

The International Biometric Society
Australasian Region Conference

Taupo, New Zealand, 30 Nov – 3 Dec 2009

An Exploratory Analysis of the Effects of Sampling in Marine Surveys for Biodiversity Estimation



Hideyasu SHIMADZU¹

Ross DARNELL²

¹ Geoscience Australia

² CSIRO Mathematical Information Sciences

Outline

1

1. Background
2. Data
3. Impact of sampling process
4. Summary

Background

2

CERF project: (<http://www.marinehub.org/>)

The Commonwealth Environment Research Facilities (CERF) Marine Biodiversity Hub prediction project analyses patterns and dynamic of marine biodiversity to determine the appropriate units and models for effectively predicting Australia's marine biodiversity.

The project administered through the Australian Government Department of the Environment, Water, Heritage and the Arts.

Major contributors: University of Tasmania (UTas); CSIRO Wealth from Oceans Flagship; Geoscience Australia (GA); Australian Institute of Marine Science (AIMS); Museum Victoria (MV).

Aims: Modelling the relationship between biodiversity (eg presence/absence, richness, biomass etc) and the environment (eg depth, %carbonate, %mud, O₂).

Niche theory

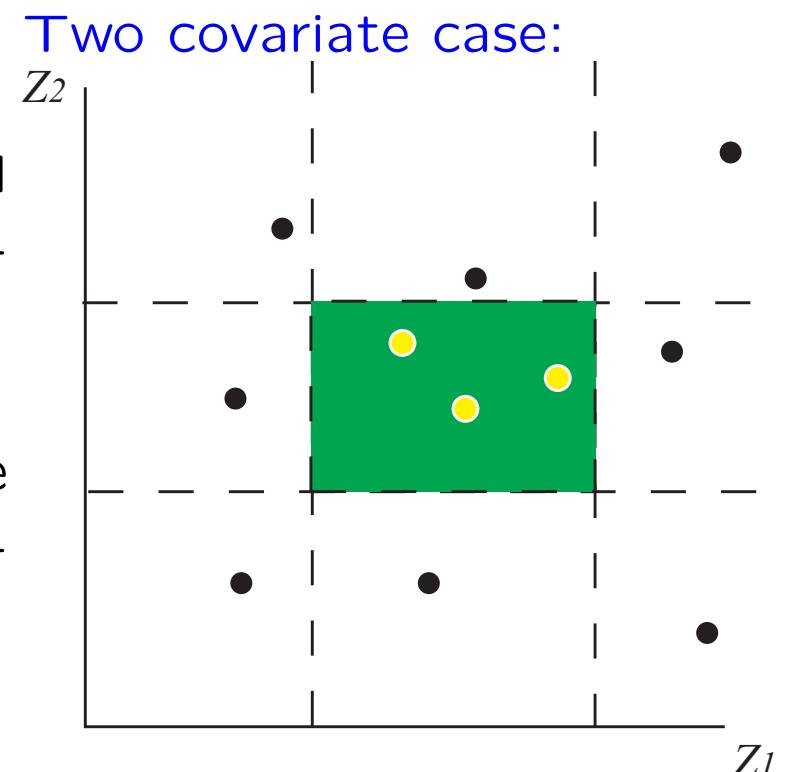
3

Hutchinson (1957)

Niche: The full range of environmental conditions under which an organism can exist.

Note:

Niche is a sub-space of a hyper space spanned by some physical covariates (environmental conditions).



Great Barrier Reef Data

4

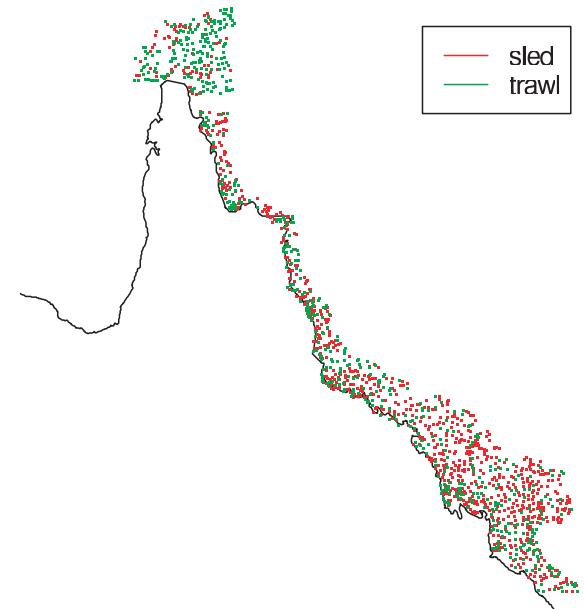
Pitcher *et al.* (2007)

