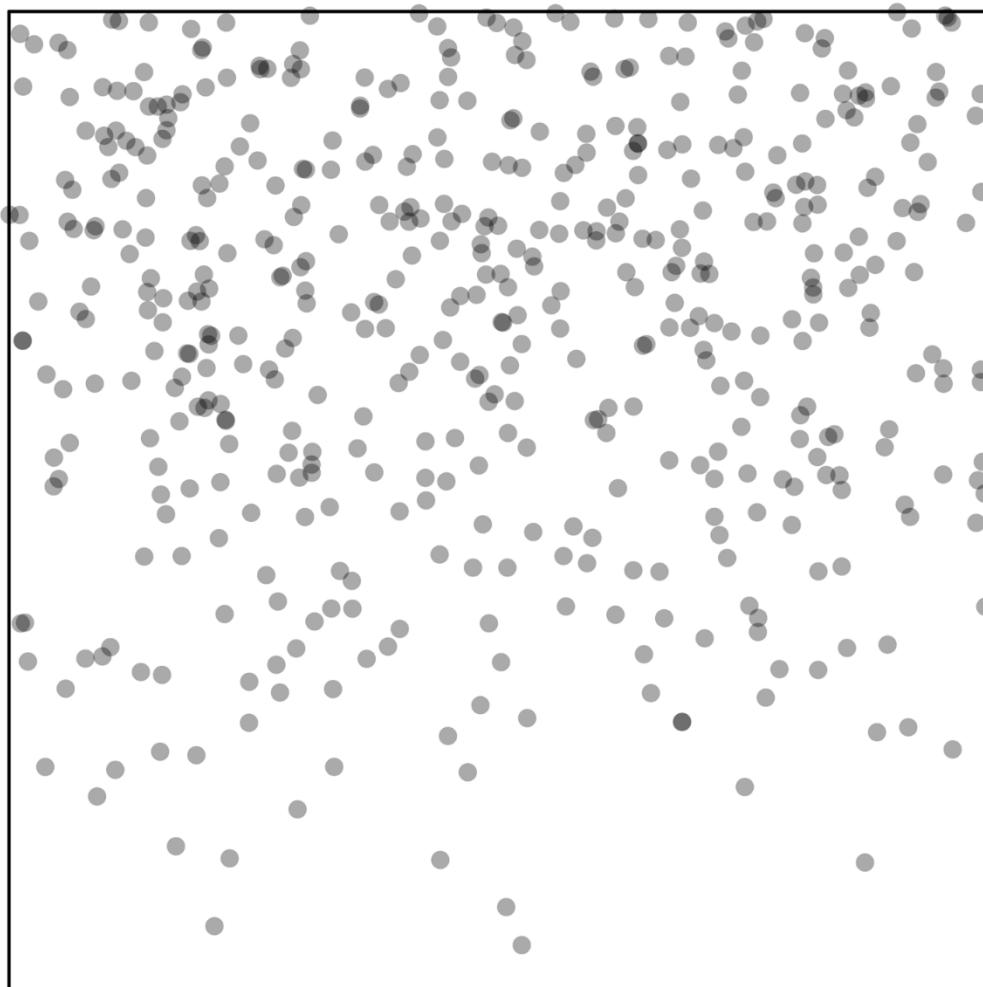
How many tigers are in India?

What data would you collect?

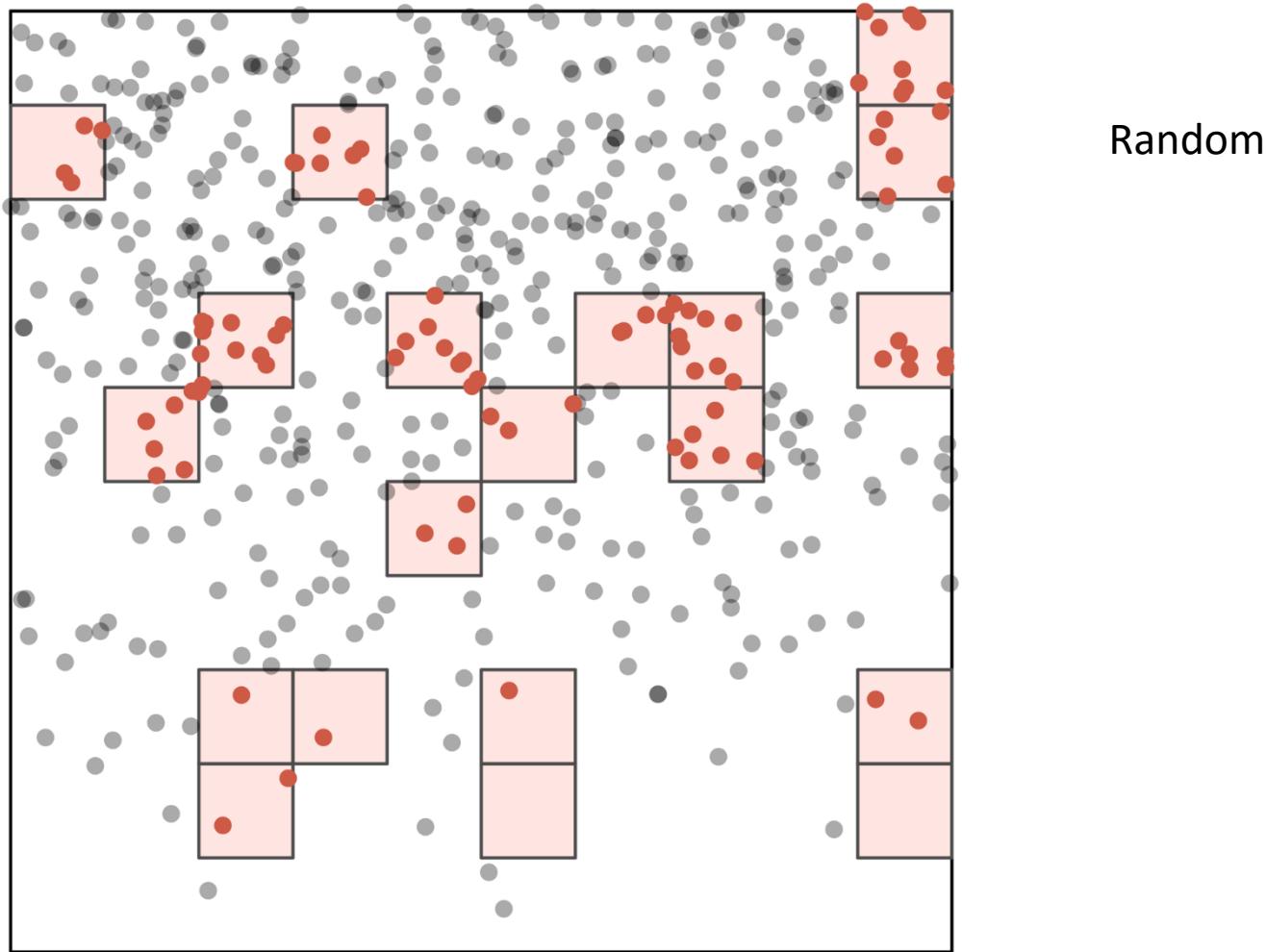


What compromises will you make?

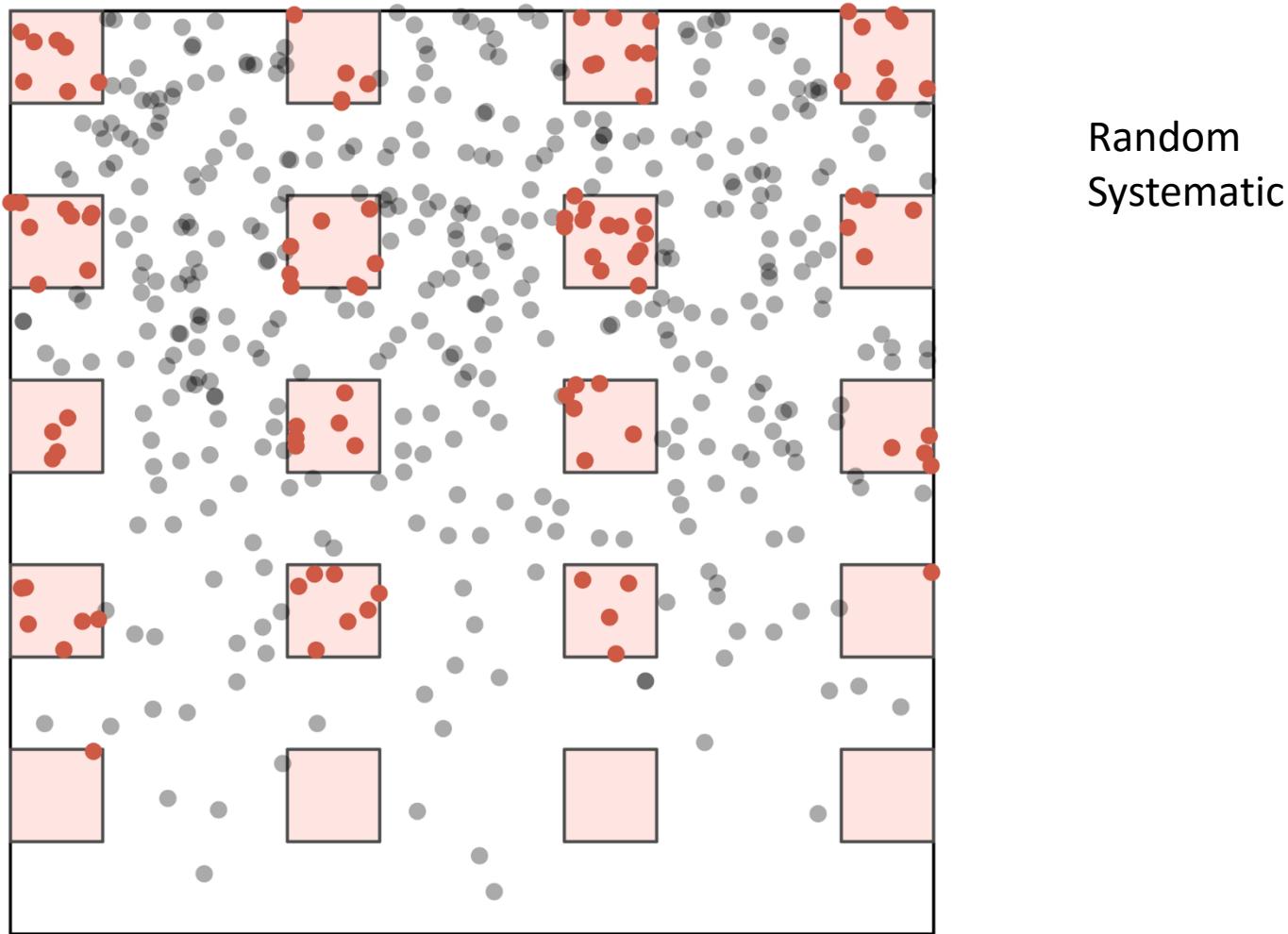
Estimate number of trees



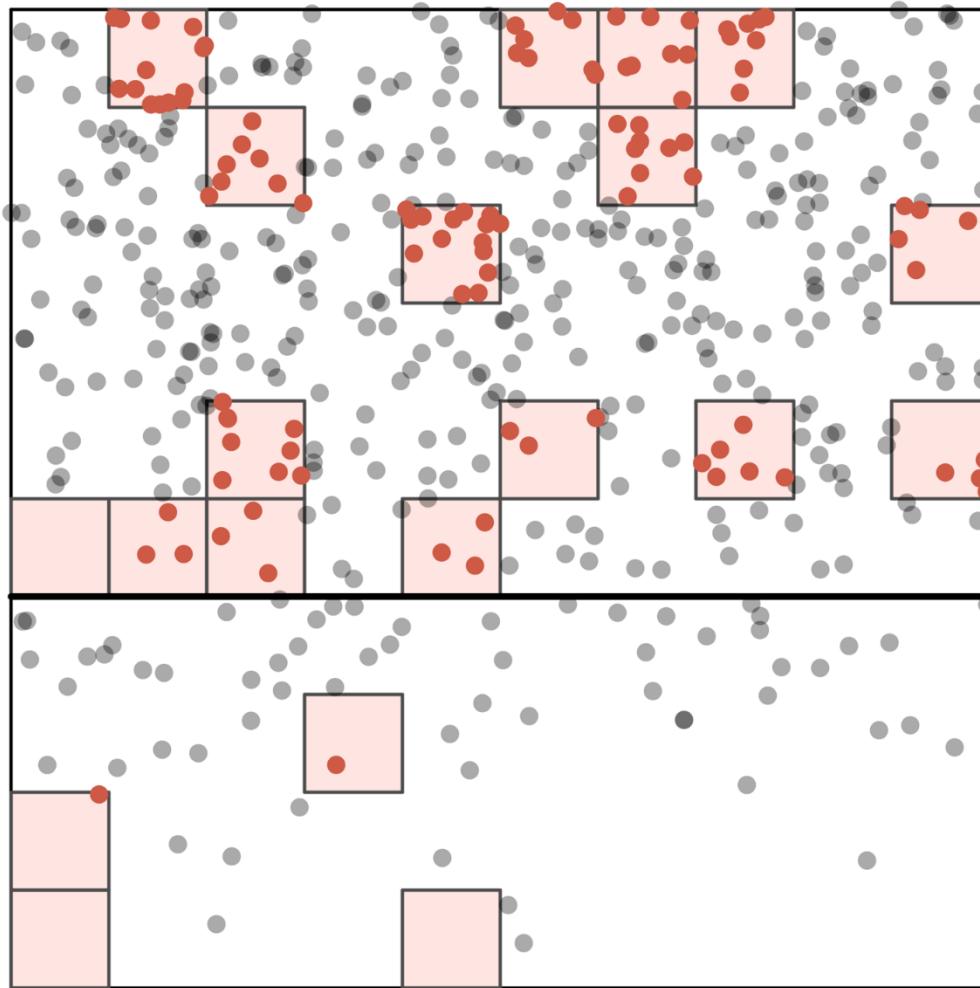
Estimate number of trees



Estimate number of trees



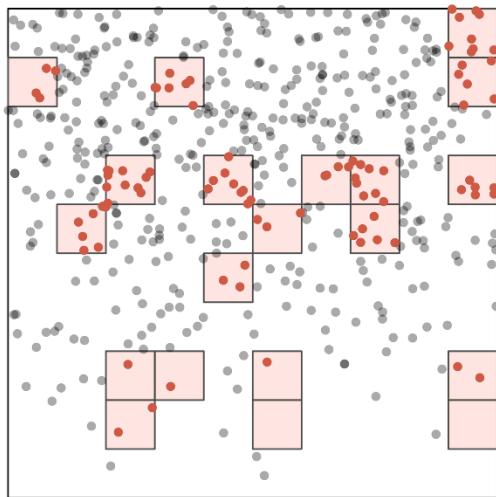
Estimate number of trees



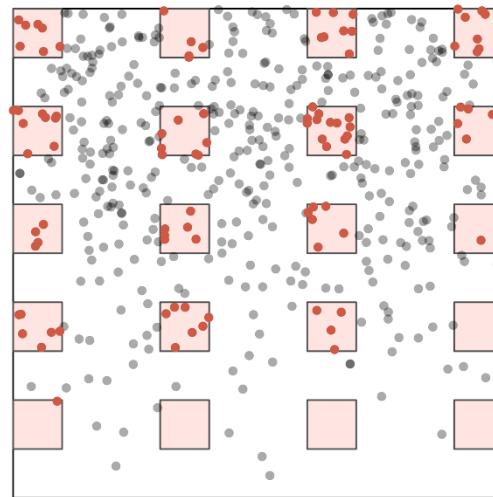
Random
Systematic
Stratified

Estimate number of trees

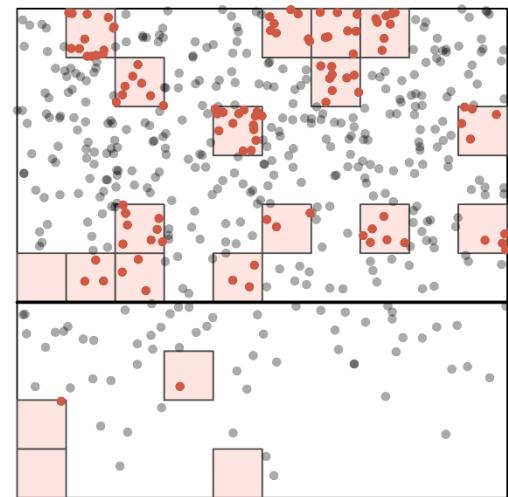
Target: REPRESENTATIVE SAMPLE



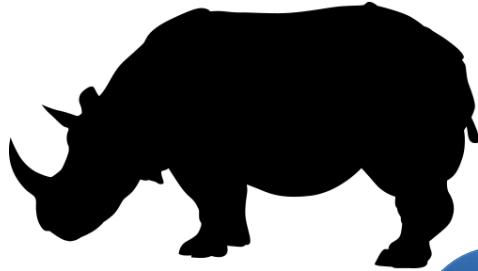
RANDOM



SYSTEMATIC



STRATIFIED



How many rhino horns are sold each year and where?

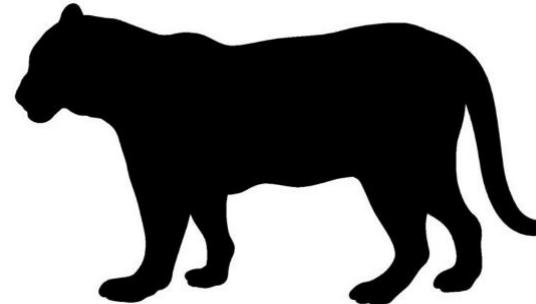
How does habitat affect blue tit nesting success?

How is sea temperature affecting coral reefs?

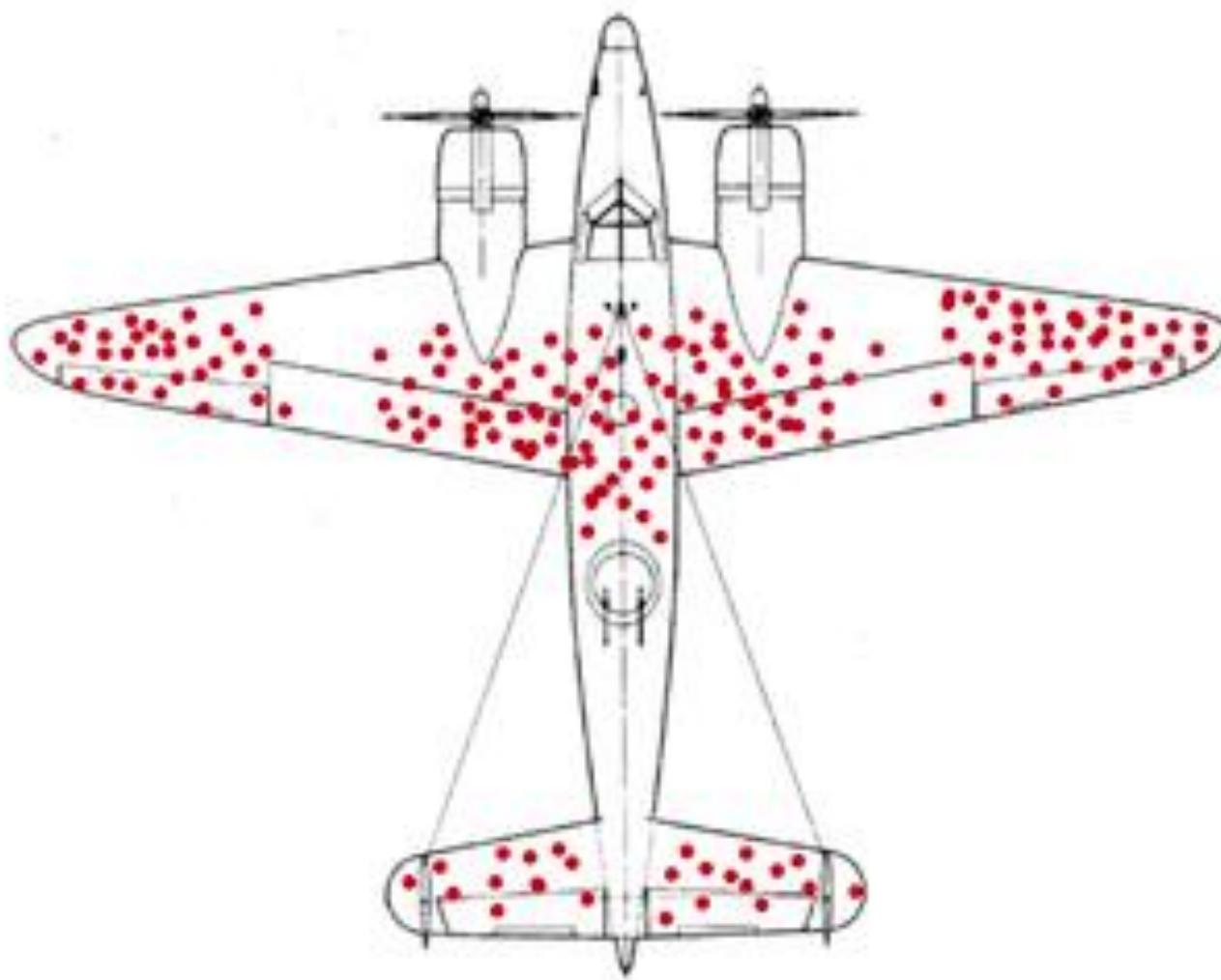
What's the maximum fragmentation that a tropical frog species can tolerate?