Survey period: 2003 – 6

Number of observation sites: 1252

Methods: Trawl, Sled

Number of species observed: 2862



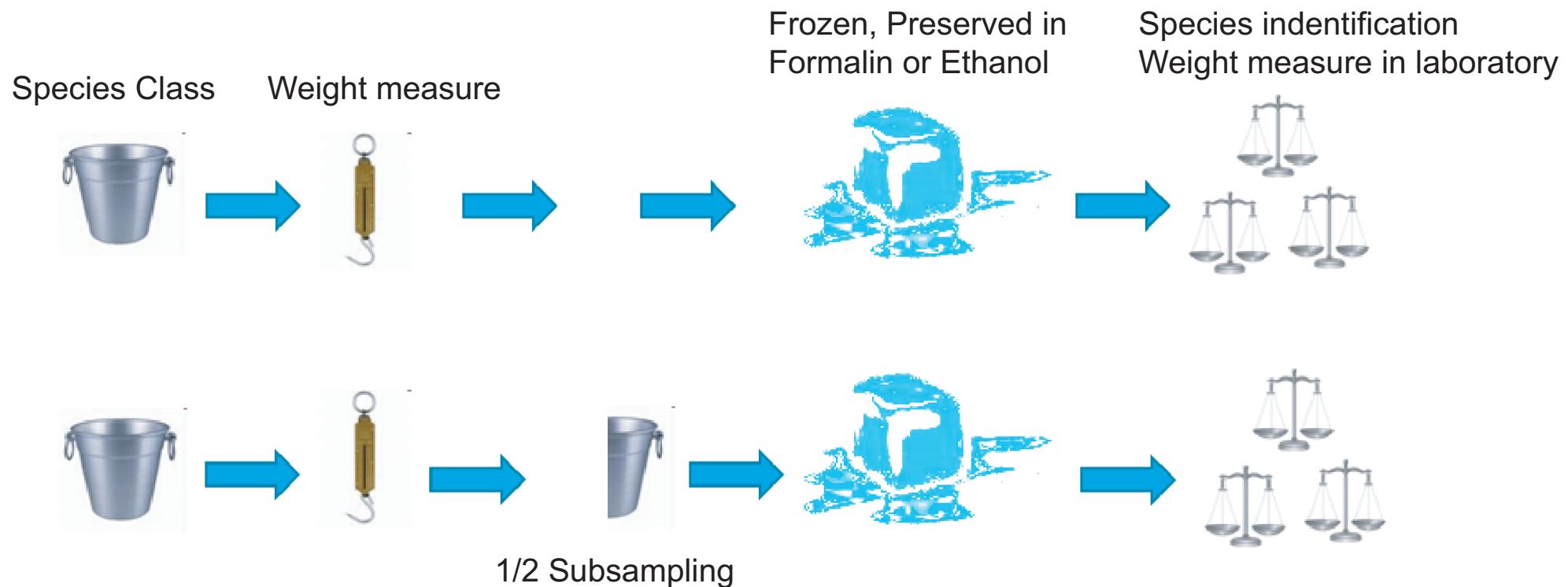
Note

Also physical covariates are given.



Sampling process

5



Species classes

Algae, Brachiopoda, Bryozoa, Crustacea, Porifera, Seagrass, etc.

Subsampling

6

Widely used method in marine surveys.
A part of the whole catch is sorted.
Many variety of subsampling procedures
(survey specific).