How many tigers are in India?

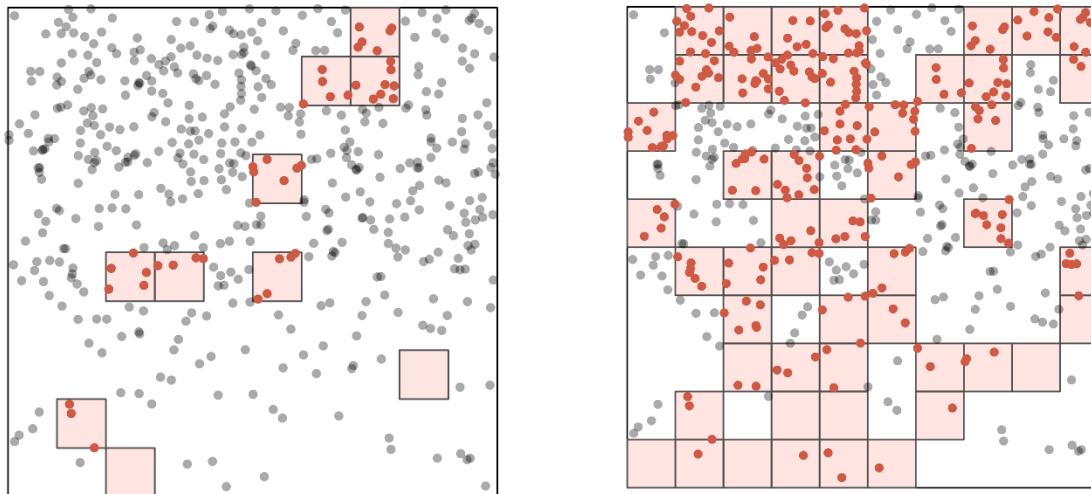
- Is your sample **representative**?
- Where do your results apply?



Can you actually observe what matters?



Replication



How many samples do I need?

How many samples?

The power of **POWER ANALYSIS**

1. State what you want to estimate

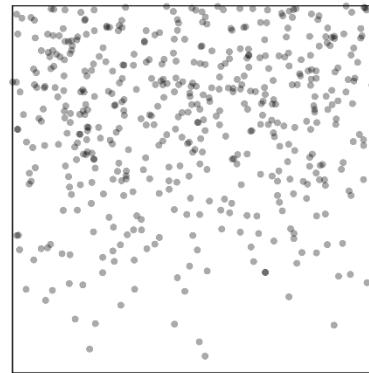
Number of trees in area
within 20% of true number

How many samples?

The power of **POWER ANALYSIS**

1. State what you want to estimate
2. Simulate a realistic version of what you want to estimate

Number of trees in area
within 20% of true number

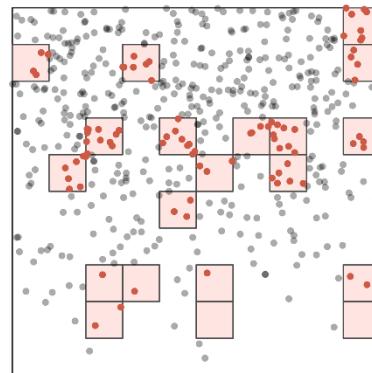
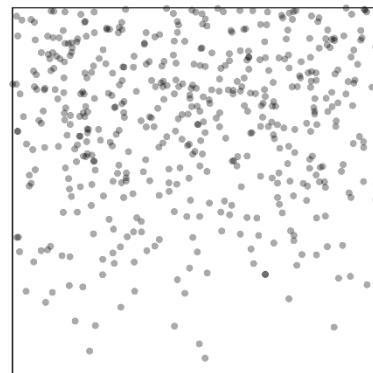


How many samples?

The power of **POWER ANALYSIS**

1. State what you want to estimate
2. Simulate a realistic version of what you want to estimate
3. Simulate sampling the population

Number of trees in area
within 20% of true number

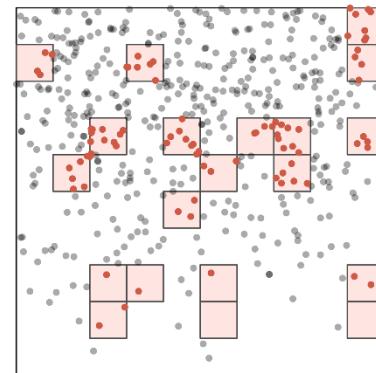
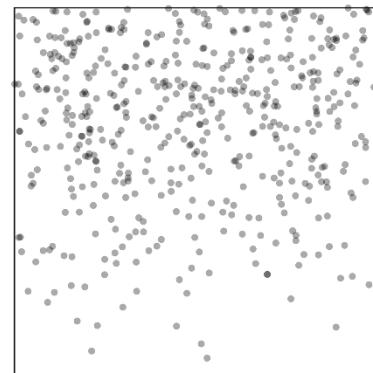


How many samples?

The power of **POWER ANALYSIS**

1. State what you want to estimate
2. Simulate a realistic version of what you want to estimate
3. Simulate sampling the population
4. Simulate analysing the sample

Number of trees in area
within 20% of true number



Number observed

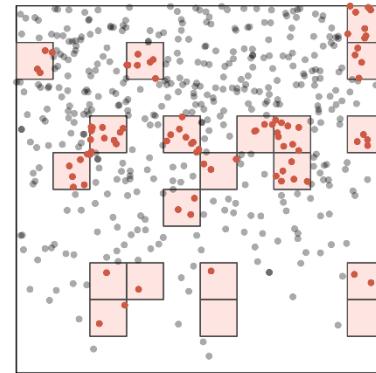
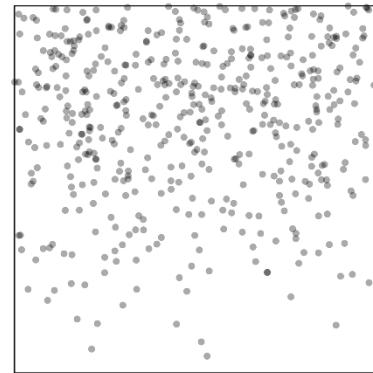
Proportion of area sampled

How many samples?

The power of **POWER ANALYSIS**

1. State what you want to estimate
2. Simulate a realistic version of what you want to estimate
3. Simulate sampling the population
4. Simulate analysing the sample
5. Calculate *how often* you meet your target

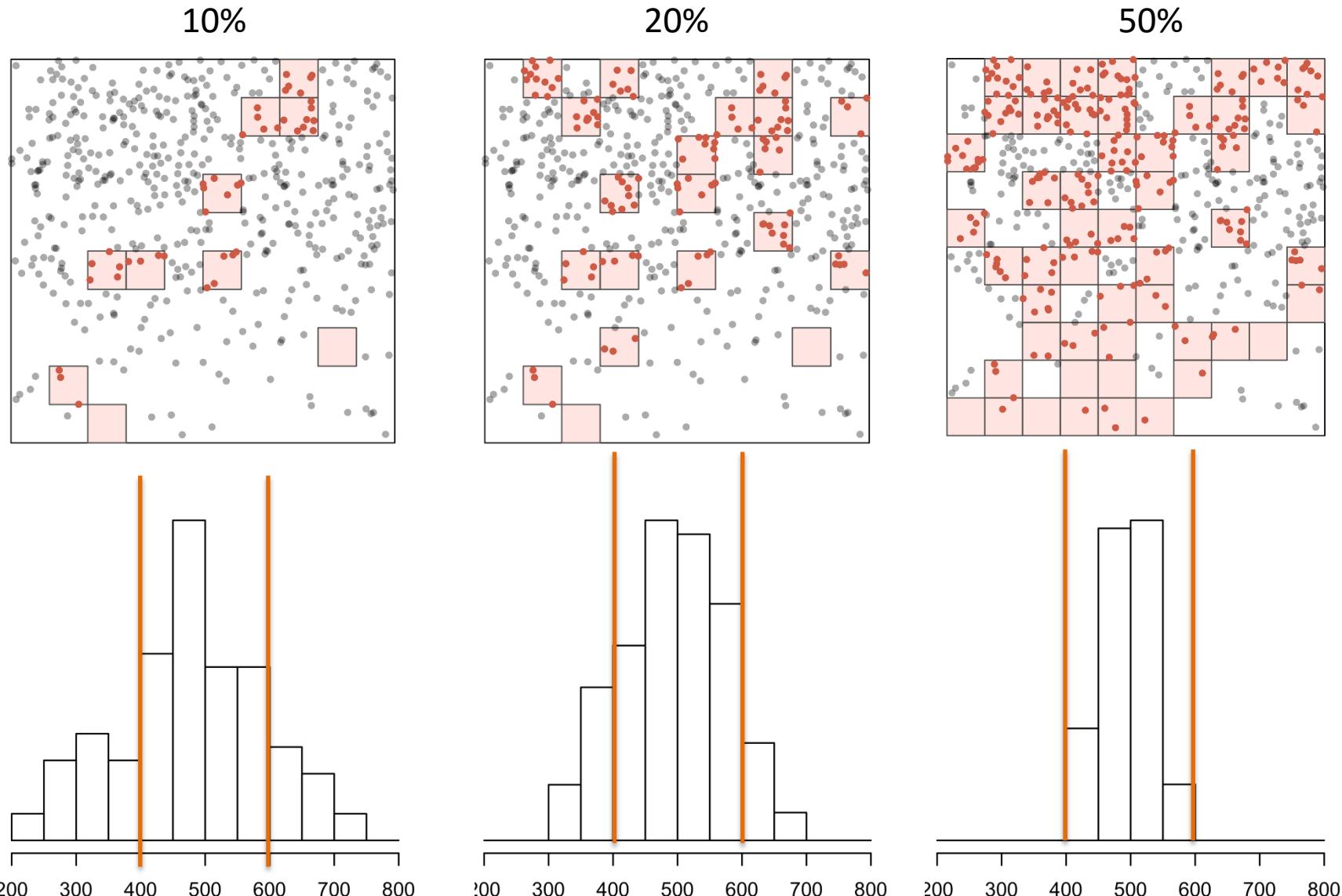
Number of trees in area within 20% of true number



Number observed

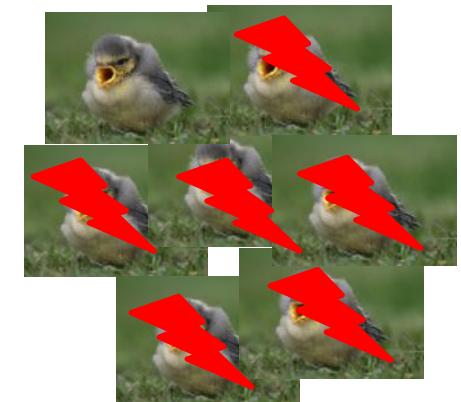
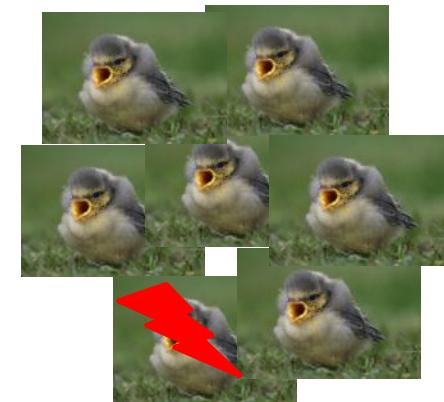
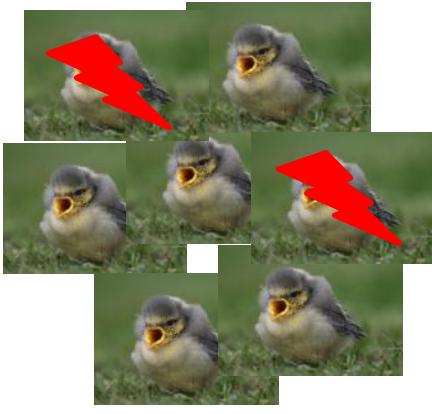
Proportion of area sampled

Estimating number of trees





What is the unit of independence?



	Extra food	No extra food
Survive to fledging	11	1
Dead	3	6

21 chicks in 3 nest boxes. What's the sample size?

Percentage data

10 x 10 grid in quadrat

```
1, 1, 1, 1, 1, 1, 1, 1, 1,  
1, 1, 1, 1, 1, 1, 1, 1, 1,  
1, 1, 1, 1, 1, 1, 1, 1, 1,  
1, 1, 1, 1, 1, 1, 1, 1, 1,  
1, 1, 1, 1, 1, 0, 0, 0, 0,  
0, 0, 0, 0, 0, 0, 0, 0, 0,  
0, 0, 0, 0, 0, 0, 0, 0, 0,  
0, 0, 0, 0, 0, 0, 0, 0, 0,  
0, 0, 0, 0, 0, 0, 0, 0, 0,
```

What's the sample size?



Survey design and statistical analysis for conservation science

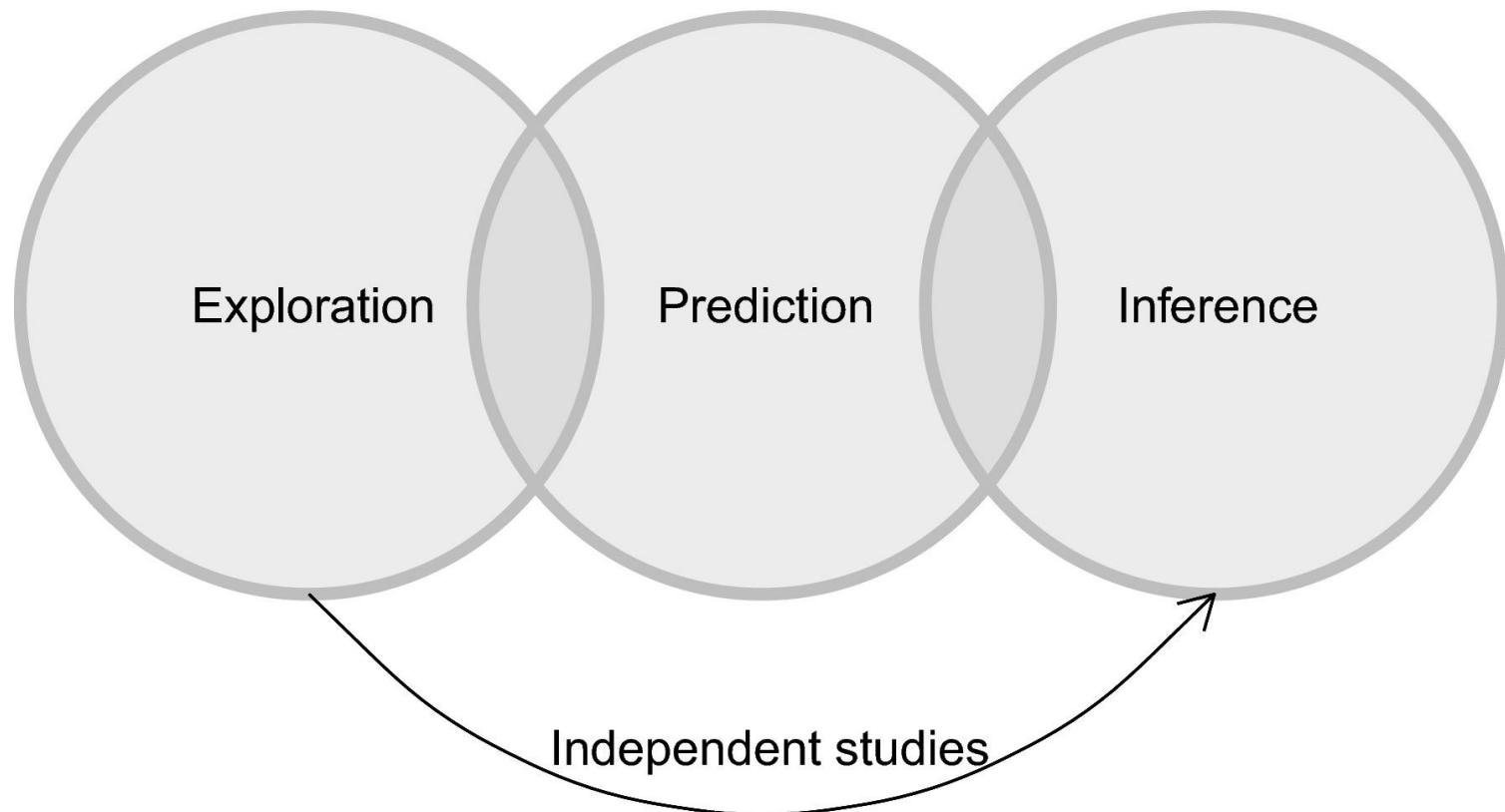
SURVEY DESIGN AND SAMPLING

1. REPRESENTATIVE samples
2. STRATIFICATION for sampling design
3. POWER ANALYSIS to assess sampling design
4. PSEUDO-REPLICATION in sampling

STATISTICAL ANALYSIS

COMMUNICATING RESULTS

STATISTICAL ANALYSIS



What is your ecological goal?

DESCRIBE a pattern

EXPLAIN a pattern

PREDICT an outcome

How many rhino horns are sold each year and where from?

How does habitat affect blue tit nesting success?

How is sea temperature affecting coral reefs?

What's the maximum fragmentation that a tropical frog species can tolerate?

How many tigers are in India?

What is your ecological goal?

DESCRIBE a pattern

EXPLAIN a pattern

PREDICT an outcome

How many rhino horns are sold each year and where from?

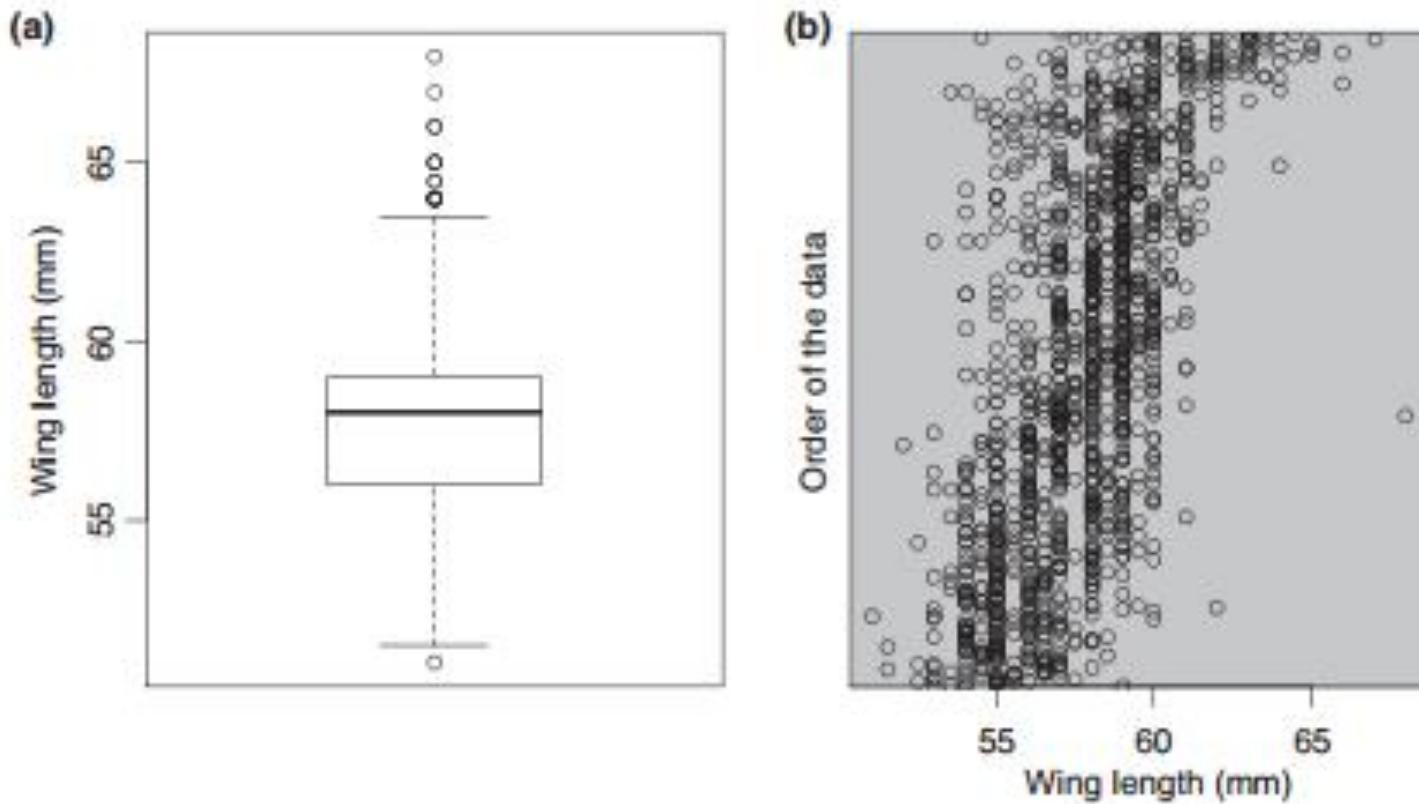
How does habitat affect blue tit nesting success?

How is sea temperature affecting coral reefs?

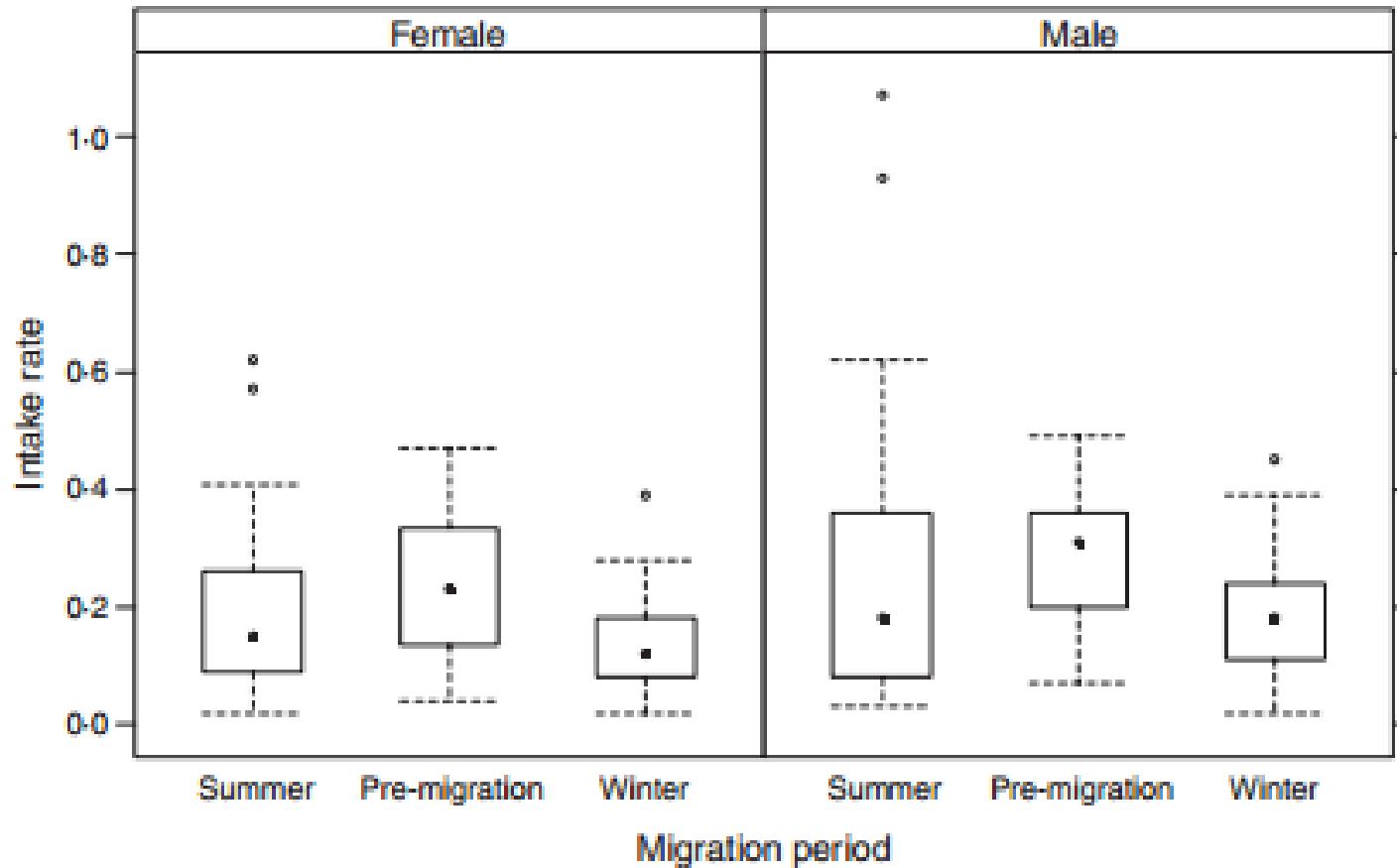
What's the maximum fragmentation that a tropical frog species can tolerate?

How many tigers are in India?

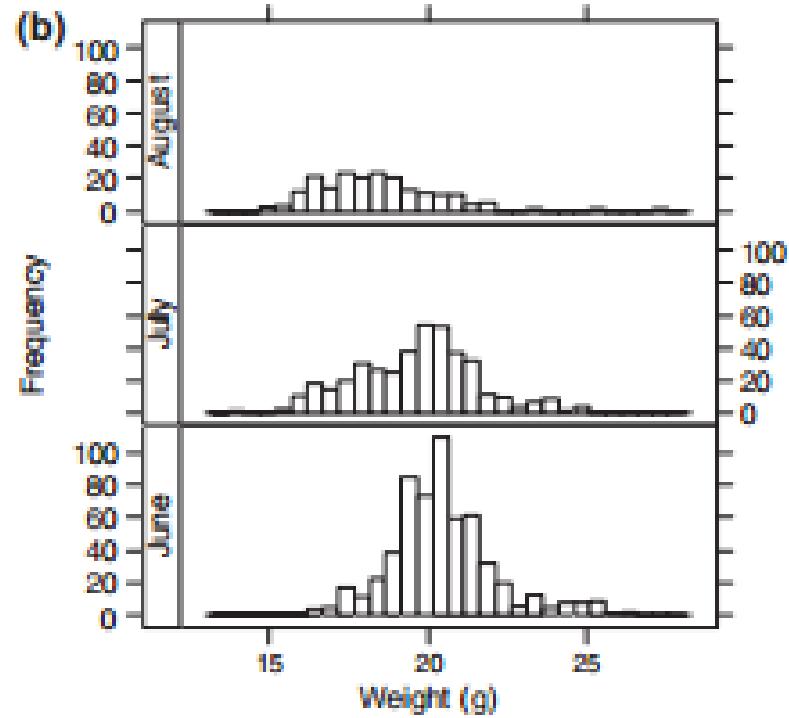
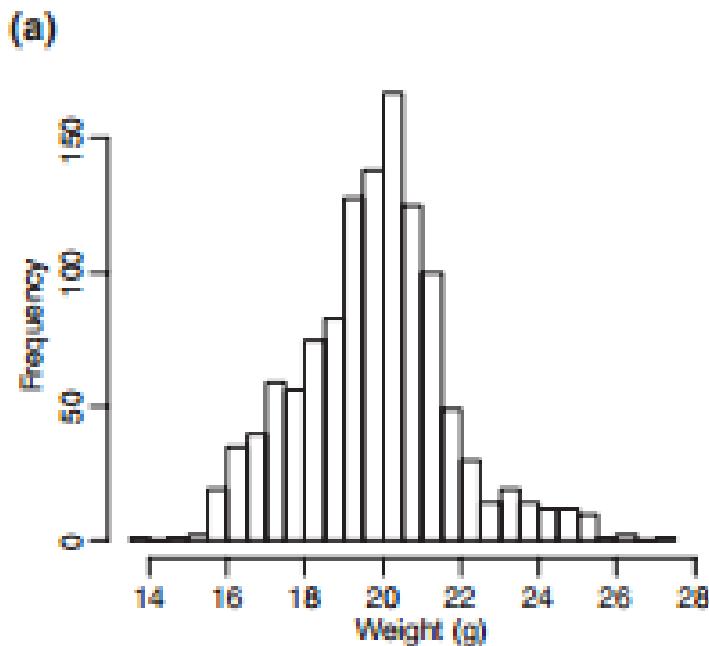
Look at your data!



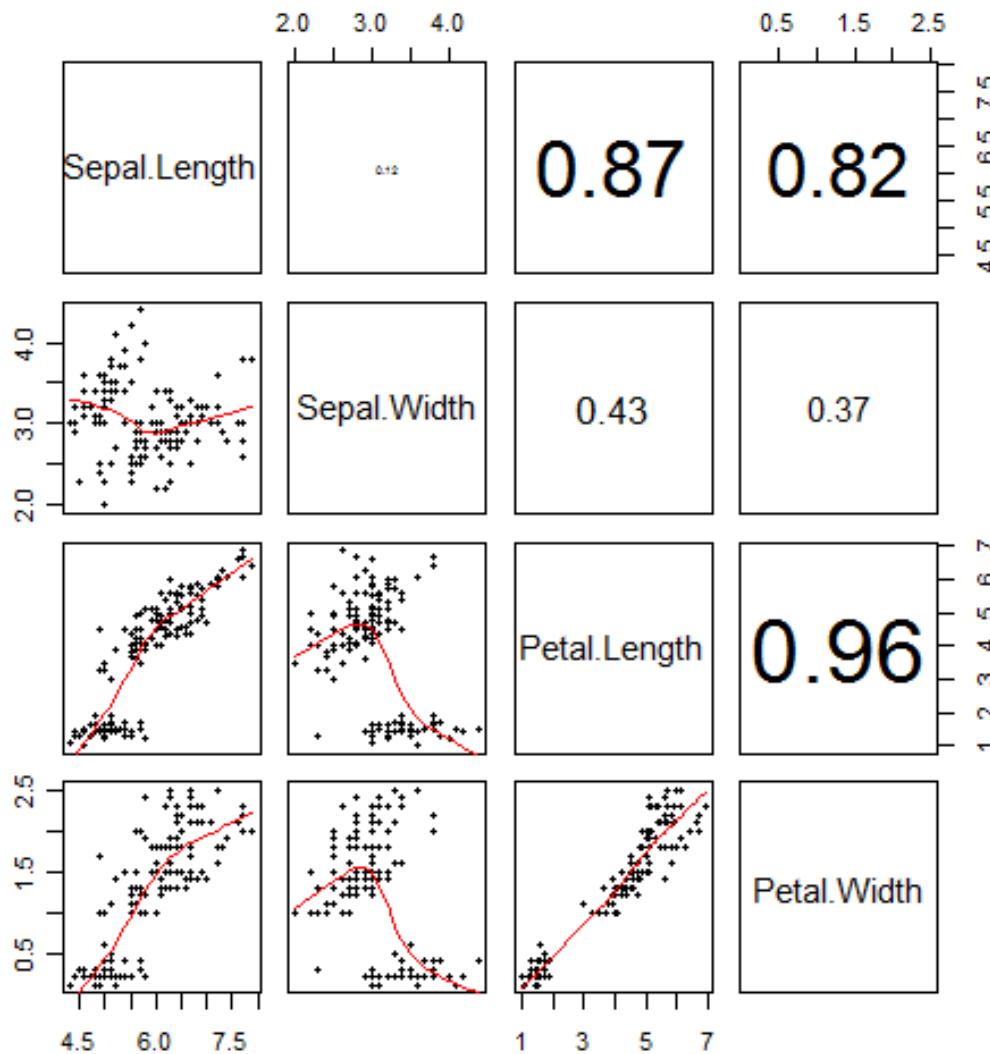
Look at your data!



Look at your data!



Iris Scatterplot Matrix



Inference - Which are significant?

	est	SE	Z	P-value	significance
Diurnal temp range	0.009693	0.0077	1.26	0.2069	
Precipitation	0.002372	0.0010	-2.49	0.086	
Temperature	0.546524	0.0650	8.41	<0.0001	***
Vapour pressure	-1.089962	0.1595	-3.834	0.0453	*
Frost days	-0.010056	0.0031	-6.786	0.0012	***
Aspect	1.54897	0.9873	-5.439	0.0556	



Statisticians issue warning over misuse of *P* values

Policy statement aims to halt missteps in the quest for certainty.

Jellies



Monya Baker

07 March 2016

Is the p-value pointless?

Posted March 16, 2016 by Lauren Richardson in Announcement, Biology, Community, Data, Debate, Education, News, Policy, Publishing, Research Resources

Opinion

Rewriting results sections in the language of evidence

Stefanie Muff ^{1,2,*}, Erlend B. Nilsen, ^{2,3,4,✉}, Robert B. O'Hara, ^{1,2,✉} and Chloé R. Nater ^{2,3,✉}

Despite much criticism, black-or-white null-hypothesis significance testing with an arbitrary *P*-value cutoff still is the standard way to report scientific findings. One obstacle to progress is likely a lack of knowledge about suitable alternatives. Here, we suggest language of evidence that allows for a more nuanced approach

Highlights

It has been known for decades that there are severe problems associated with null-hypothesis significance testing (NHST) based on arbitrary *P*-value

What is a p-value?

p-value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true

This is not the probability that the null hypothesis is true:

The probability of observing a result given that some hypothesis is true
is not equivalent to

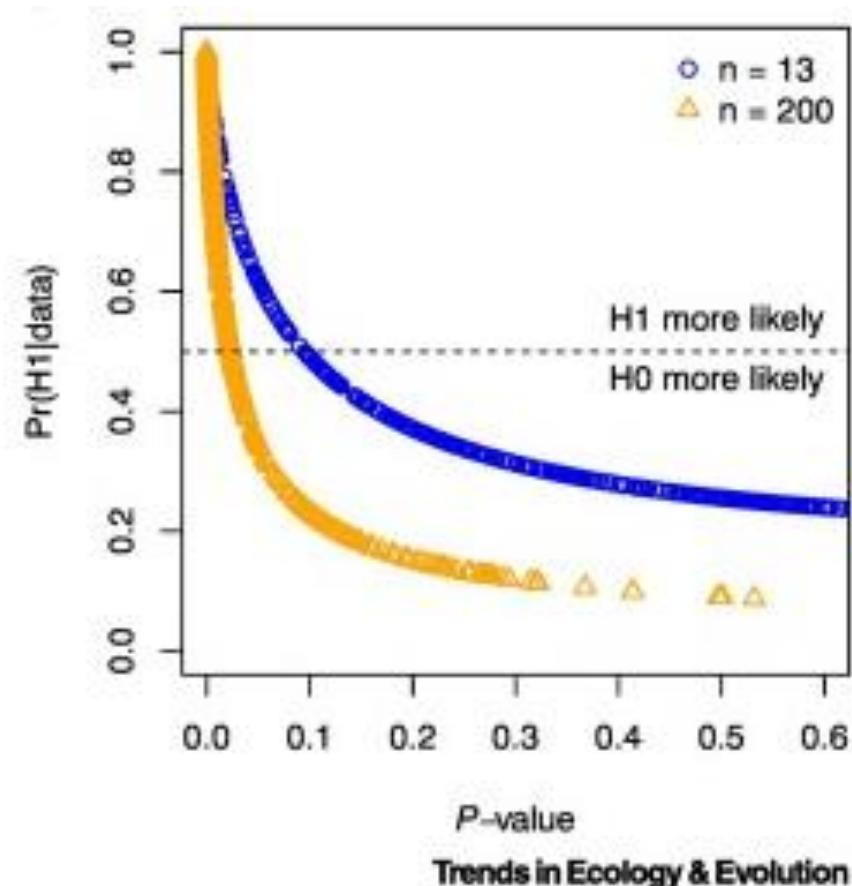
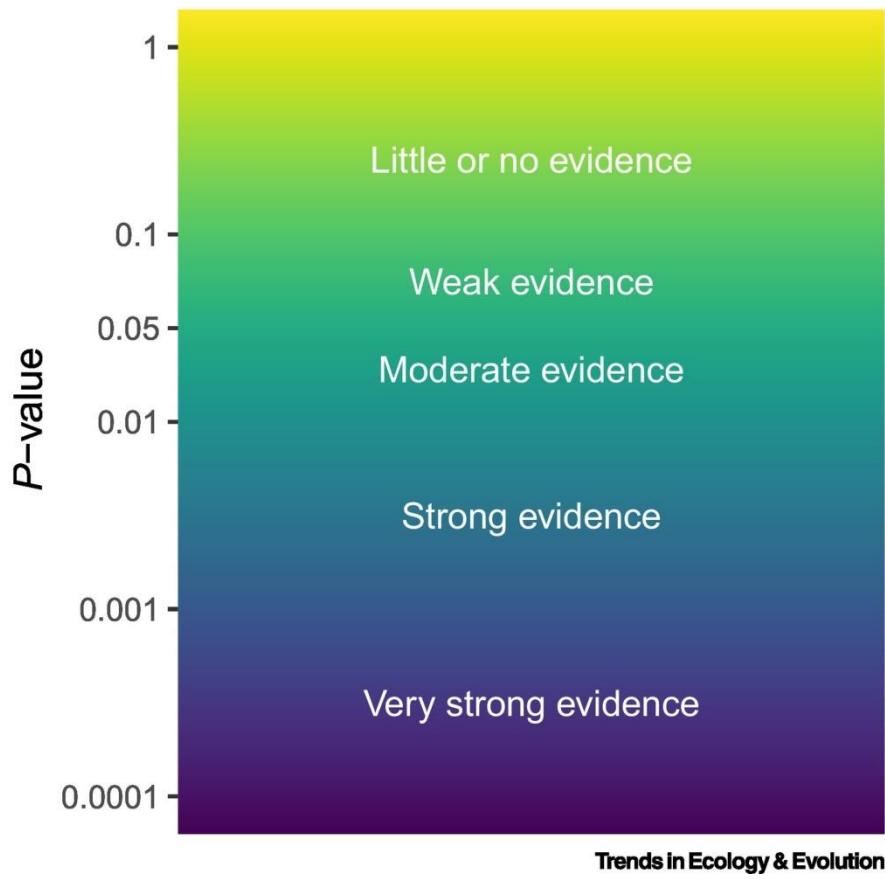
the probability that a hypothesis is true given that some result has been observed.

What is a p-value?

By itself, a *p*-value does not provide a good measure of evidence regarding a model or hypothesis

Whether and where to discretizing the p-value into significant/not-significant (e.g. at 0.05) is completely arbitrary!

Statistical significance is continuous, but also context dependent



3. *Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.*



COMMENT · 20 MARCH 2019

Scientists rise up against statistical significance

Valentin Amrhein, Sander Greenland, Blake McShane and more than 800 signatories call for an end to hyped claims and the dismissal of possibly crucial effects.

Valentin Amrhein , Sander Greenland & Blake McShane

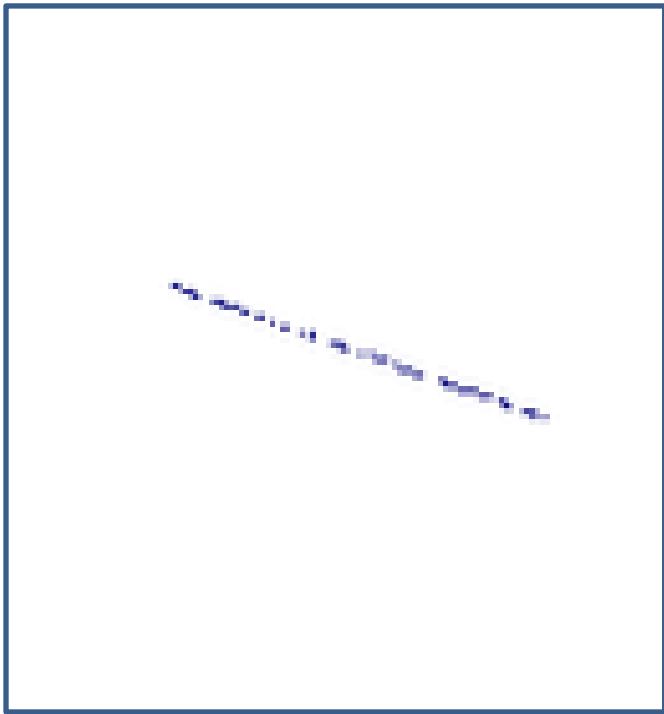
All biologists should be ultimately interested in **biological importance**, which may be assessed using the **magnitude of an effect**, but not its **statistical significance**.

It is almost always more useful for interpretation to look at / report **effect estimates and their uncertainty range** (e.g. confidence interval), than the p-value alone.