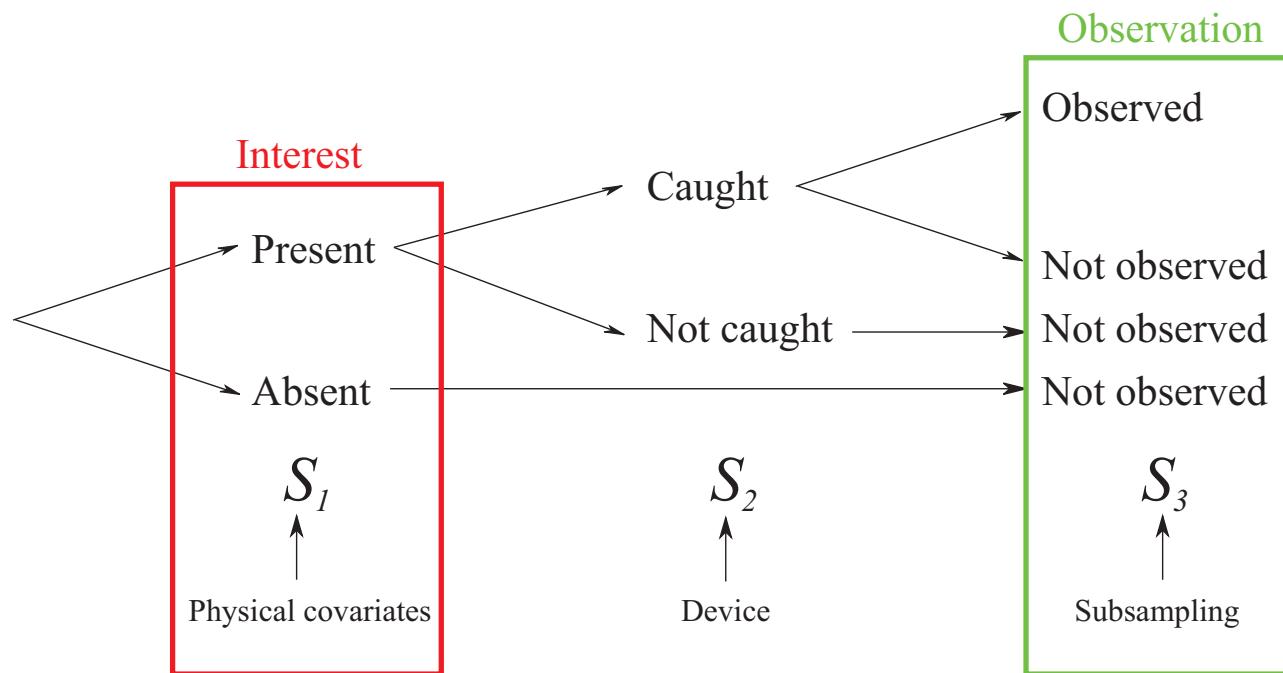
Note: Very little research on subsampling
(eg Heales *et al.*, 2003).



Model

7

Species presence/absence



Occupancy – Detectability – Sampling (Compound events)

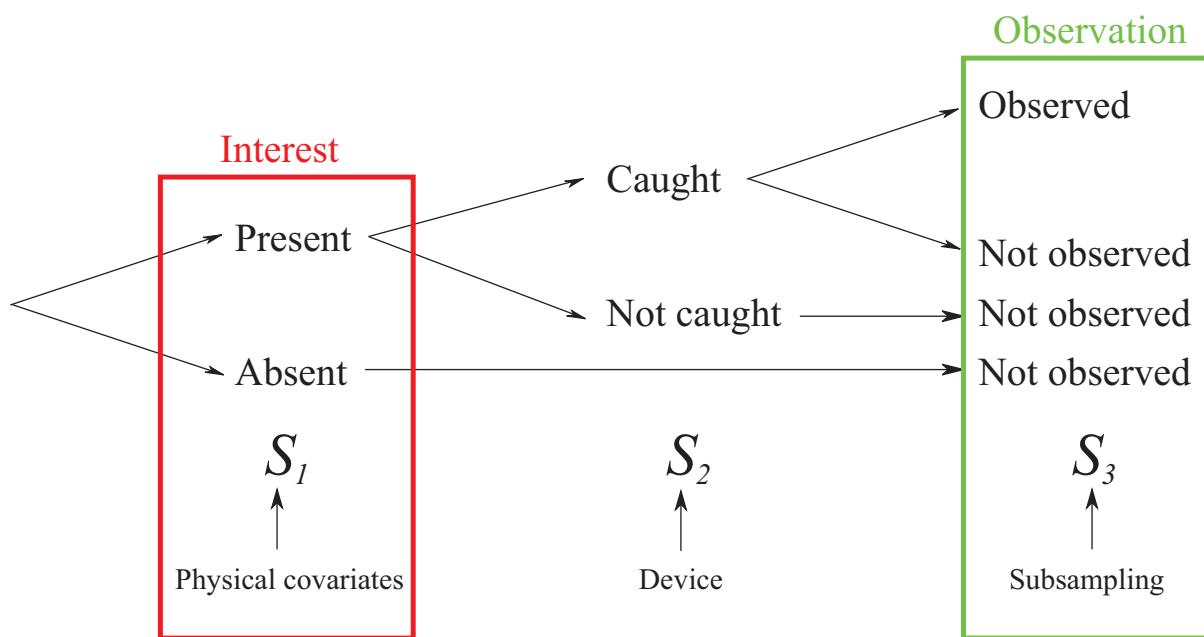
$$\Pr(S_3 = s_3) = \sum_{S_2} \sum_{S_1} \Pr(S_1, S_2, S_3) = \sum_{S_2} \sum_{S_1} \Pr(S_1) \Pr(S_2|S_1) \Pr(S_3|S_2, S_1)$$