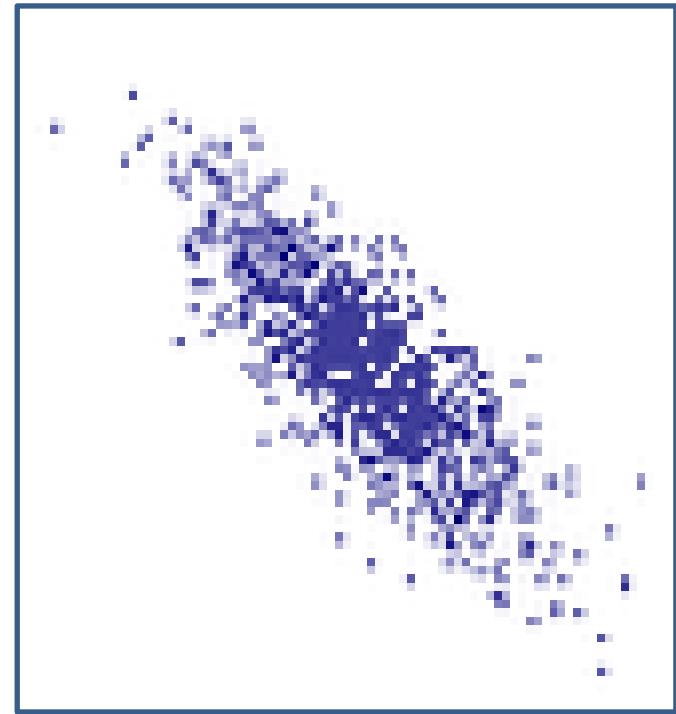
Nakagawa & Cuthill (2007)

Effect size, confidence interval and statistical significance: a practical guide for biologists
Biological Reviews 82(4):591 – 605. <https://doi.org/10.1111/j.1469-185X.2007.00027.x>

response

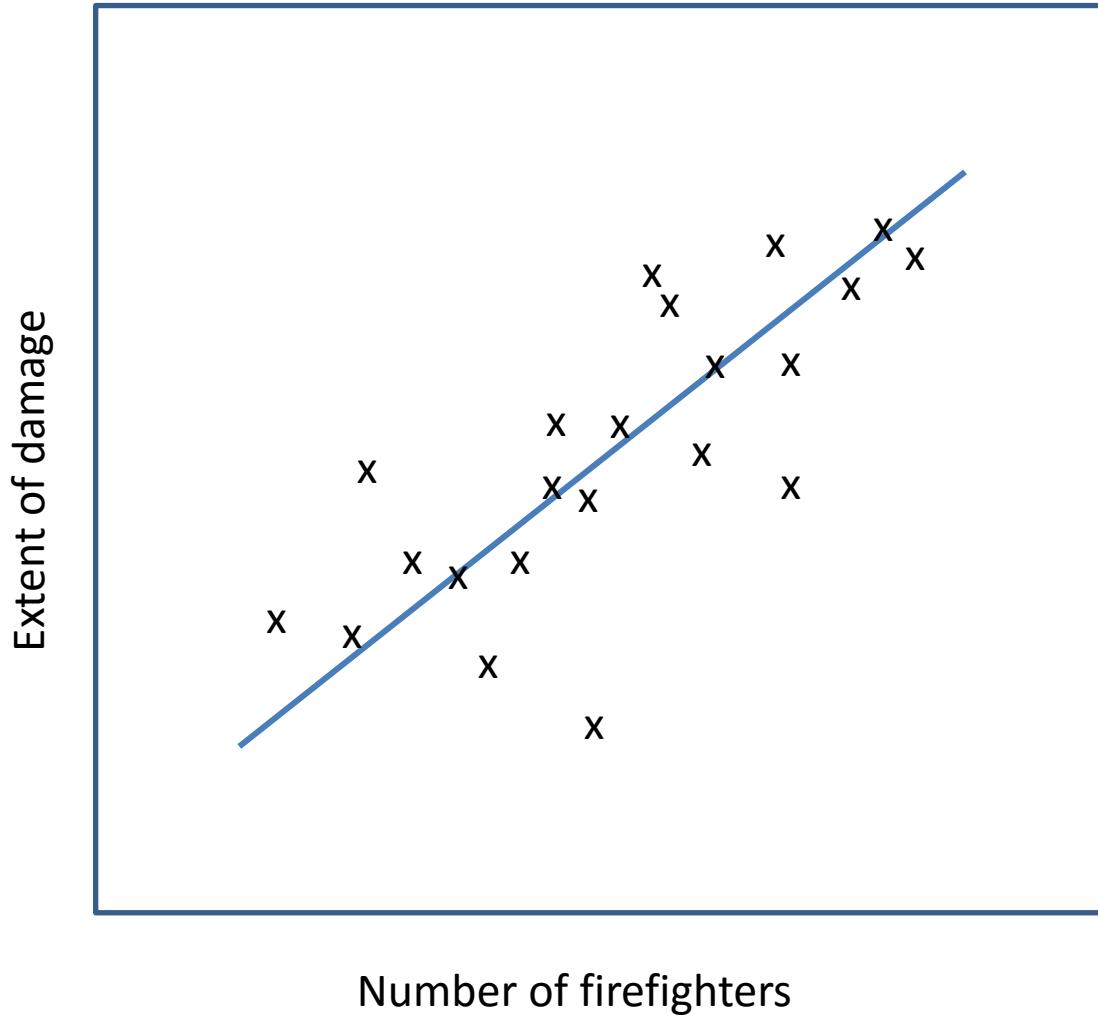


Variable A

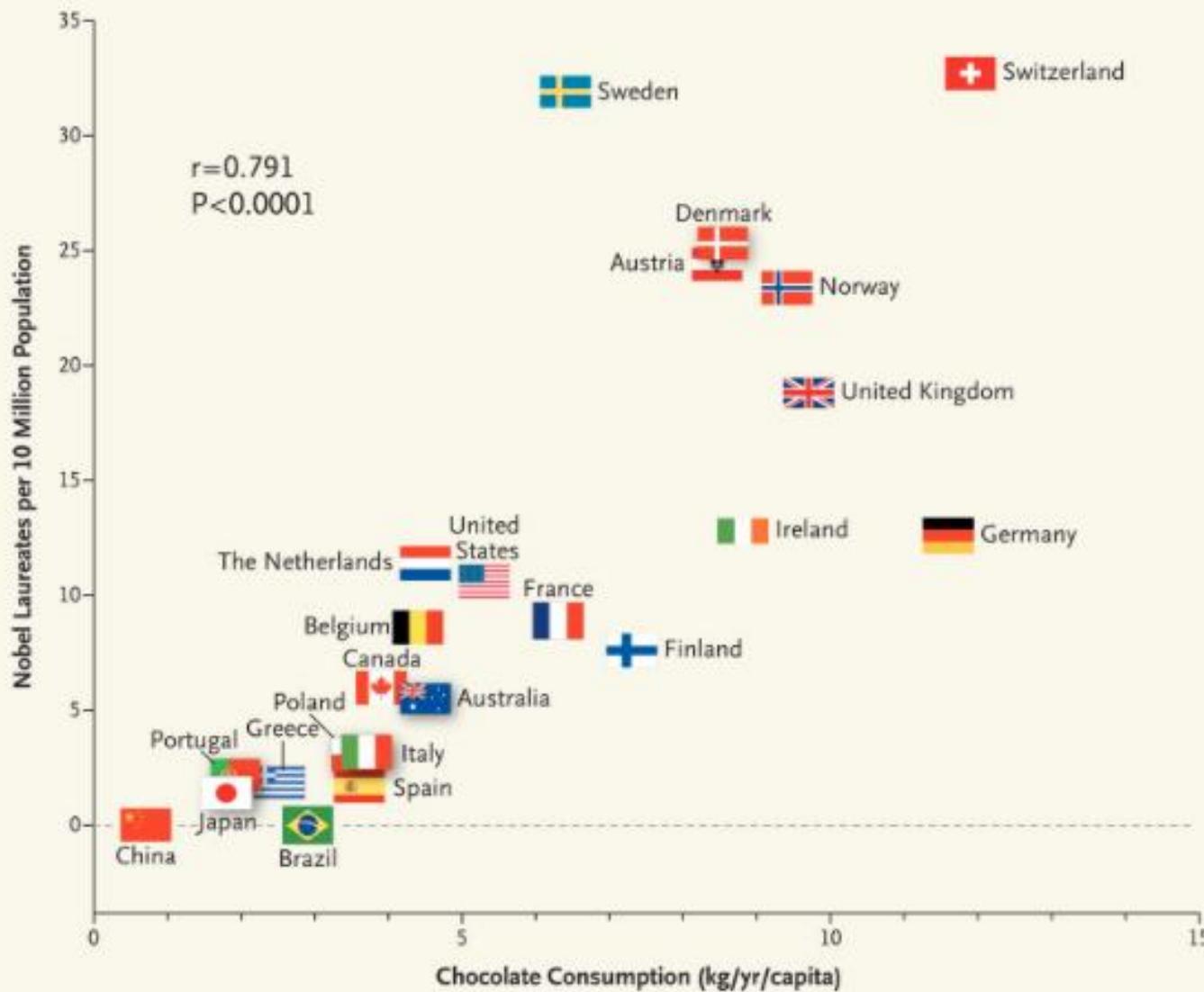


Variable B

correlation *is not equal to* causation

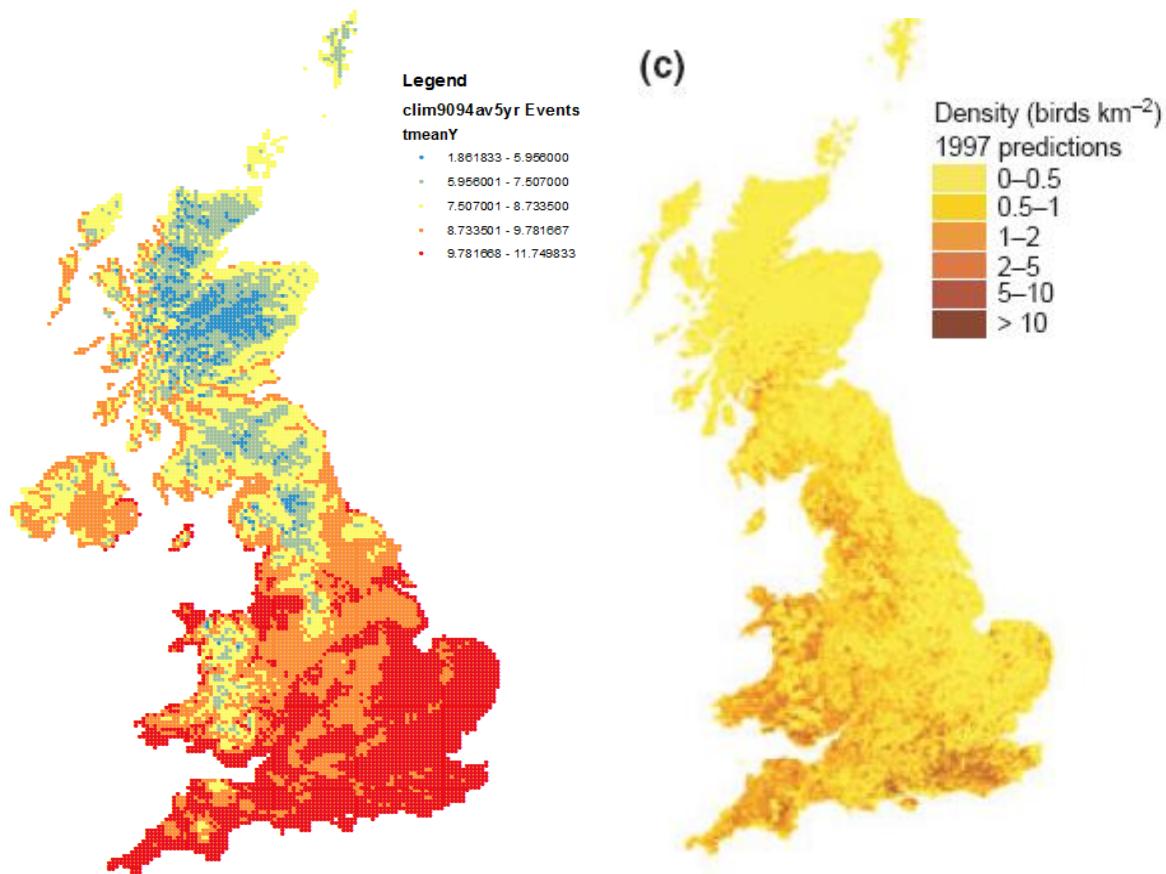


Conclusion: firefighters cause damage



"It remains to be determined whether the consumption of chocolate is the underlying mechanism for the observed association with improved cognitive function."

Climate envelope models



Renwick et al. (2011) *Diversity & Distributions*

Thinking about uncertainty

- How big is the uncertainty?
- How can you best communicate the uncertainty?
- What are the consequences of concluding the wrong thing?
- How much uncertainty are you comfortable with?
- How much uncertainty are policy makers comfortable with?

Survey design and statistical analysis for conservation science

SURVEY DESIGN AND SAMPLING

1. REPRESENTATIVE samples
2. STRATIFICATION for sampling design
3. POWER ANALYSIS to assess sampling design
4. PSEUDO-REPLICATION in sampling

STATISTICAL ANALYSIS

5. DESCRIPTION or EXPLANATION
6. SIGNIFICANCE is continuous
7. Separate SIGNIFICANCE and EFFECT SIZE
8. Separate CORRELATION AND CAUSATION
9. Consider UNCERTAINTY

COMMUNICATING RESULTS

COMMUNICATING ANALYSES AND RESULTS

- Methods:
 - Always fully describe the analysis you are conducting
 - Write out the exact model you are fitting
 - Provide details on the software you are using
 - If you need to be terse in the main manuscript, put a full description in an appendix/supplement

COMMUNICATING ANALYSES AND RESULTS

Name & cite the software

Analysis

All statistical analyses were carried out in *R v. 3.1* (R Development Core Team, 2013). We used generalized linear models with Poisson-distributed errors to estimate the species-area relationship by predator status for all islands (Supplementary Material 4). The models of species richness and island area assume there is a linear increase in the number of species as island area increases but in reality this is limited to the number of seabird species in the region. We assume that the maximum species richness is constrained by the number of species included in the analysis ($n = 16$). Model assumptions of a linear mean variance relationship and of spatially uncorrelated errors were checked using the dispersion test of the *AER* package in *R* (Kleiber & Zeileis, 2008), and spatial correlation tests of the *sp*, *ape* and *gstat* packages (Supplementary Material 4; Paradis et al., 2004; Pebesma, 2004; Bivand et al., 2013).

Describe the model structures fully

We considered three models for the species-area relationship by predator status, all being Poisson regressions of a semilog model (i.e. species count, S , as response, $\log_{10}(A)$ as predictor). The three models are all Poisson generalized linear models: $S \sim \text{Poisson}(\mu)$ with log link, such that $\log \mu(S) = \beta X_p$, where X and β are the design matrix and the parameter vector, respectively.

Effect of $\log(\text{area})$ only:

$$\log \mu(S) = \beta_0 + \beta_1 \log_{10} A \quad (1)$$

Additive effects of $\log(\text{area})$ and predator status (i.e. separate intercept per predator status but same slope):

$$\log \mu(S) = \beta_0 + \beta_1 \log_{10} A + \beta_2 \text{Predator} \quad (2)$$

Interaction of $\log(\text{area})$ and predator status (i.e. separate slope and intercept per predator status):

$$\log \mu(S) = \beta_1 \text{Predator} + \beta_2 \log_{10} A \times \text{Predator} \quad (3)$$

Describe the model and its assumptions (using words)

Equations or pseudocode can be less ambiguous than words

COMMUNICATING ANALYSES AND RESULTS

- Results:
 - Always provide full model summary tables (if need be in an appendix)
 - BUT key to a good paper/report is translating statistical summaries into human language and images

COMMUNICATING RESULTS

- Always provide full model summary tables (if need be in an appendix)

TABLE 2 Generalized linear model parameter estimates for the species-area relationship by predator status. Parameter values are given on the link scale.

Parameter	Parameter estimate \pm SE	P
Predator_statusCleared	1.462 ± 0.160	< 0.001
Predator_statusInvaded	-0.480 ± 0.285	0.092
Predator_statusUninvaded	0.890 ± 0.156	< 0.001
log(ha):Predator_statusCleared	0.038 ± 0.036	0.295
log(ha):Predator_statusInvaded	0.249 ± 0.096	0.009
log(ha):Predator_statusUninvaded	0.259 ± 0.062	< 0.001

COMMUNICATING RESULTS

Restate the purpose / ecological question of the analysis

TABLE 2 Generalized linear model parameter estimates for the species-area relationship by predator status. Parameter values are given on the link scale.

Parameter	Parameter estimate \pm SE	P
Predator_statusCleared	1.462 ± 0.160	< 0.001
Predator_statusInvaded	-0.480 ± 0.285	0.092
Predator_statusUninvaded	0.890 ± 0.156	< 0.001
log(ha):Predator_statusCleared	0.038 ± 0.036	0.295
log(ha):Predator_statusInvaded	0.249 ± 0.096	0.009
log(ha):Predator_statusUninvaded	0.259 ± 0.062	< 0.001

P-values / significance measures

Use human-readable descriptions of your variables

Parameter estimates and their uncertainties

COMMUNICATING RESULTS

- BUT good communication means translating statistical summaries into human language!

Translate parameter estimates back to words about ecology!

Restate the model

The final model for the species-area relationship included separate slopes and intercepts by predator status. No significant overdispersion was detected (overdispersion test: dispersion parameter = 0.99; $P = 0.53$). Moran's I ($I = 0.1$, $P = 0.011$) indicated weak but statistically significant spatial autocorrelation in the residuals, and visual inspection of spatial residuals indicated that this was because the model overpredicted species richness for the islands of the inner Hauraki Gulf (Supplementary Fig. S1). Generalized linear model parameter estimates are in Table 2, and model predictions are illustrated in Fig. 2.

The final model for the species-area relationship included separate slopes and intercepts by predator status. No significant overdispersion was detected (overdispersion test: dispersion parameter = 0.99; $P = 0.53$). Moran's I ($I = 0.1$, $P = 0.011$) indicated weak but statistically significant spatial autocorrelation in the residuals, and visual inspection of spatial residuals indicated that this was because the model overpredicted species richness for the islands of the inner Hauraki Gulf (Supplementary Fig. S1). Generalized linear model parameter estimates are in Table 2, and model predictions are illustrated in Fig. 2.

Describe how you checked the model

For uninvaded islands the intercept (i.e. expected species count on a 1 ha island) is 2.43 (95% CI 1.77–3.28; $\beta = 0.89$, $P < 0.001$), and for a 10-fold increase in area the species count is expected to increase 1.81-fold (95% CI 1.36–2.39; $\beta = 0.26$, $P < 0.001$). For cleared islands the intercept is 4.31 (95% CI 3.12–5.84; $\beta = 1.46$, $P < 0.001$), and no significant relationship was found between species richness and area ($\beta = 0.04$, $P = 0.295$). For invaded islands the intercept is 0.62 (95% CI 0.34–1.05; $\beta = -0.48$, $P = 0.092$), and for a 10-fold increase in area the species count is expected to increase 1.77-fold (95% CI 1.14–2.71; $\beta = 0.25$, $P = 0.009$). Regardless of size, invaded islands did not have more than four species (Fig. 2).

Estimate and CI on interpretable scale are more useful than raw coefficients/p-values

COMMUNICATING RESULTS

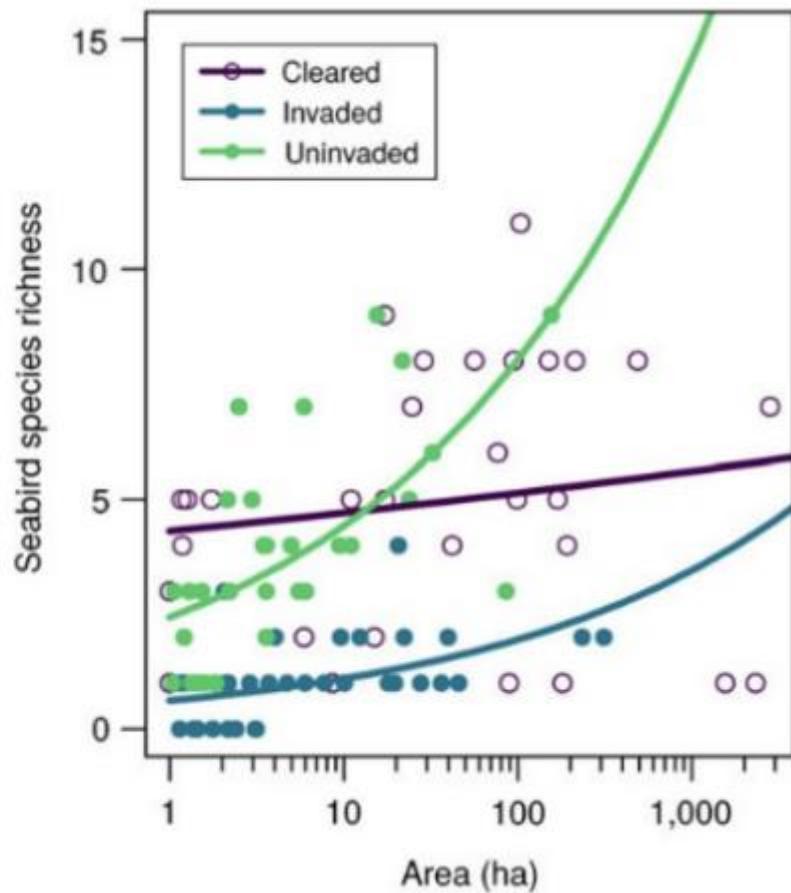
– or images!

You can use model predictions to illustrate the key take-aways from your model

Label Axes and Symbols clearly
Use simple colours

Use the caption to summarise the findings.
This could be improved here, e.g.:
“Species richness increases with increasing area, and is higher on uninvaded and cleared islands”

FIG. 2 Observed species richness and generalized linear model predictions as a function of island area for uninvaded (intercept = 2.43, slope = 0.26), cleared (intercept = 4.31, slope = 0.04), and invaded islands (intercept = 0.62, slope = 0.25) in the Hauraki Gulf (Fig. 1). Model parameter estimates are in Table 2.



COMMUNICATING RESULTS: GRAPHS

What is the **key** message you want the graph to convey?

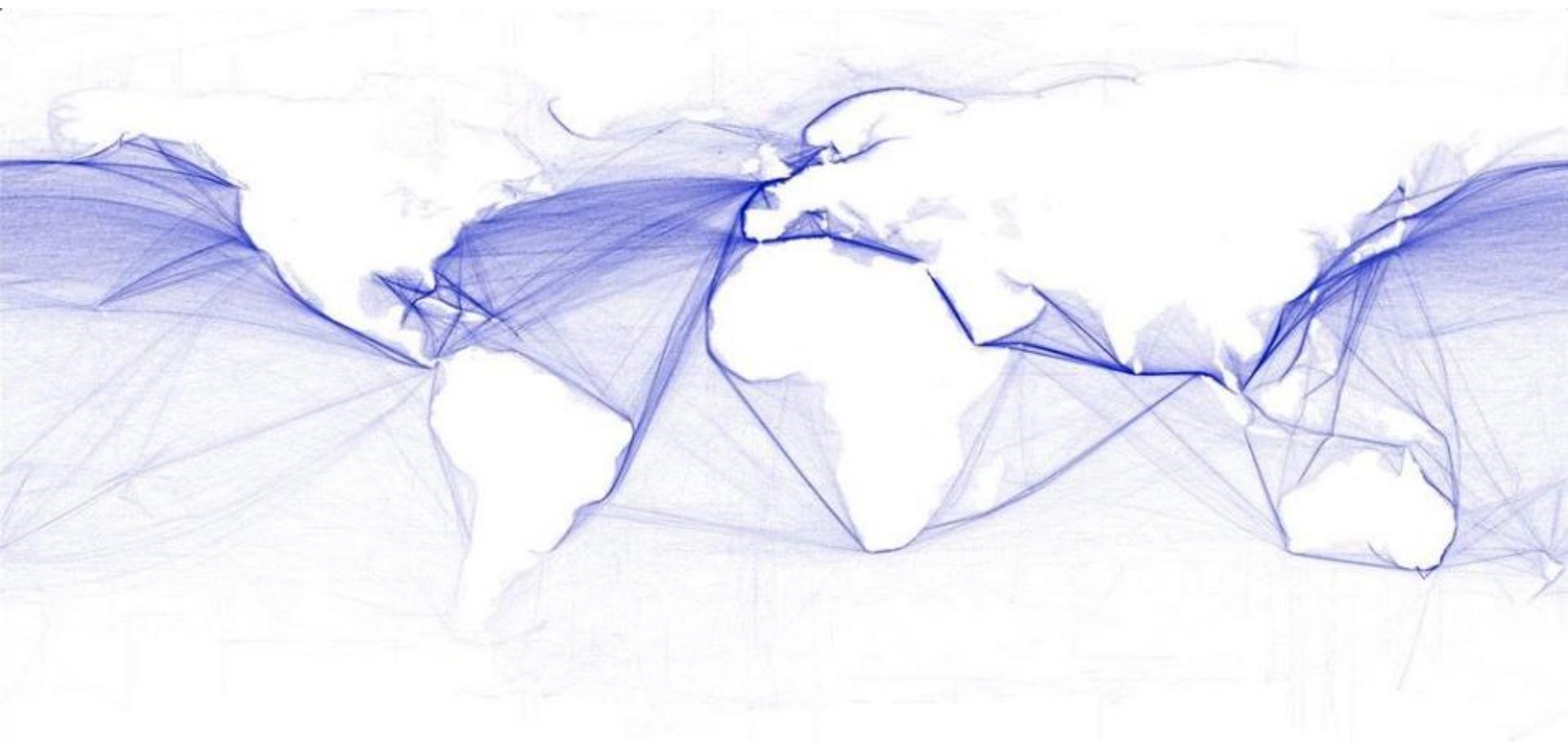
What is the **simplest** way of communicating that message?

Is the end graph intuitive to understand?

What might distract readers from the key message?

Add the data where possible

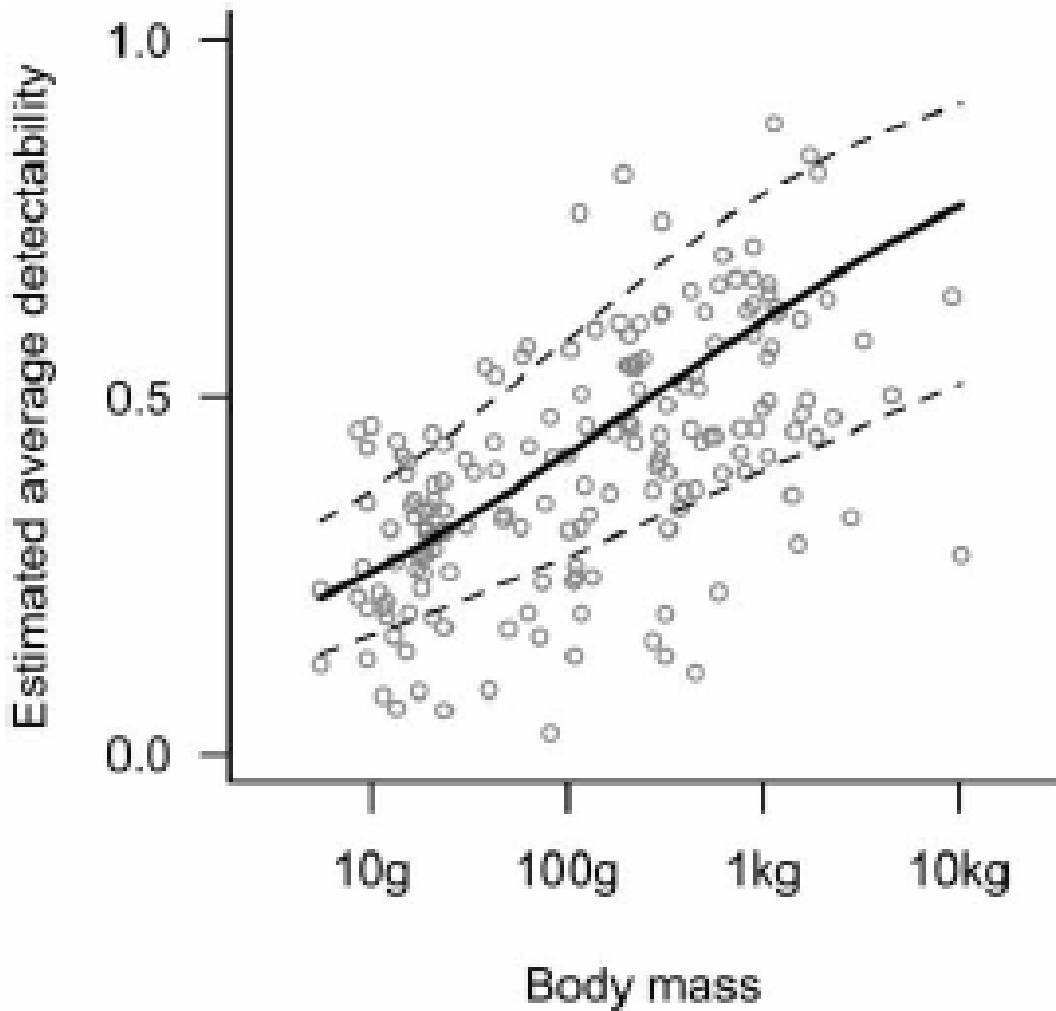
Sometimes the raw data is all you need:
World shipping routes



(but most of the time we need
to work a bit harder...)

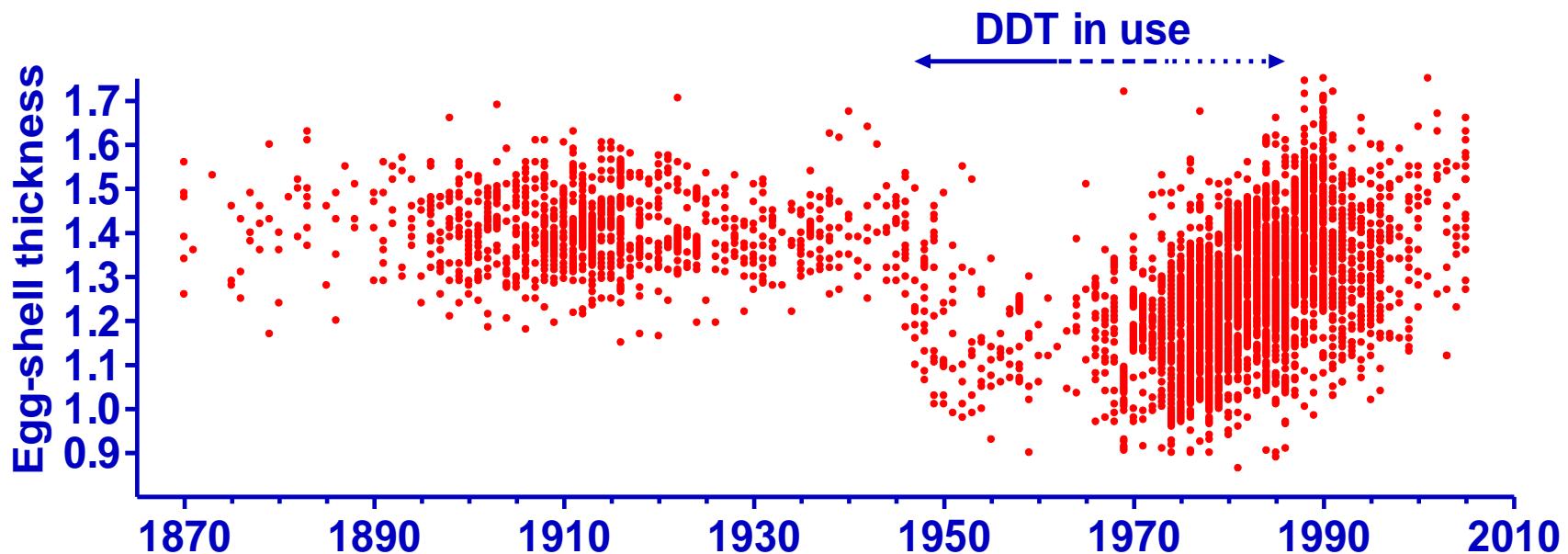


- Use clear labels
- Include units
- Show the data
- Show the model
- Show uncertainty

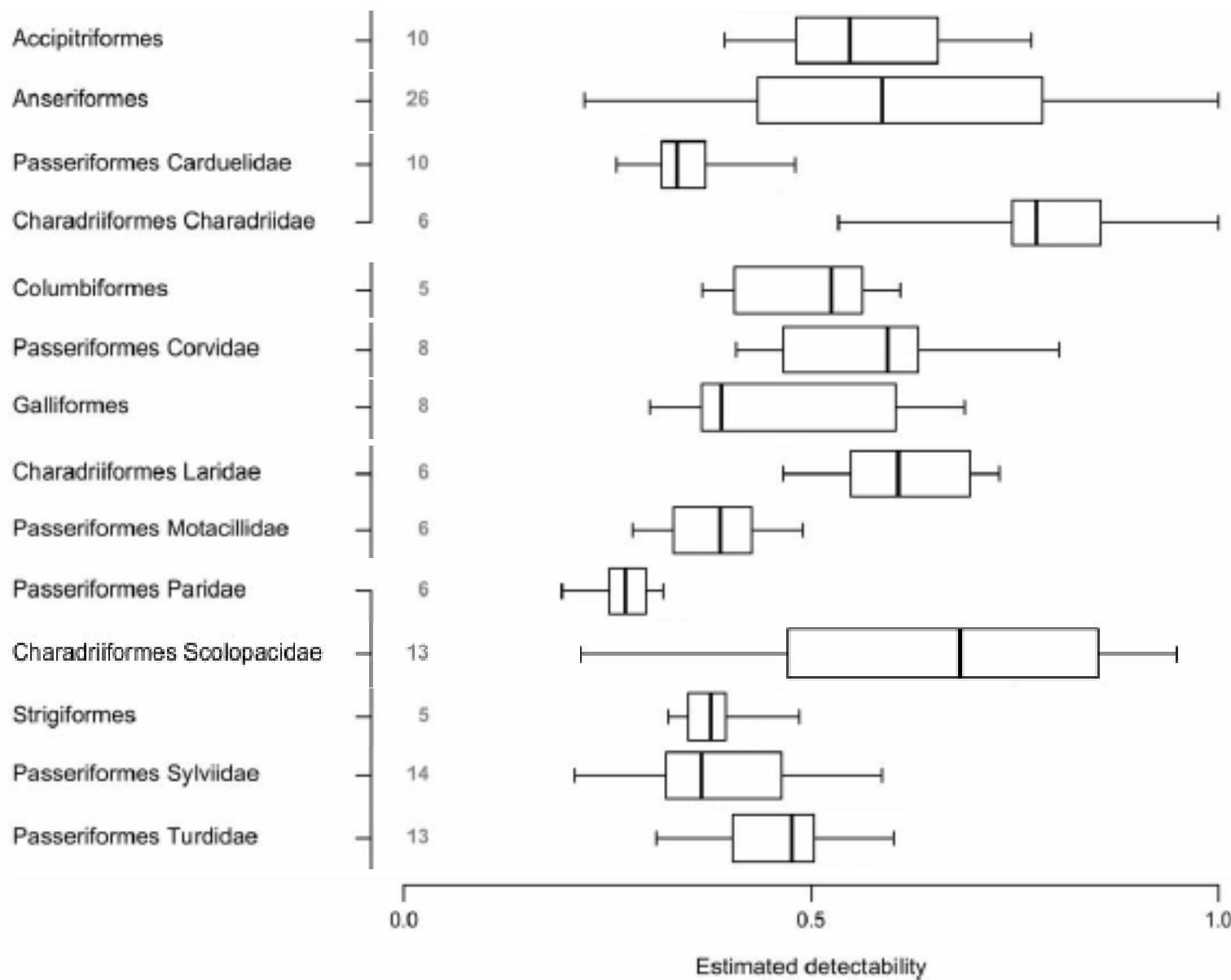


Using annotations to add context:

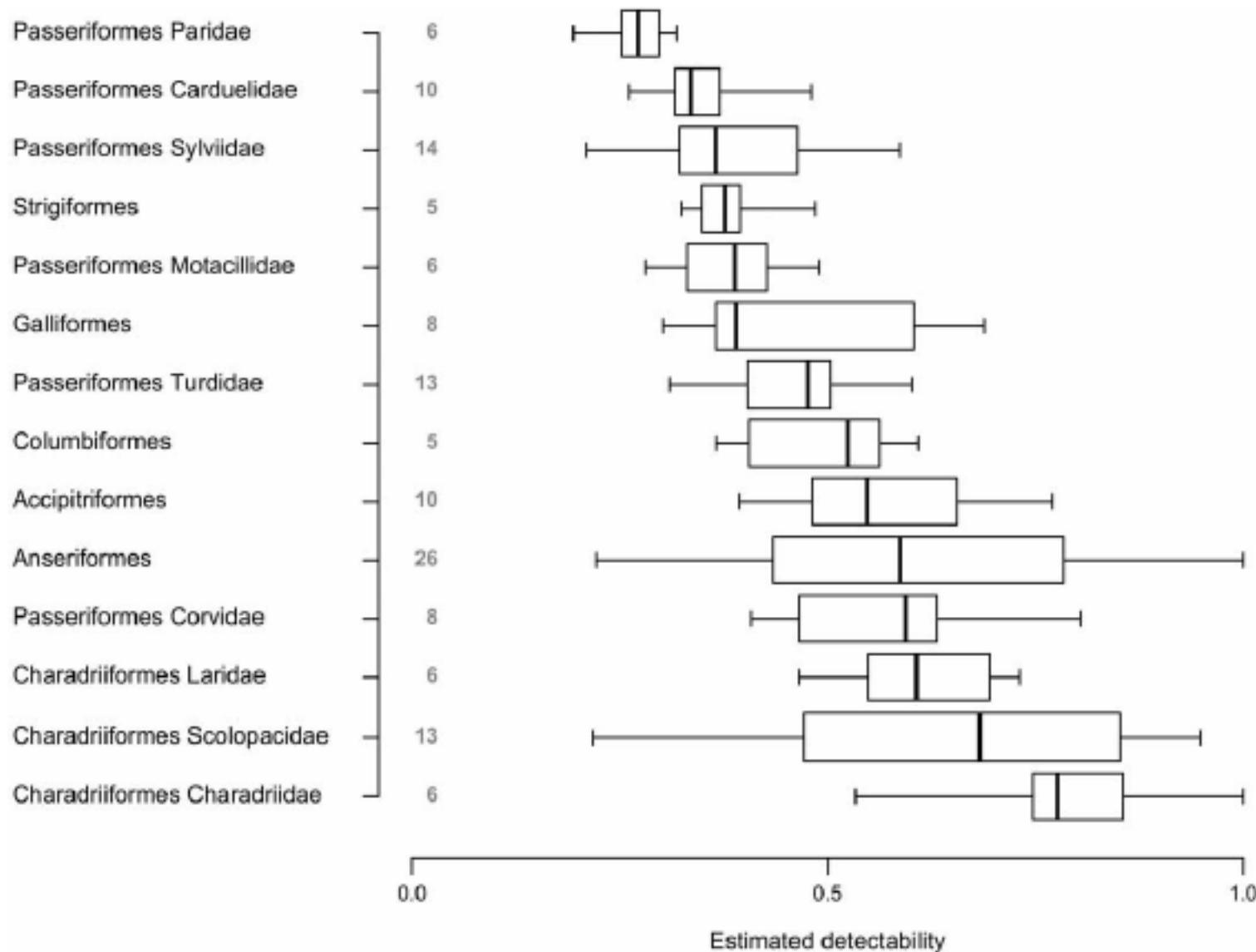
Sparrowhawk eggs and pesticide contamination



Thinking about composition: Ordering of categories can be important



Thinking about composition: Ordering of categories can be important



Colours

Select **intuitive** colours to communicate your message

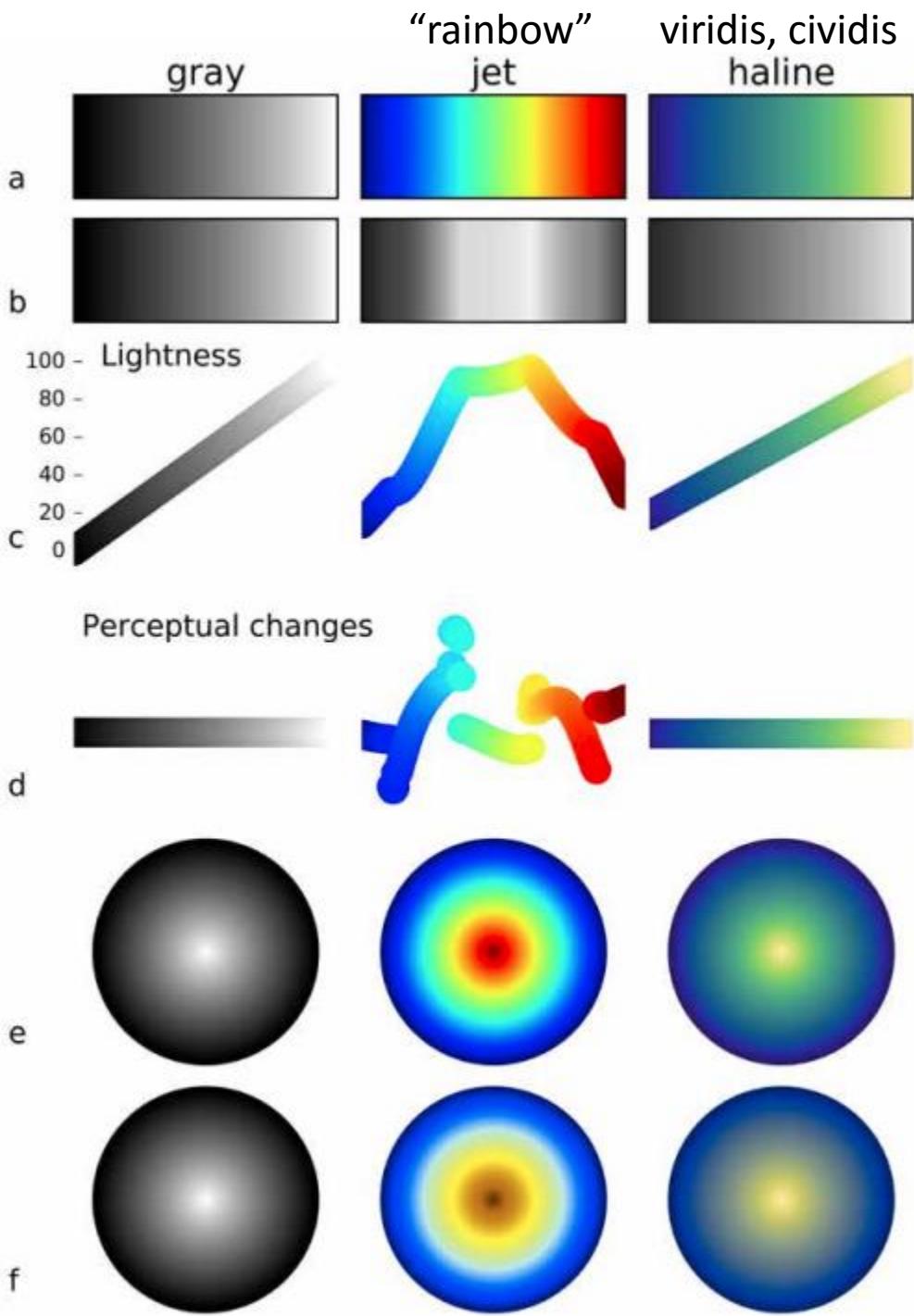
Use the **simplest** colour scheme that communicates the story

Think about the data:
Continuous or Categorical?
Ordered? Diverging?