Our approach

8

Restrict our attention on the effect of subsampling:

- Fully sampled vs. subsampled;
- Sled;
- Site matching



Site matching

9

For survey sites $\{j : j = 1, 2, \dots, n\}$, matched pair sites such that one is fully sampled and the other is subsampled are

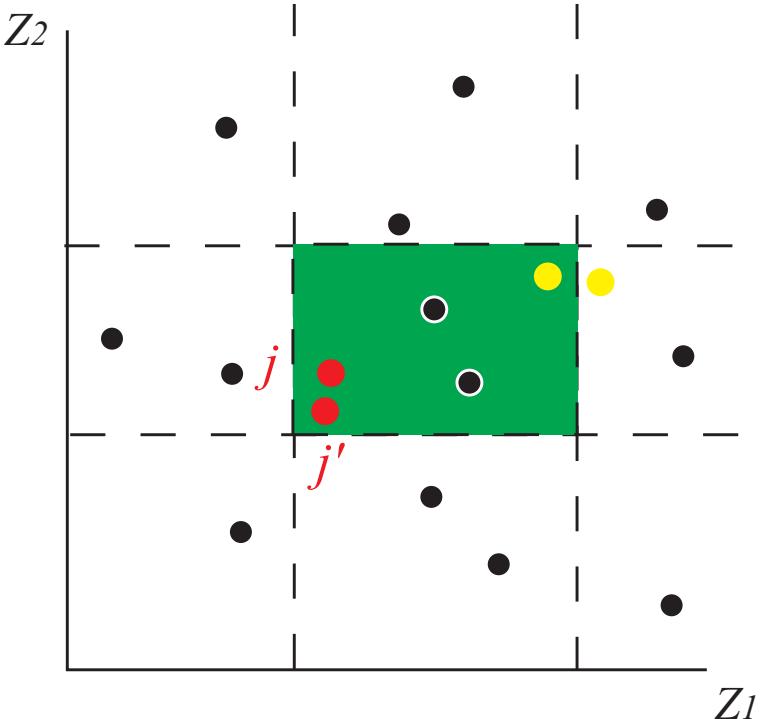
$$\left\{ (j, j') : \min_{j' \in \mathcal{J}'} \|z_j - z_{j'}\|, j \in \mathcal{J} \right\},$$

where z_j is a vector of physical covariates at site j .

Given a pair of species presence/absence $(Y_j, Y_{j'})$, odds ratio is given as

$$\text{OR} = \frac{\sum_j y_j}{|\mathcal{J}|} / \frac{\sum_{j'} y_{j'}}{|\mathcal{J}'|},$$