Think about the medium:
print vs screen, photocopying?

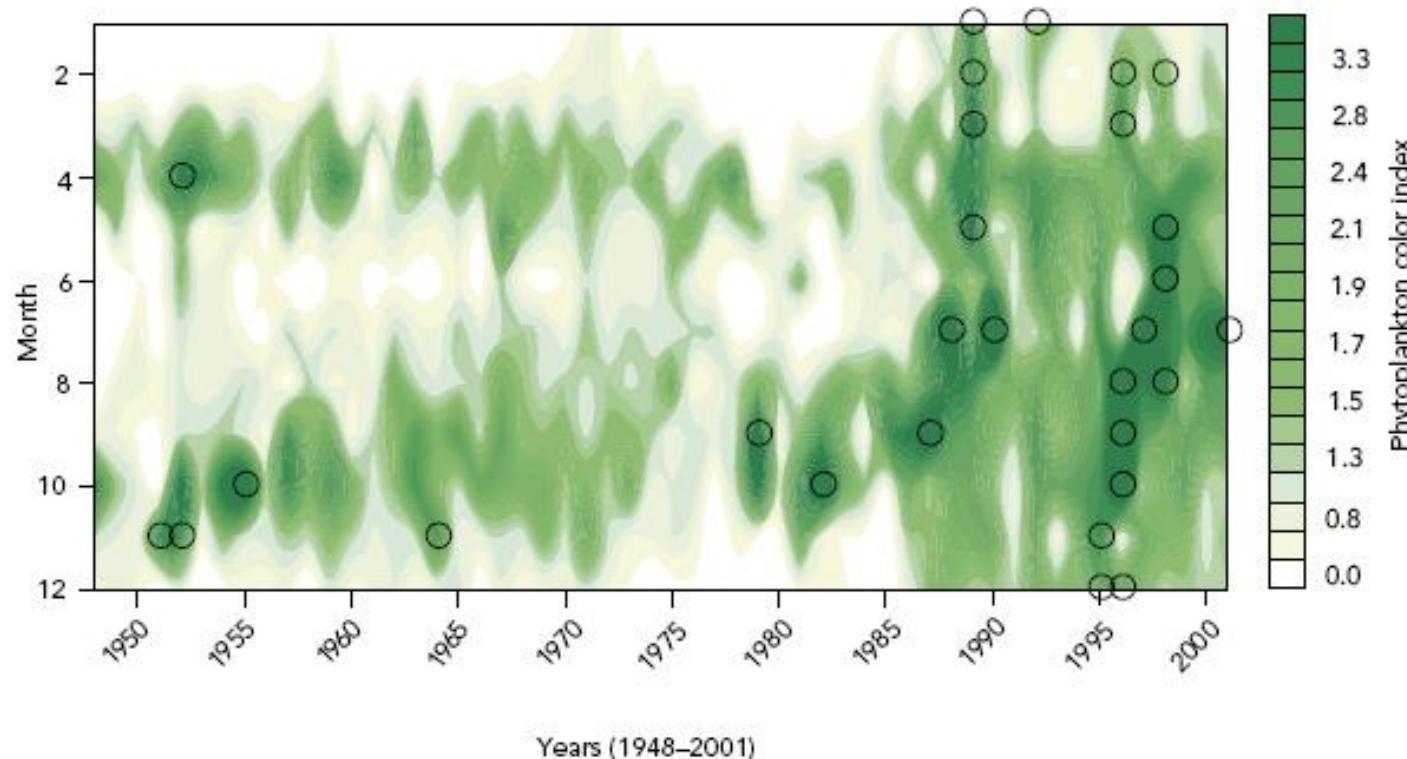
Think about the audience:
Colour blindness?

Tool for good palettes:
<https://colorbrewer2.org/>



Using colour as the third dimension: Changing timing of Phytoplankton bloom

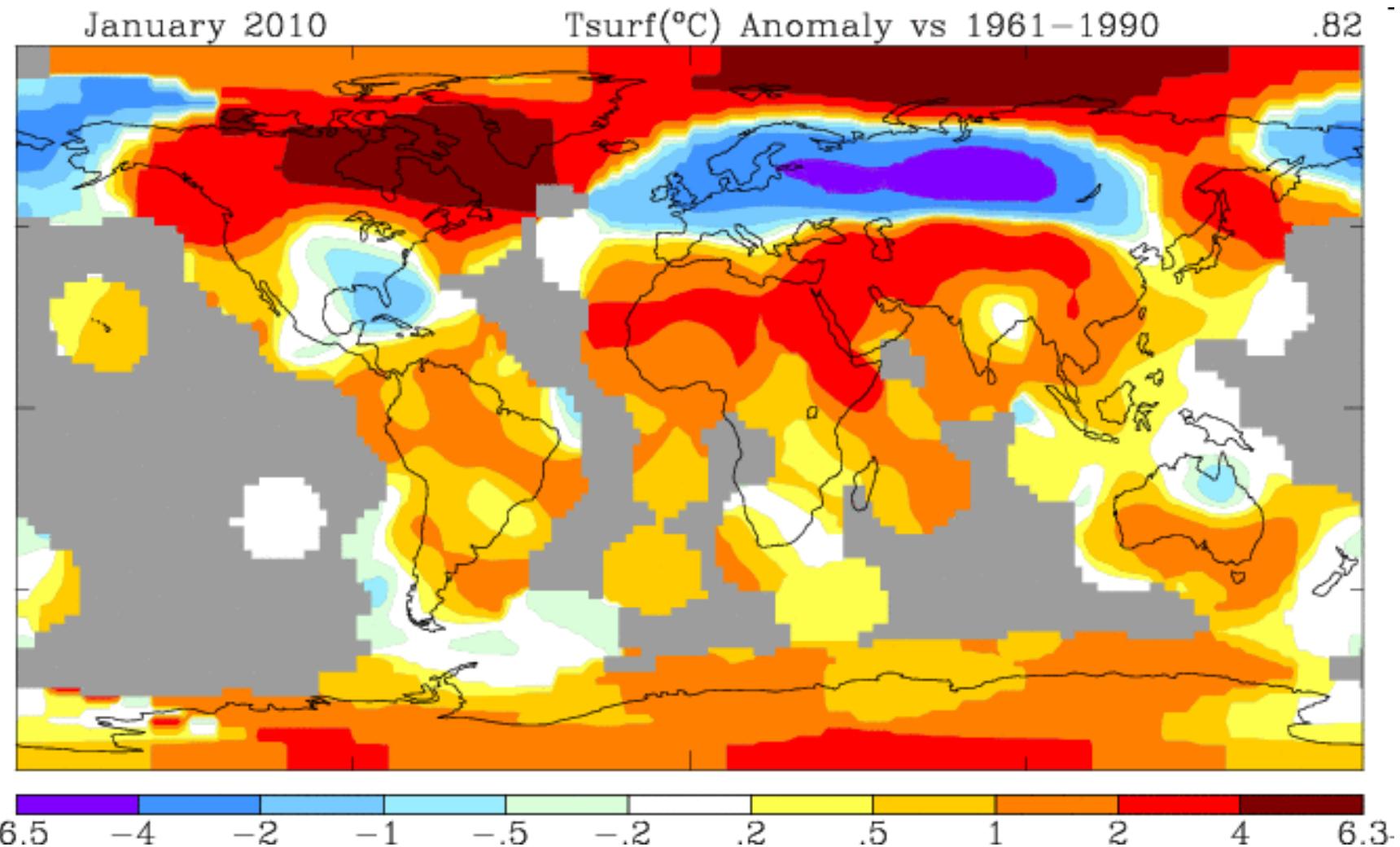
Single colour gradient -> darker = more



Source: M. Edwards, (SAHFOS), 2003.

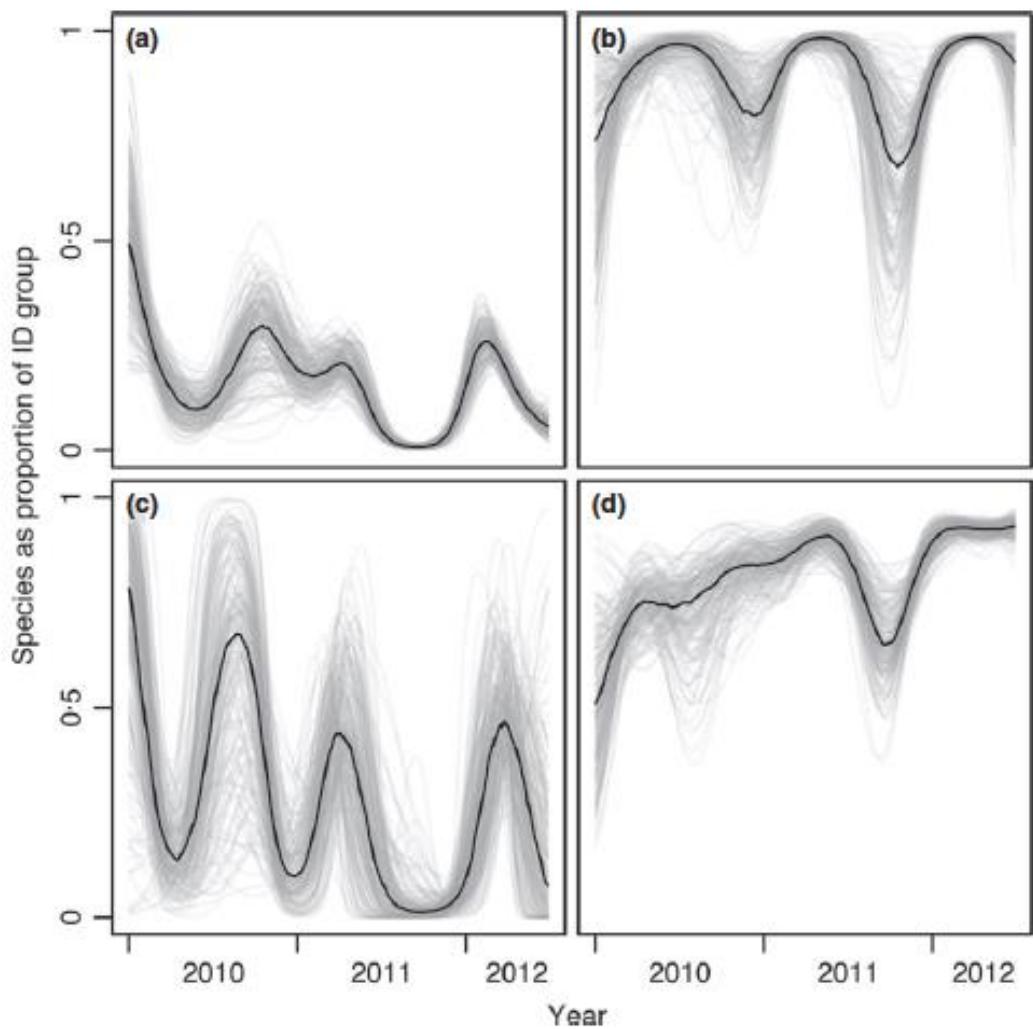
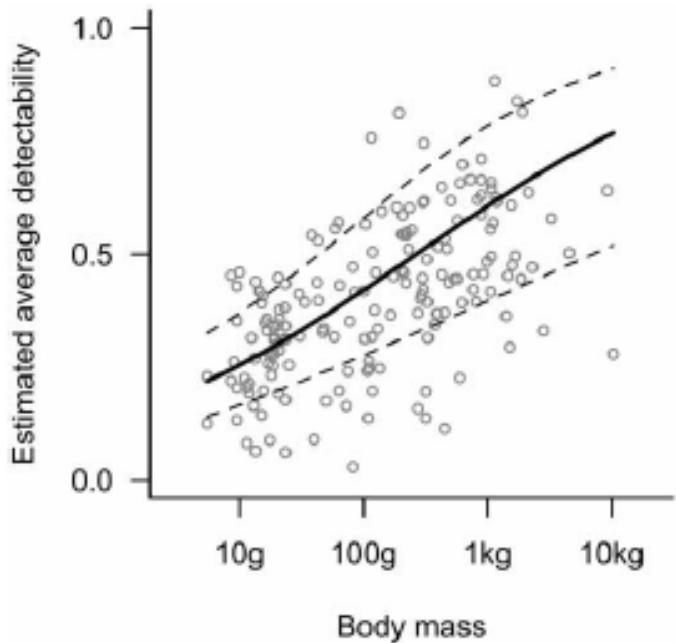
EEA Report 2004/2

**Using colour as the third dimension:
Diverging colour schemes highlight deviations from a baseline**

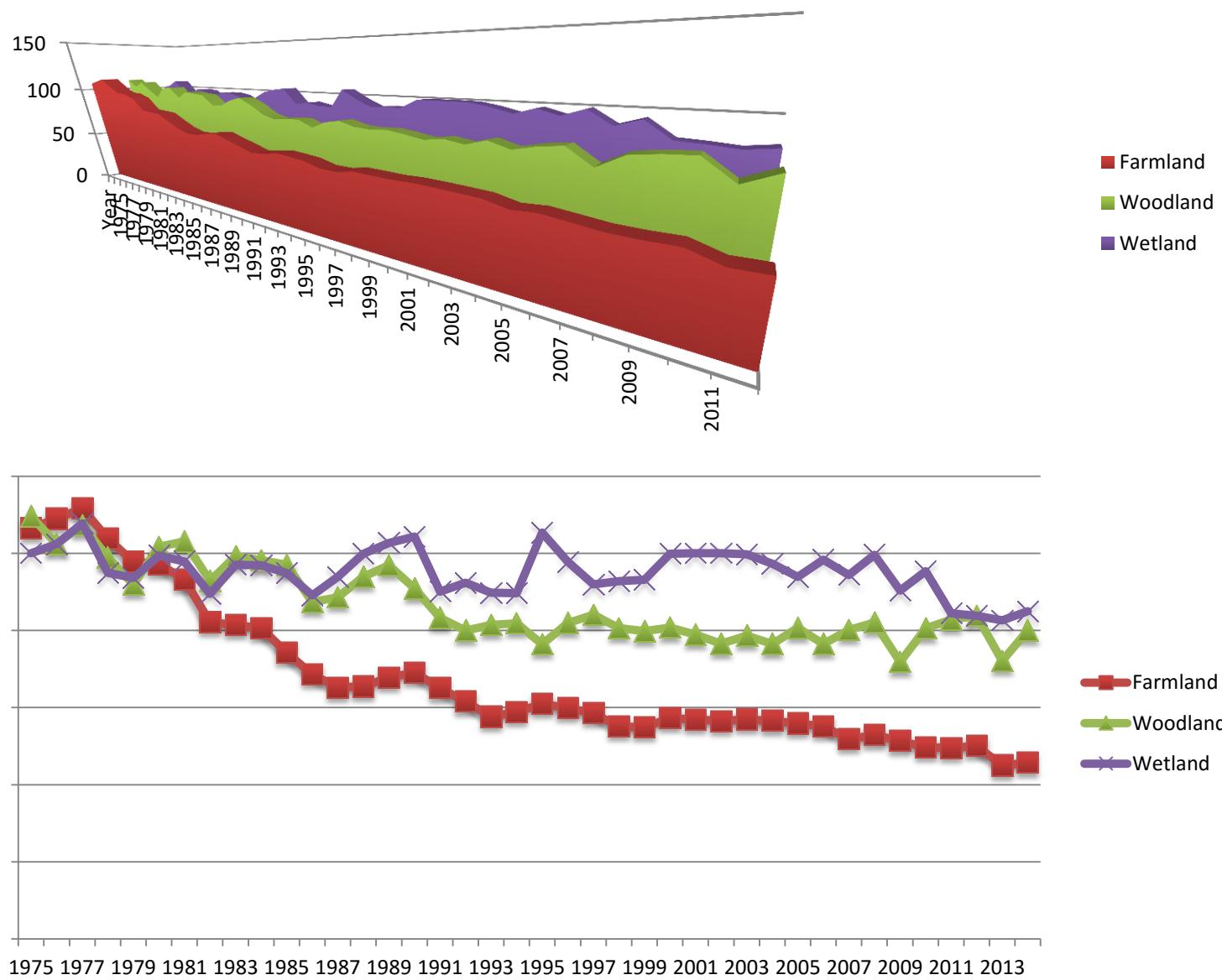


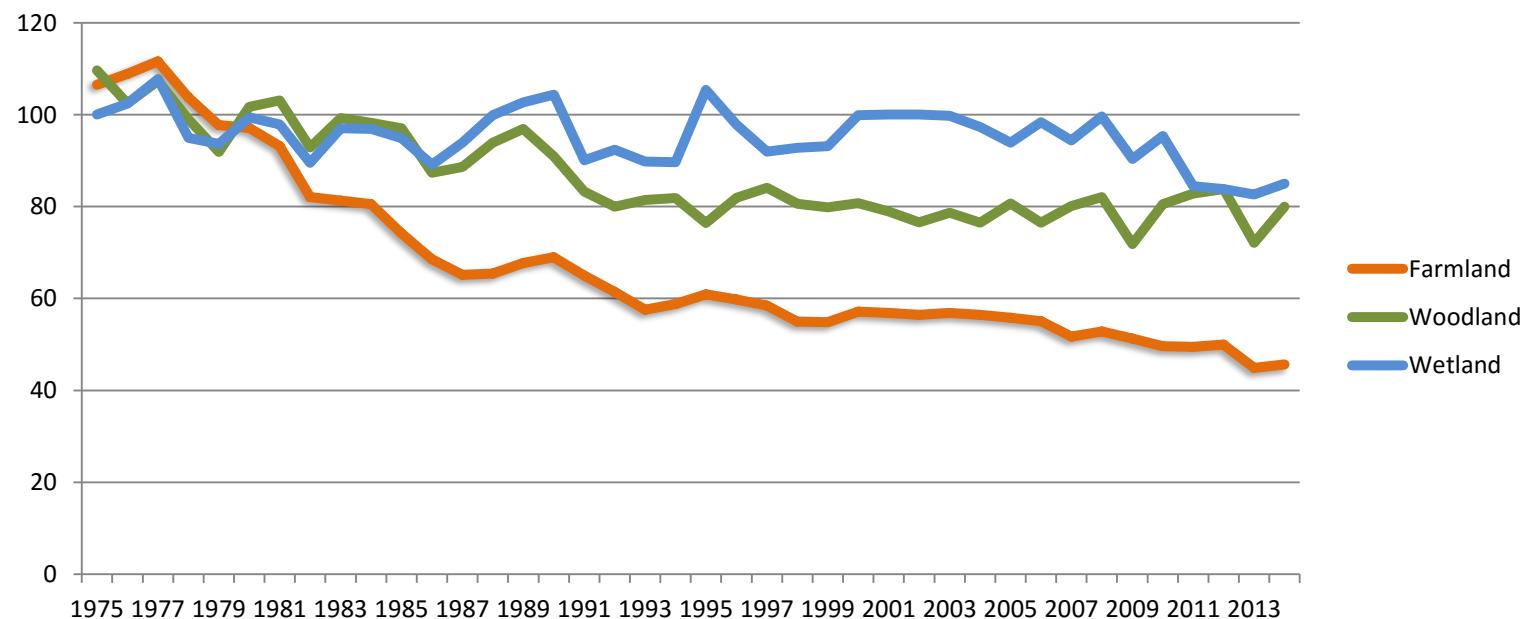
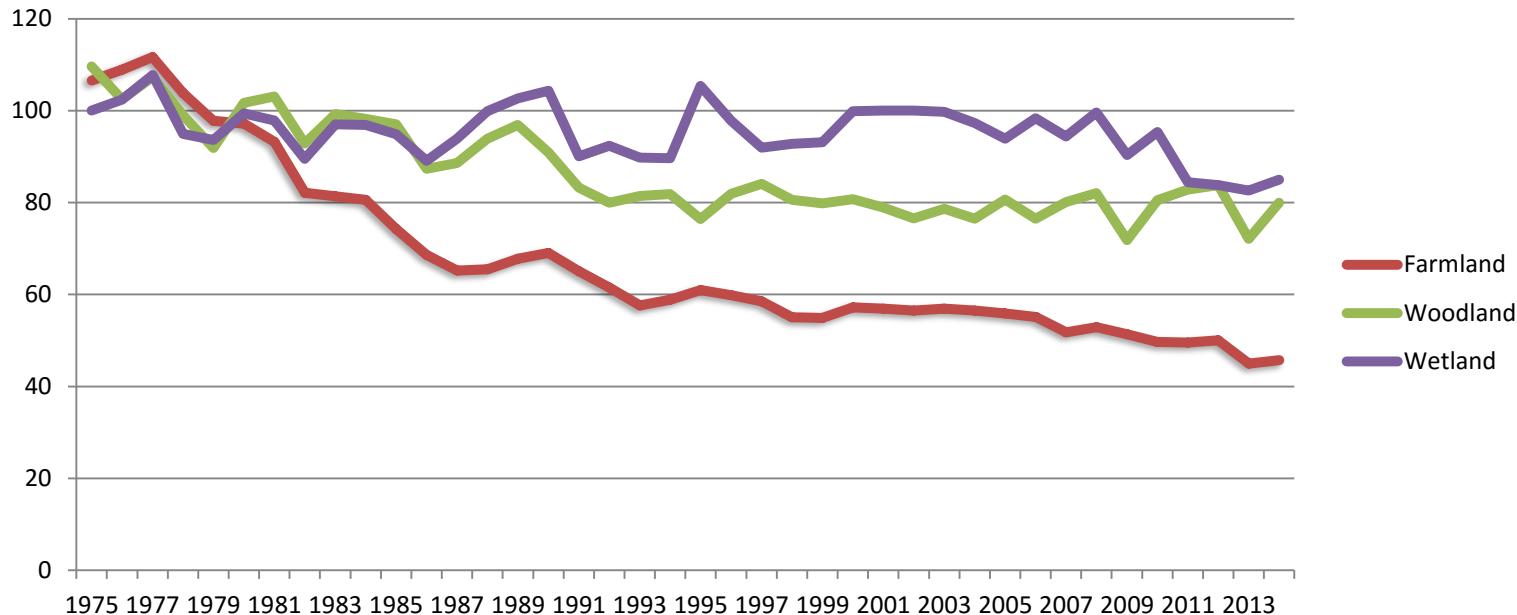
Two single colour gradients -> darker = more extreme
Blue <-> red to encode direction

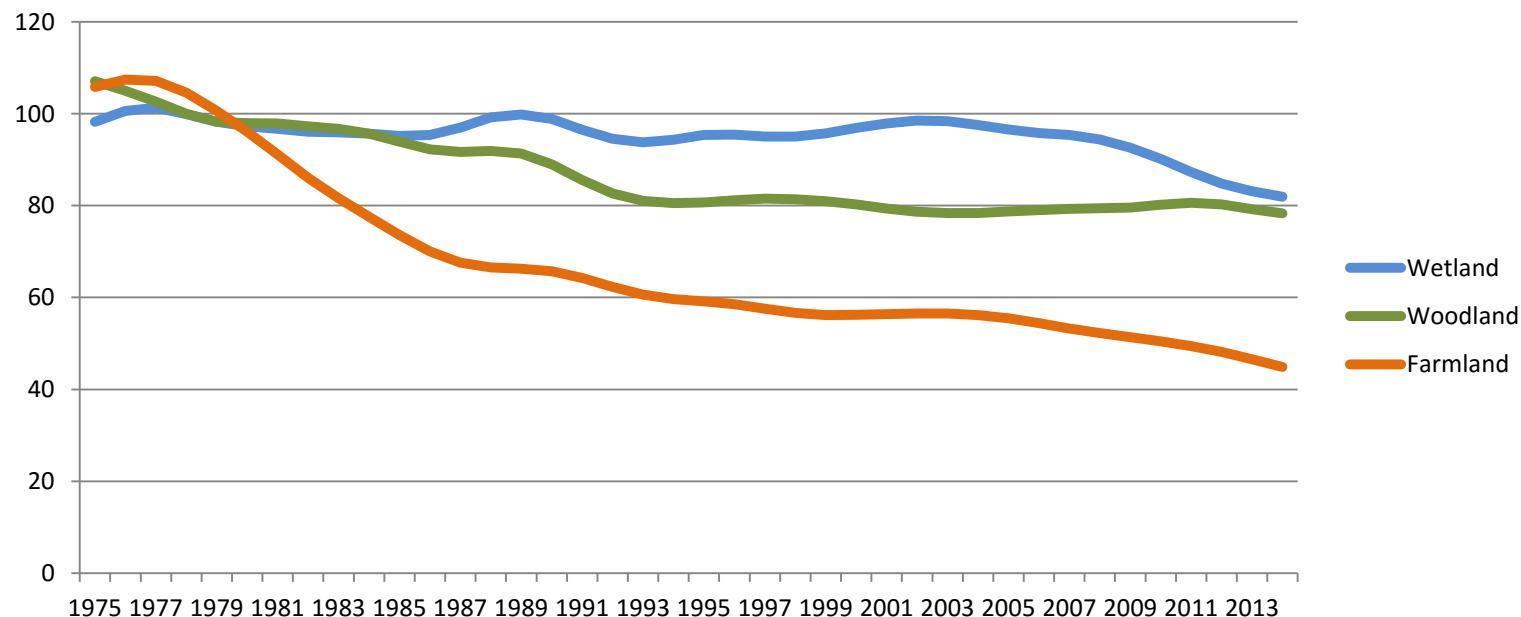
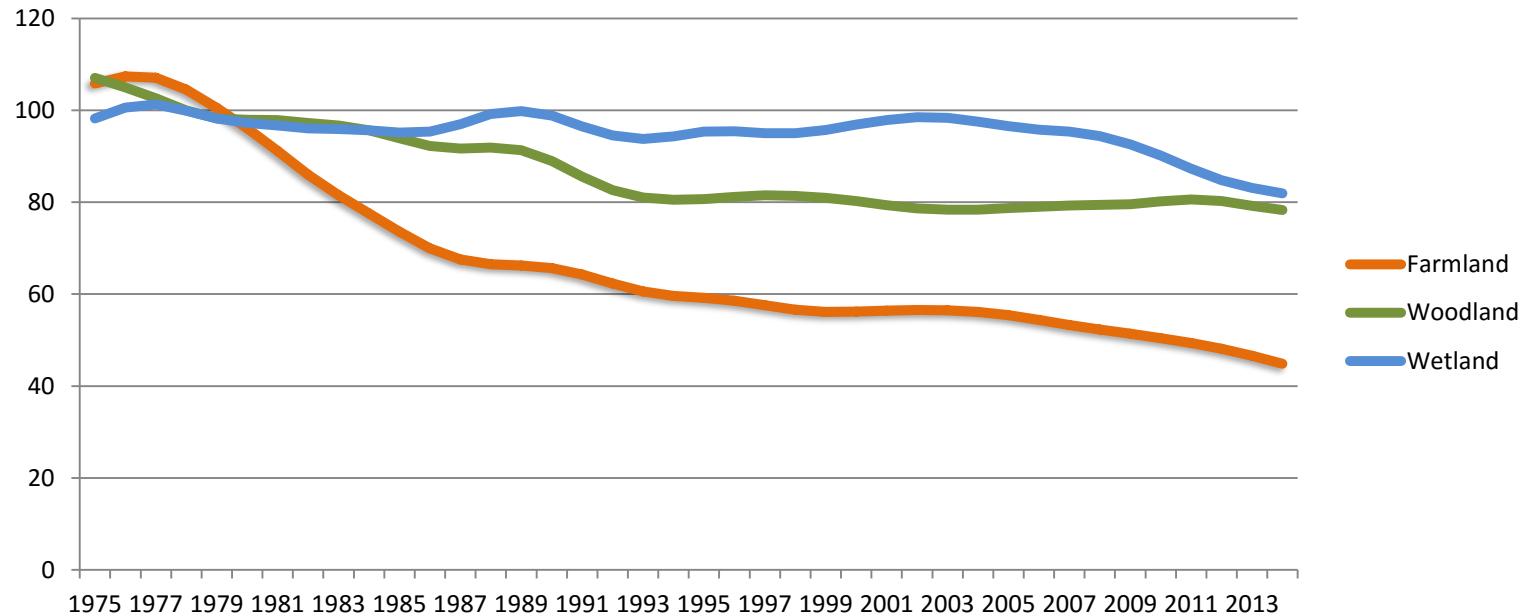
Visualising uncertainty: Confidence bands and “spaghetti plots”

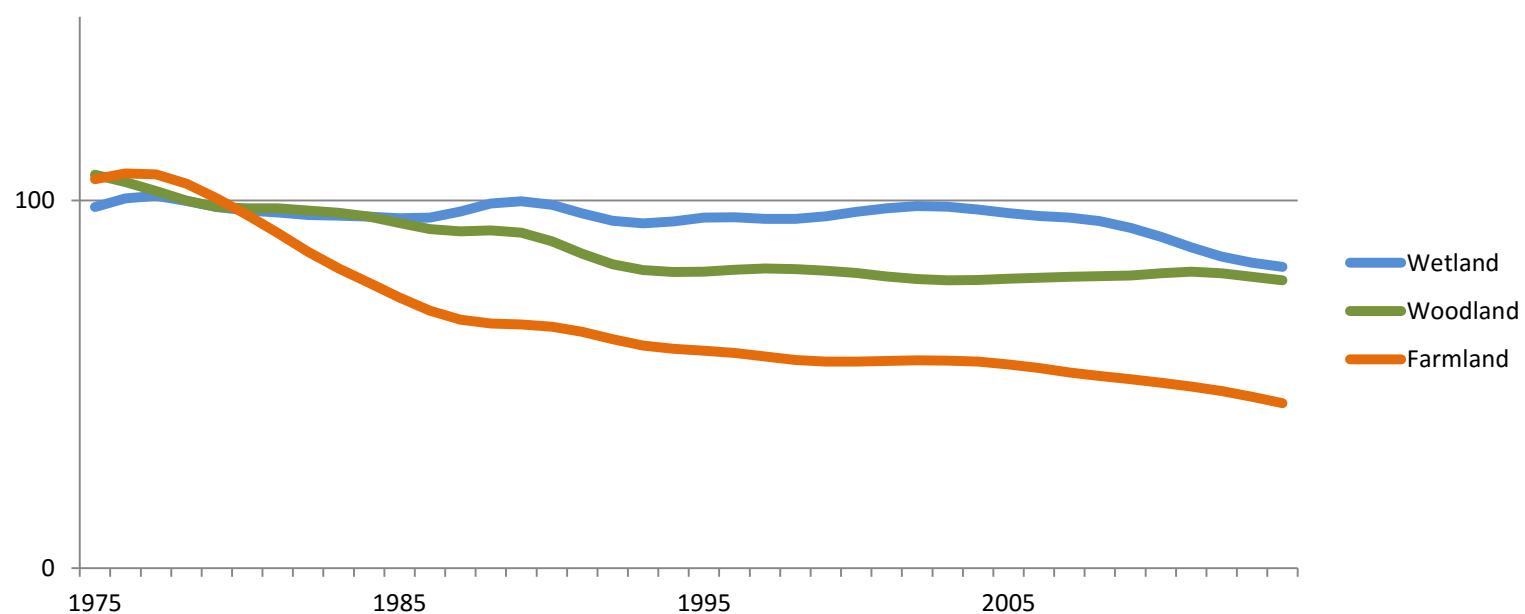
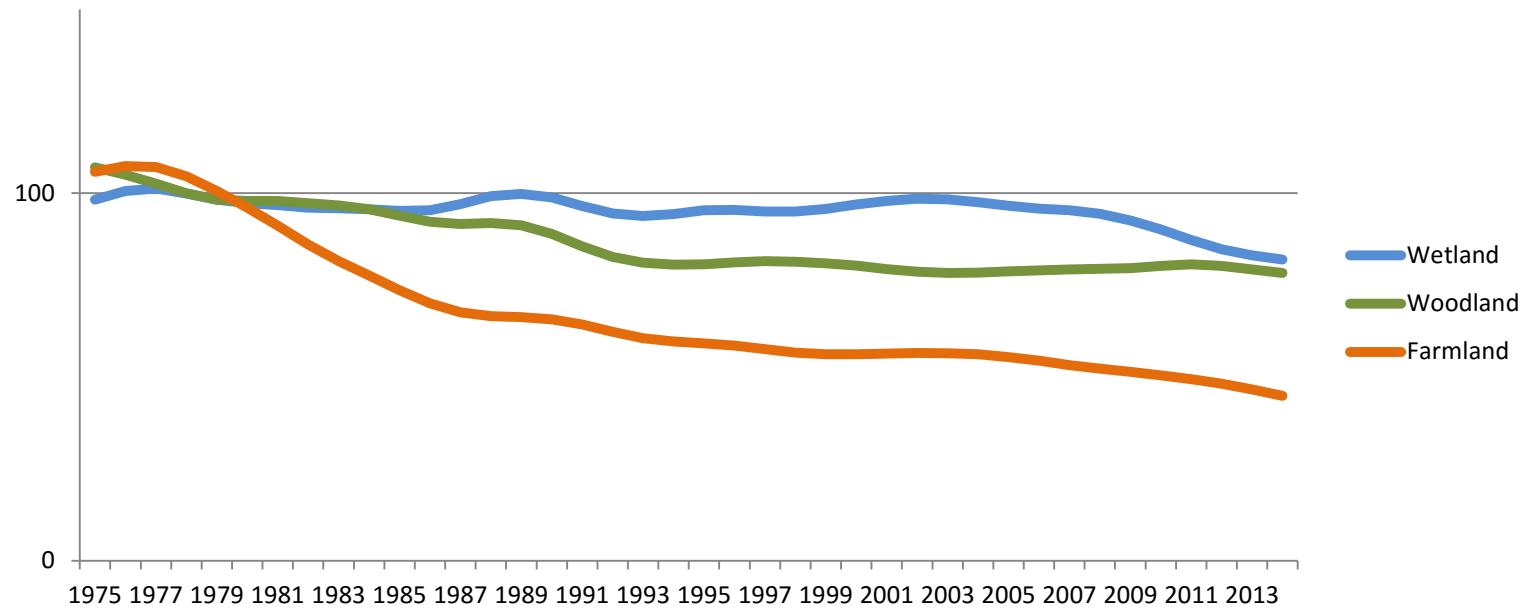


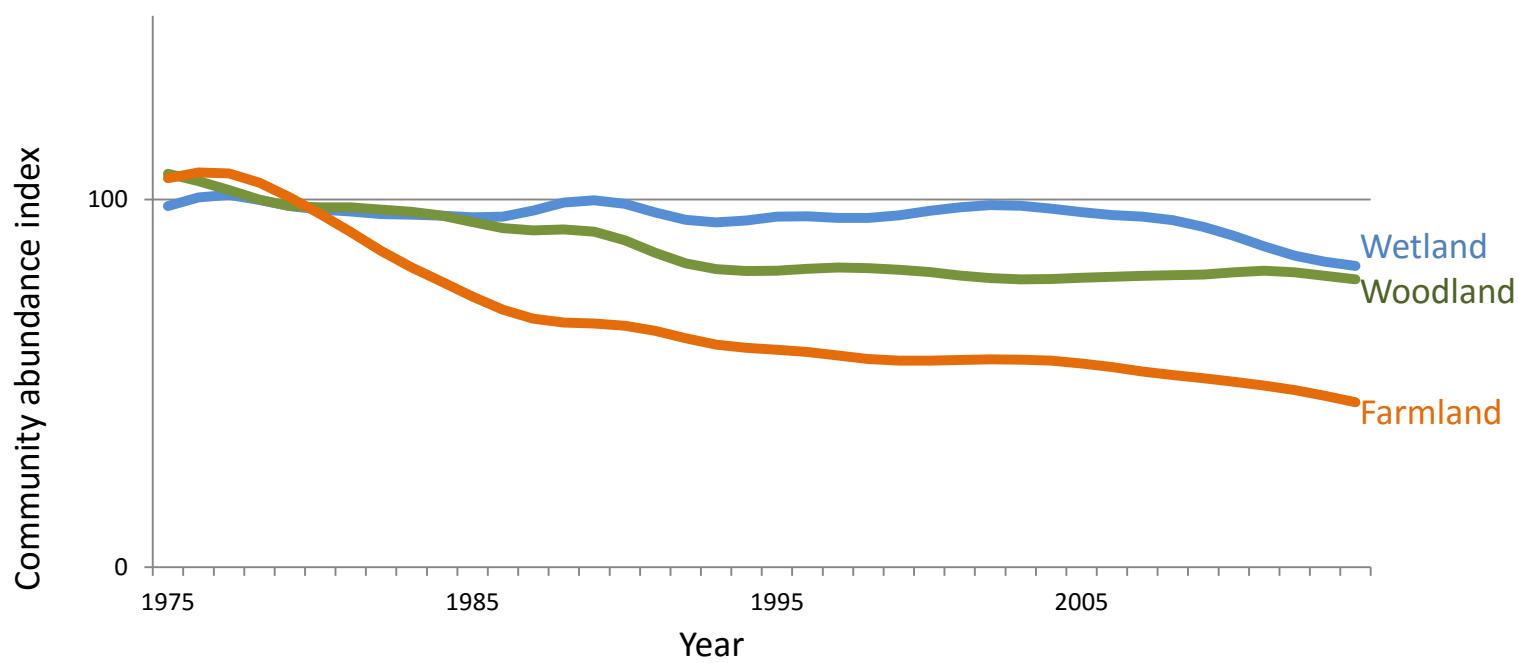
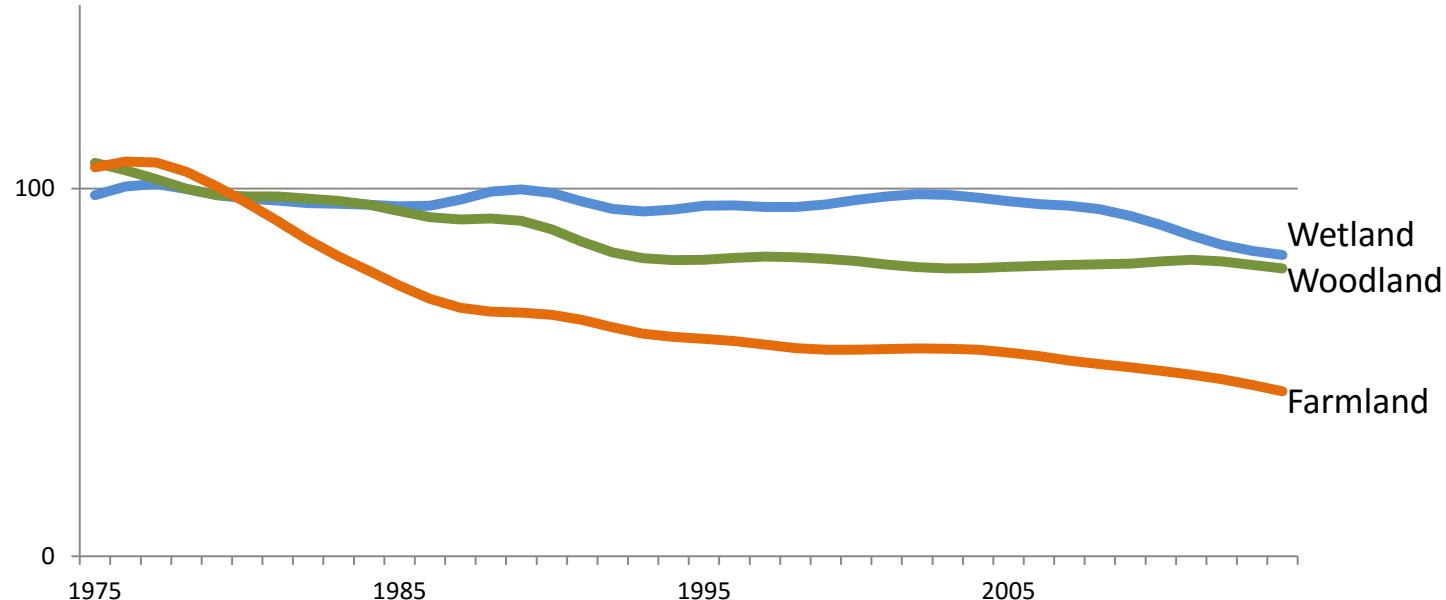
Simplifying graphs to the key message