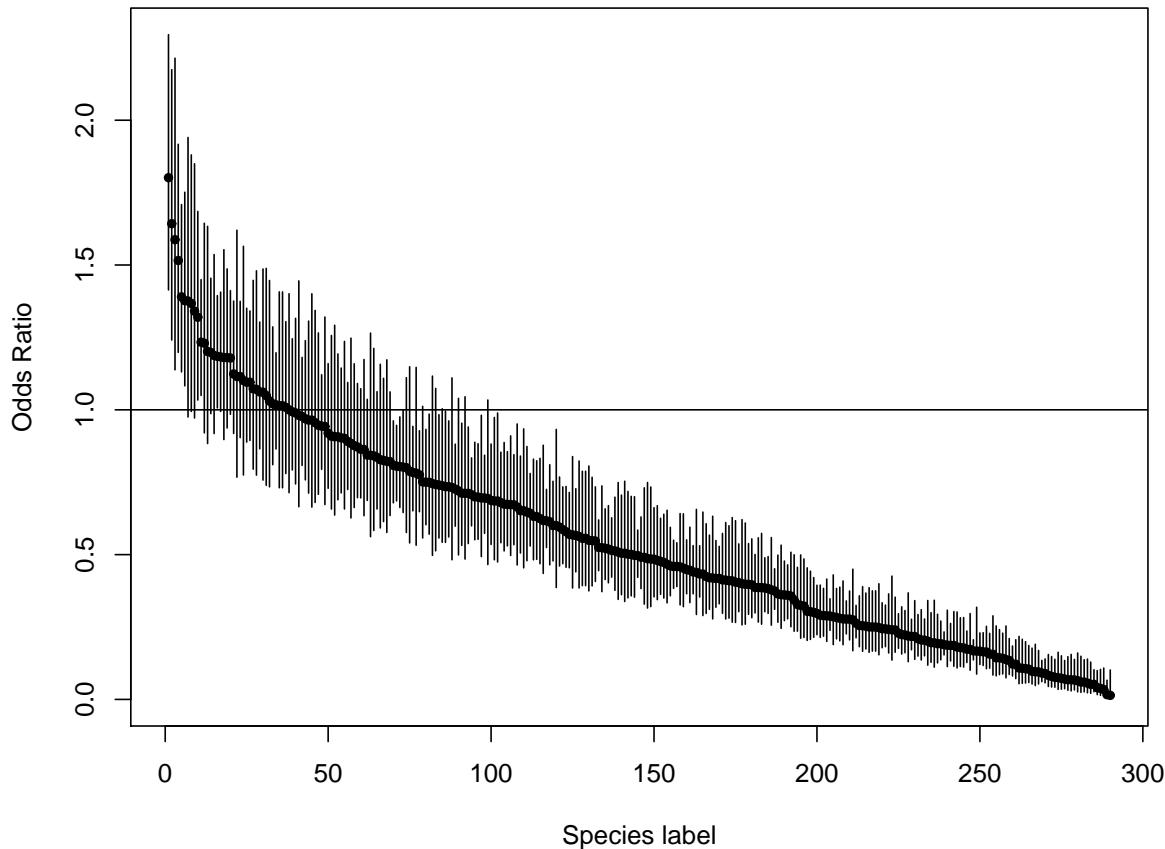
where $|\cdot|$ denotes the number of elements of a set.