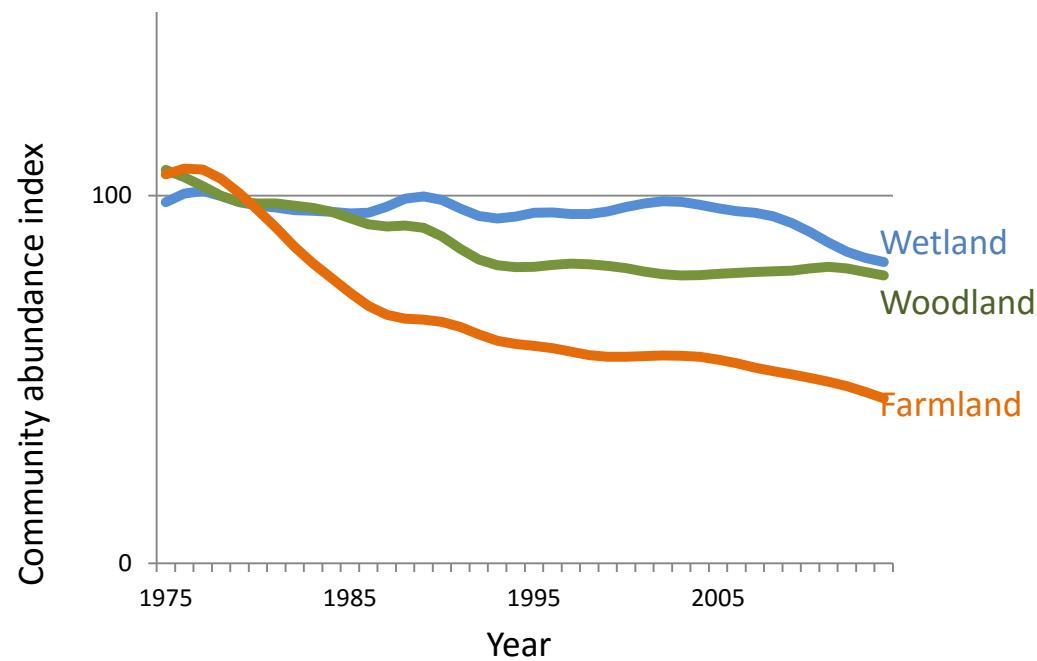
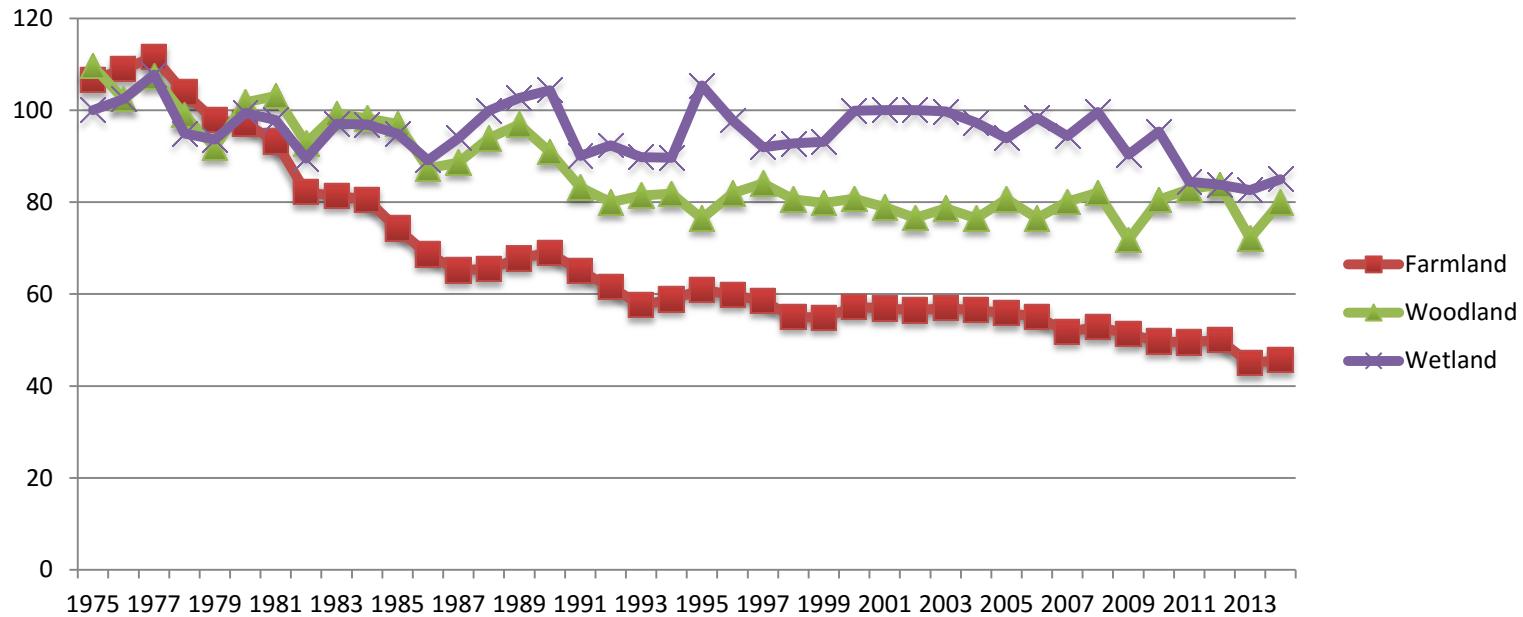




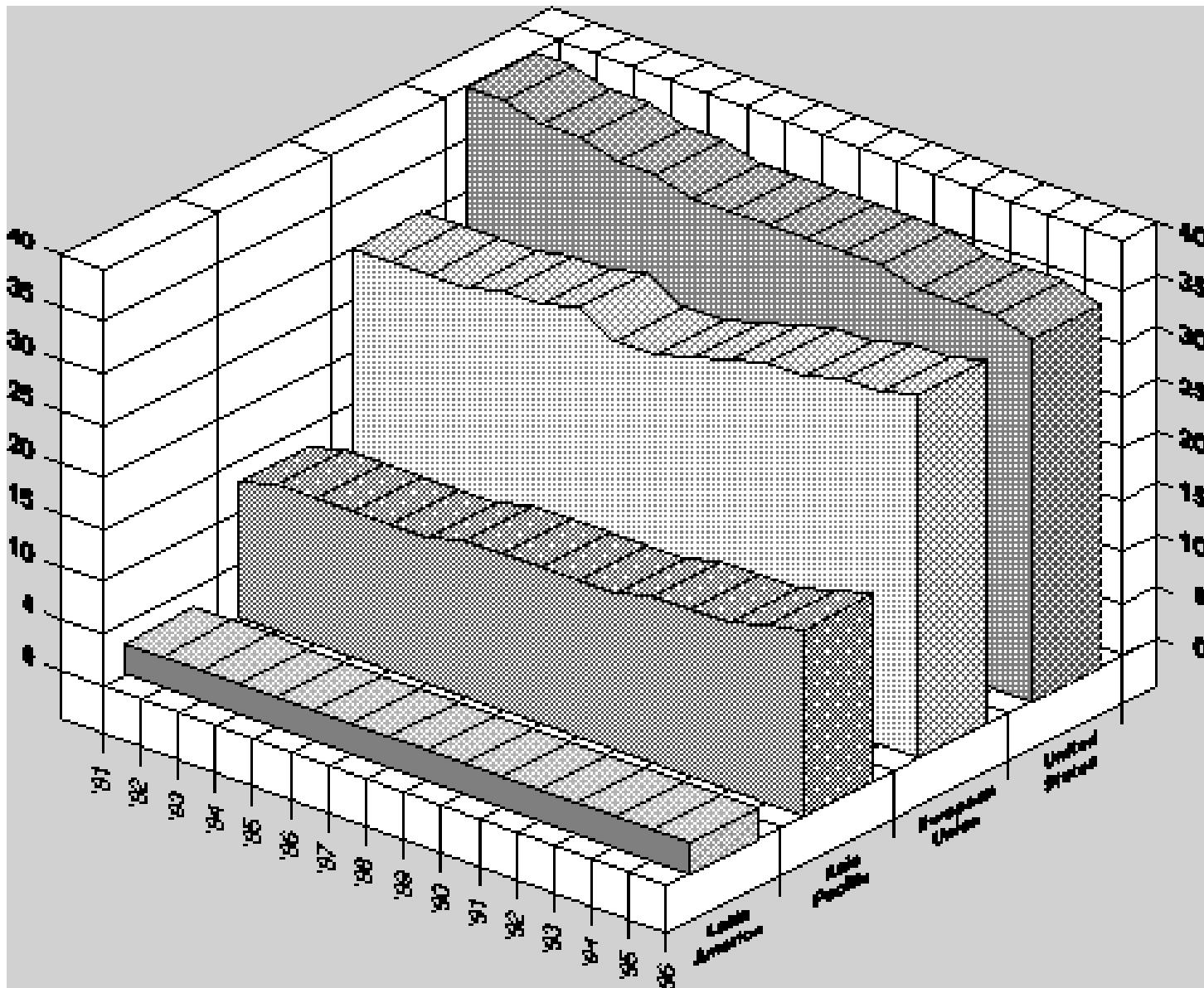






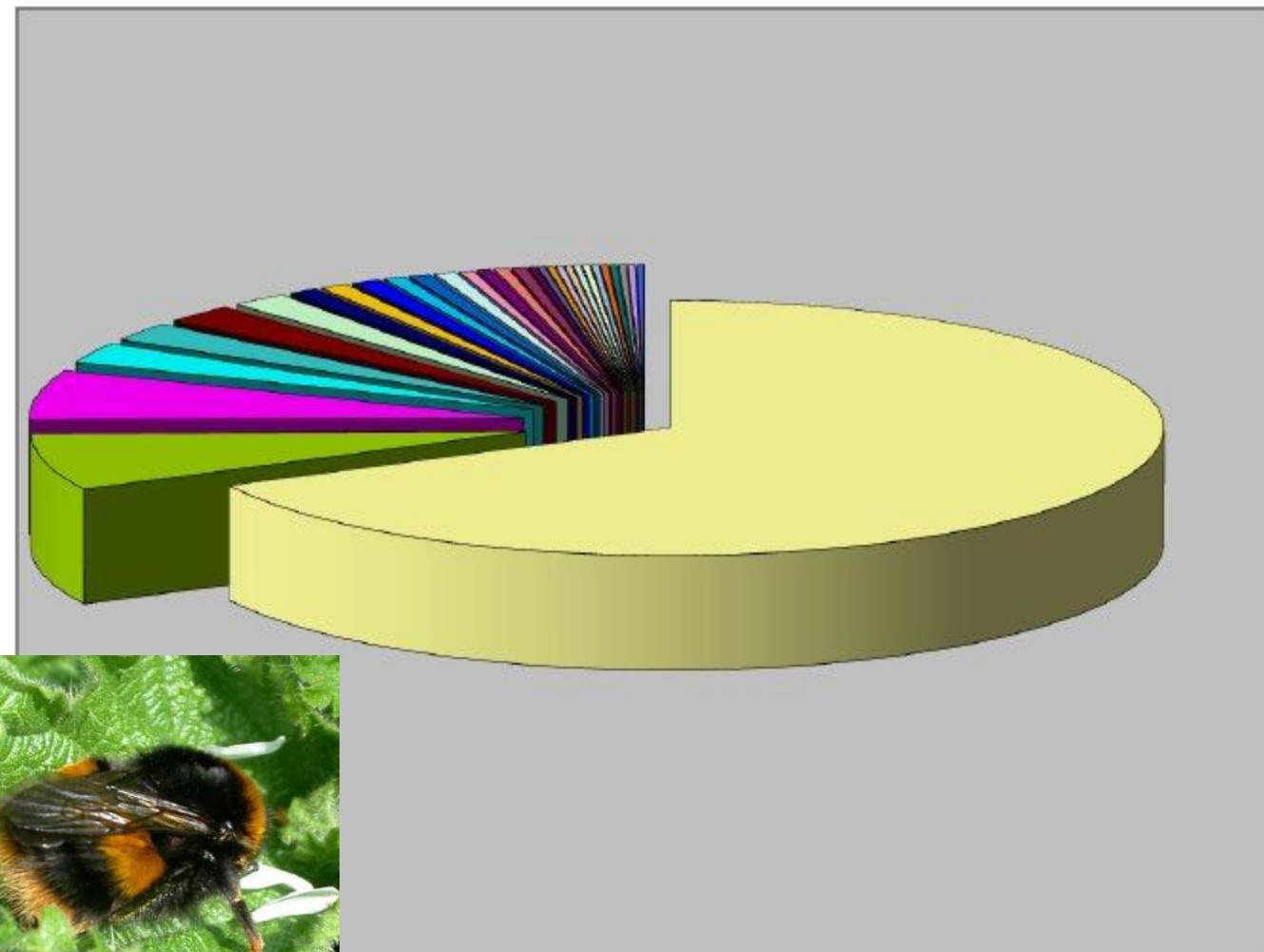


Bad graphs



Bad graphs

Total number of flower visits *Bombus terrestris*
Oct-Dec 2009



Bad graphs

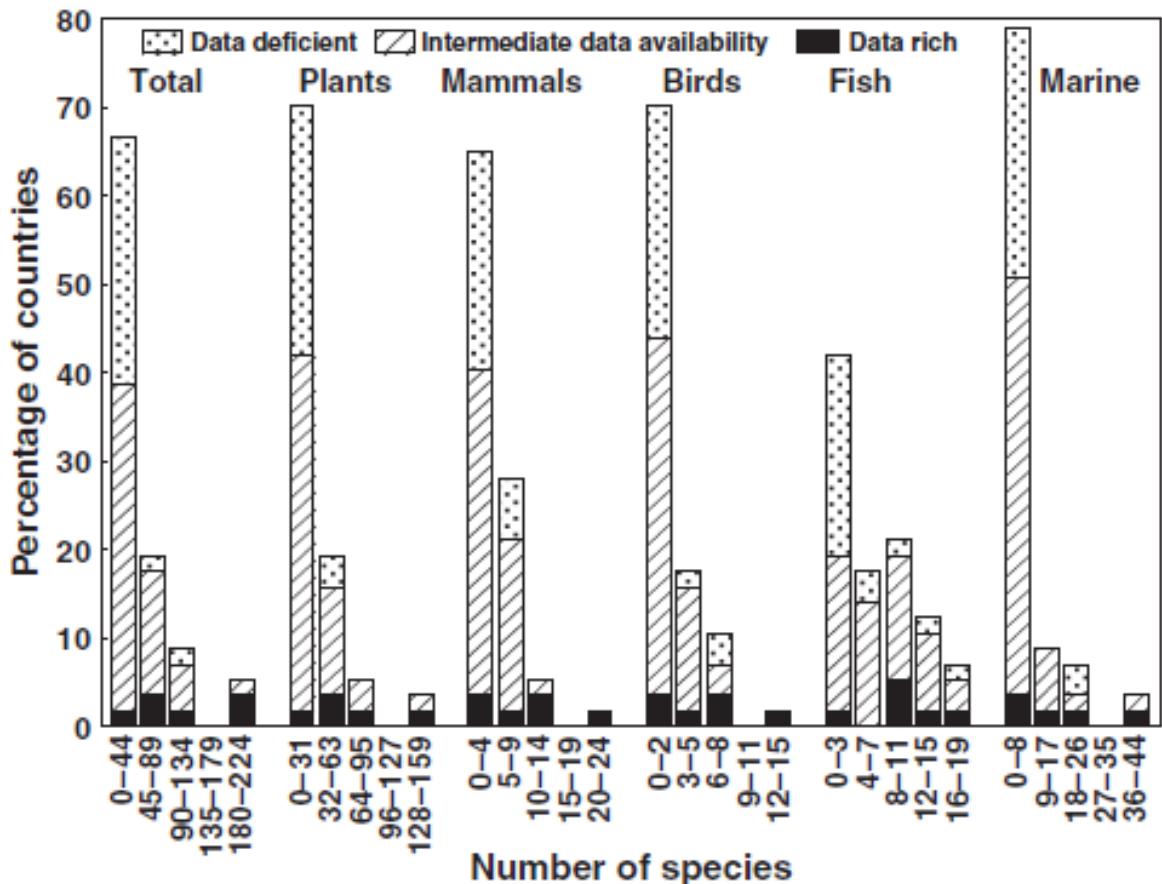
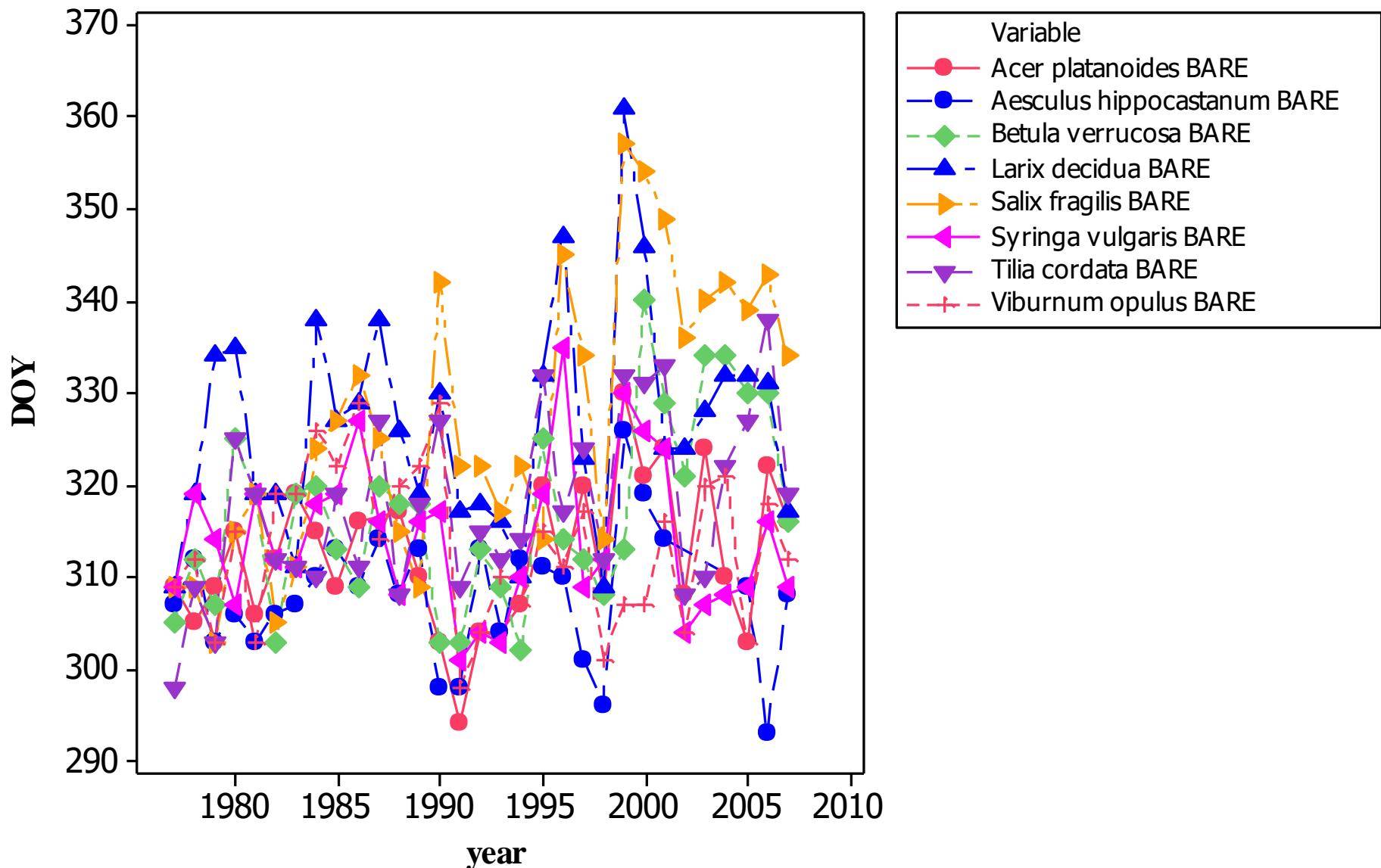


Figure 2 Frequency distributions of the number of Invasive alien species (IAS) across countries ($n = 57$) with different degrees of data adequacy ($n = 18$ data deficient (DD), $n = 33$ intermediate (IDA) and $n = 6$ data rich (DR) countries). (Vascular plants, mammals, birds, freshwater fish and marine organisms; amphibians not shown individually here because of low frequencies.)

Bad graphs



Survey design and statistical analysis for conservation science

SURVEY DESIGN AND SAMPLING

1. REPRESENTATIVE samples
2. STRATIFICATION for sampling design
3. POWER ANALYSIS to assess sampling design
4. PSEUDO-REPLICATION in sampling

STATISTICAL ANALYSIS

5. DESCRIPTION or EXPLANATION
6. SIGNIFICANCE is continuous
7. Separate SIGNIFICANCE and EFFECT SIZE
8. Separate CORRELATION AND CAUSATION
9. Consider UNCERTAINTY

COMMUNICATING RESULTS

10. Translate statistics to HUMAN LANGUAGE
11. Draw GOOD GRAPHS

Where to get support - Statistics

Statistics books are a bit like shoes – different people fit different types...

- **Regression and Other Stories**
 - Free PDF: <https://avehtari.github.io/ROS-Examples/>
 - Comprehensive introduction to modern applied statistics
- **Statistical Rethinking**
 - <https://xcelab.net/rm/statistical-rethinking/>
 - one of the most readable (intermediate - advanced) statistics books
 - free lecture series on youtube
- Wildlife population assessment (bird ringing, camera traps, etc) often require techniques that differ from regression models
 - Books by Kery, Schaub, Royle provide a good starting point
 - E.g. “Bayesian Population Analysis using WinBUGS”
 - <https://www.mbr-pwrc.usgs.gov/pubanalysis/roylebook/>

Where to get support - Coding

- The Carpentries www.carpentries.org
 - Non-profit that runs low cost data analysis / coding workshops worldwide (R, python, UNIX, ...)
 - All materials freely available
 - most are suitable for self study, e.g.
<https://datacarpentry.org/ecology-workshop/>
- Coding Club <https://ourcodingclub.github.io/>
 - Free online data science course for ecologists (R, some python)
 - Includes statistics tutorials
- R-Ladies <https://rladies.org/>
 - Global diversity initiative within in the R community

Where to get support - Collaboration

- You don't have to learn or do everything yourself!
(Even in your Masters/PhD dissertation)
- Conservation Science is highly interdisciplinary
and collaborative, so think about building
collaborations with Statisticians
 - the earlier the better!
- Statistics form a core intellectual contribution to
a study, so give credit where it's due!
 - inclusion in grant proposals, budgets, authorship
 - beer, chocolate, etc