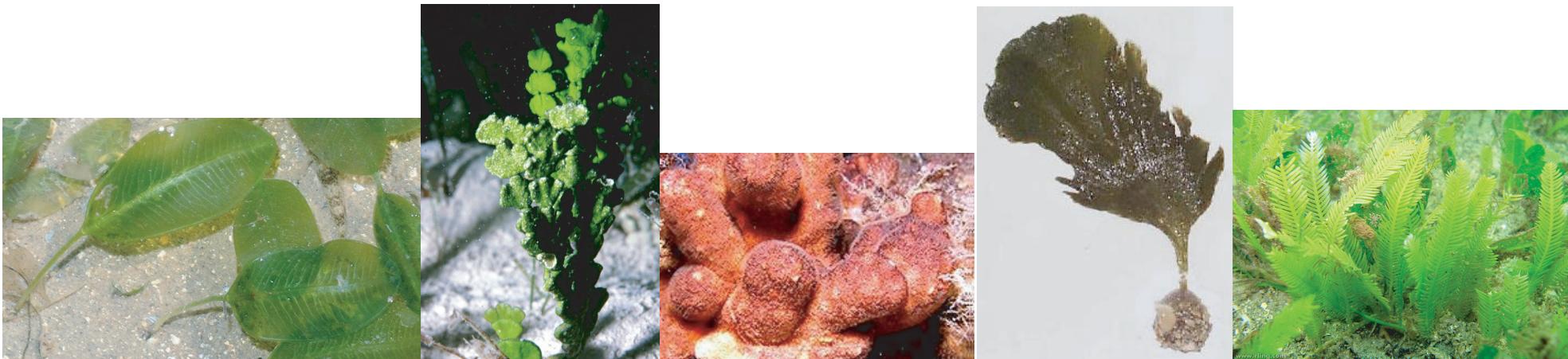
Subsampled vs. fully sampled

10

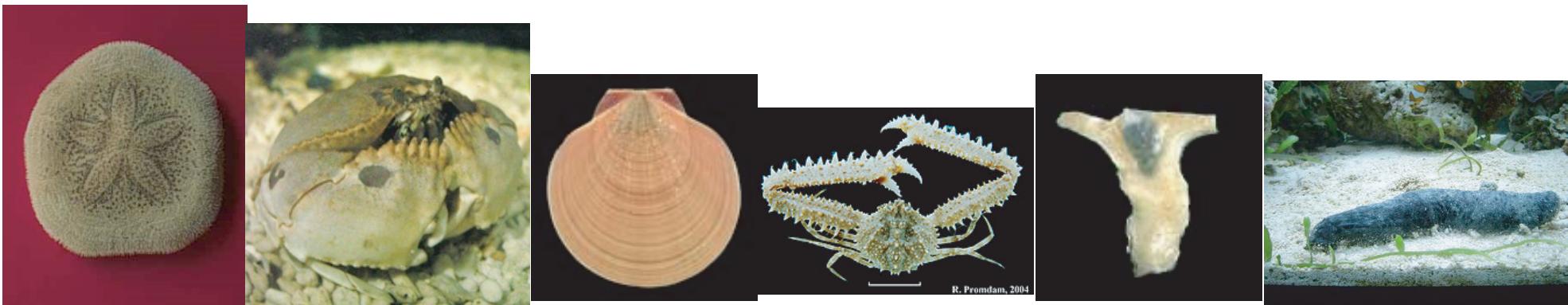
When subsample was taken species are more likely to be observed...

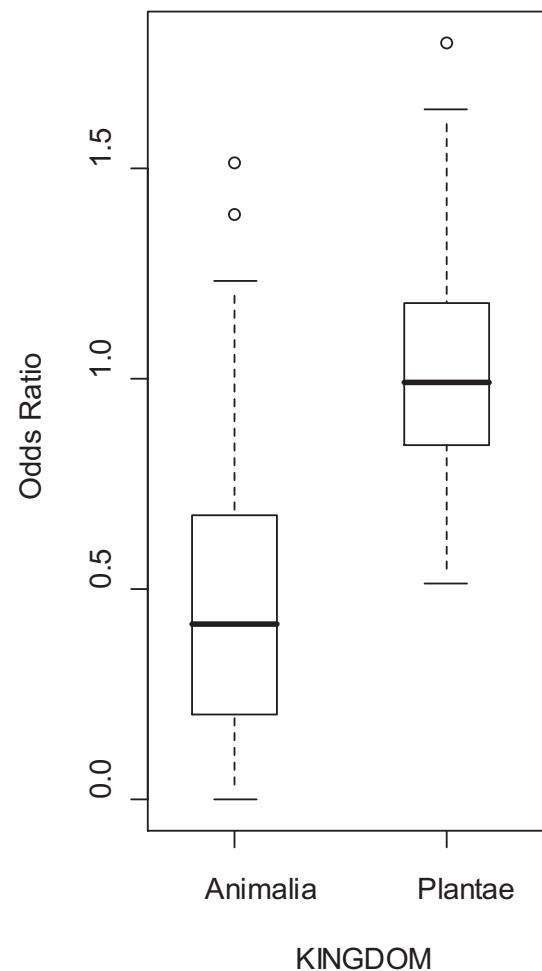


Odds ratio	Scientific name
1.801	<i>Magnoliophyta Liliopsida Hydrocharitales Hydrocharitaceae Halophila ovalis</i>
1.643	<i>Chlorophyta Chlorophyceae Caulerpales Codiaceae Halimeda gigas</i>
1.587	<i>Chlorophyta Chlorophyceae Caulerpales Codiaceae Halimeda borneenses</i>
1.516	<i>Bryozoa Gymnolaemata Cheilostomata Margarettidae Margaretta spp</i>
1.390	<i>Bryozoa Gymnolaemata Cheilostomata Schizoporellidae Stylopoma spp</i>
1.377	<i>Chlorophyta Chlorophyceae Caulerpales Codiaceae Udotea flabellum</i>
1.376	<i>Chlorophyta Chlorophyceae Caulerpales Caulerpaceae Caulerpa taxifolia</i>
1.367	<i>Chlorophyta Chlorophyceae Cladophorales Boodleaceae Phyllocladion sp1</i>
1.341	<i>Chlorophyta Chlorophyceae Caulerpales Caulerpaceae Caulerpa brachypus</i>
1.320	<i>Phaeophyta Phaeophyceae Dictyotales Dictyotaceae Lobophora variegata</i>

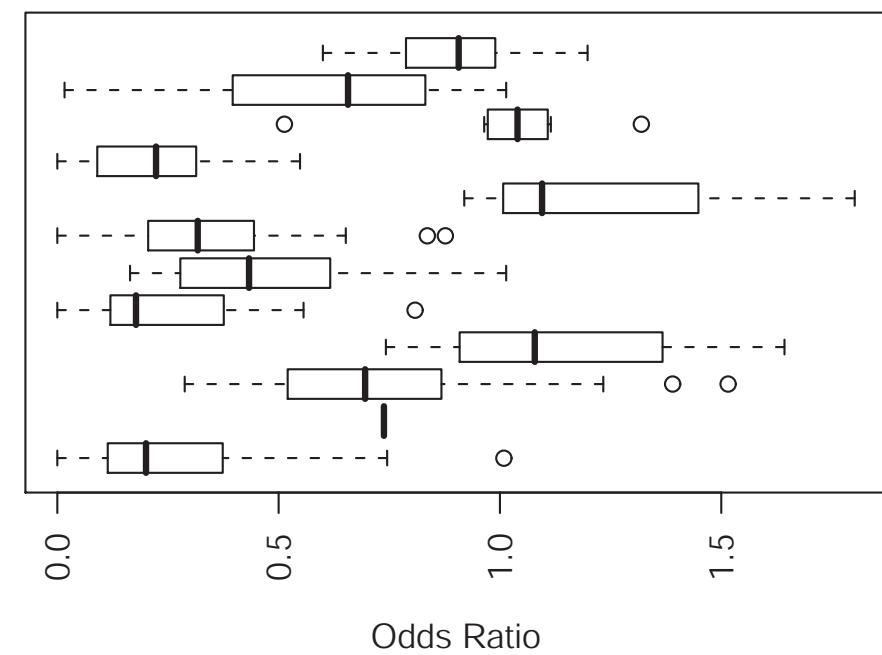


Odds ratio	Scientific names
0.014	<i>Echinodermata Echinoidea Clypeasteroida Laganidae Laganidae</i>
0.016	<i>Porifera Demospongiae Hadromerida Tethyidae Xenospongia patelliformis</i>
0.034	<i>Arthropoda Crustacea Decapoda Calappidae Calappa</i>
0.038	<i>Mollusca Bivalvia Ostreoida Pectinidae Amusium pleuronectes</i>
0.041	<i>Arthropoda Crustacea Decapoda Xanthidae Actumnus squamosus</i>
0.052	<i>Arthropoda Crustacea Decapoda Parthenopidae Parthenope longimanus</i>
0.052	<i>Mollusca Bivalvia Pterioida Malleidae Malleus albus</i>
0.056	<i>Echinodermata Holothuroidea Holothuroidea Holothuroidea Holothuroidea</i>
0.060	<i>Mollusca Gastropoda Sorbeoconcha Bursidae Bufonaria rana</i>
0.060	<i>Chordata Actinopterygii Tetraodontiformes Monacanthidae Paramonacanth</i>





Rhodophyta
Porifera
Phaeophyta
Mollusca
Magnoliophyta
Echinodermata
Cnidaria
Chordata
Chlorophyta
Bryozoa
Brachiopoda
Arthropoda



Summary

14

- Challenges in marine survey data;
- The impact of subsampling;
- Need more investigation of the factor related to odds ratio.

Thank you for your kind attentions.
Comments and suggestions are welcome!

Ross DARNELL
Ross.Darnell@csiro.au

Hideyasu SHIMADZU
Hideyasu.Shimadzu@gc.gov.au