



# Habitat prediction:

Can we produce probable, practical  
*and* ecologically meaningful results?



Katrina Baxter  
University of Western Australia

Gary Kendrick, Andy Bickers, Kimberly Van Niel

# Overview



- Expected outcomes of habitat prediction
- Some Key Considerations
- Recherche Archipelago
- Prediction methods applied using available VIDEO data.
- Implications for applying these methods to full coverage spatial data

# Expected outcomes



Why predict?

- Describe where habitat types are found  
*“Classify the **combination of identifiable and distinct** characteristics that distinguish the environment and the ecological communities that live there”*
- Apply patterns to other data to predict into unsurveyed areas > Target survey efforts, minimise costs
- **Produce maps of probable habitat distribution**  
**i.e. where are we MOST LIKELY to find dense seagrass?**

# Key Considerations



## **1. What information and level of detail do we need to best explain habitat distribution?**

- Is how we classify / map habitats ecologically meaningful from a prediction perspective?

## **2. How can we best guide our sampling to fulfil both mapping and prediction goals?**

- WHAT DATA DO WE HAVE? WHAT DATA DO WE NEED ?  
CAN WE COLLECT IT? IS IT COST EFFECTIVE TO DO SO?

## **3. Is it really possible then to *spatially predict* the probable extents or occurrence of habitats?**

# The Recherche Archipelago



## Isolated and unstudied

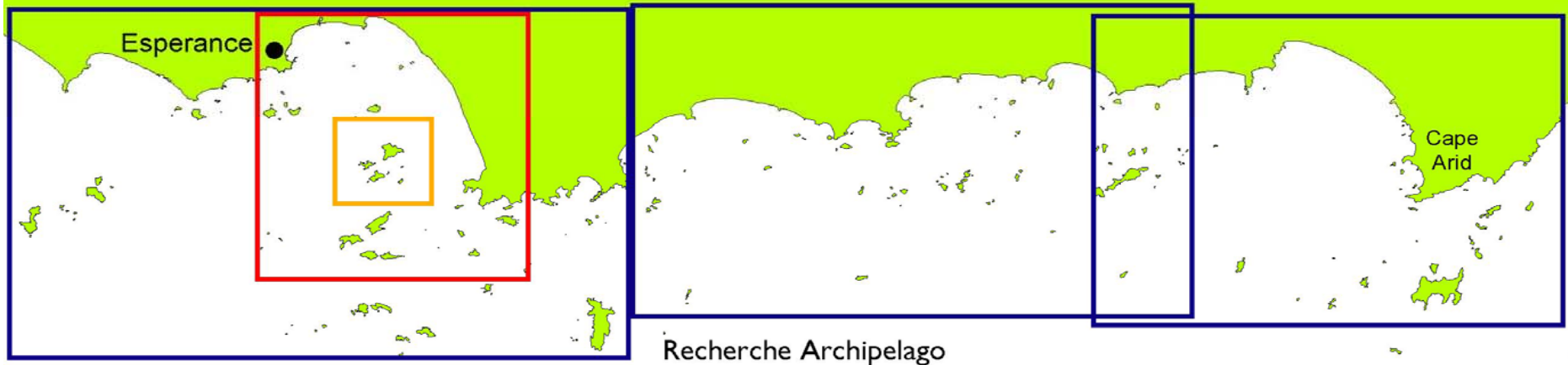
- Sparse bathymetry
- Little existing habitat data
- Isolated biological data

## FRDC / CRC project - 2 Main goals

- Map habitats for community
- Predict / explain habitat occurrence



# What we need to map with

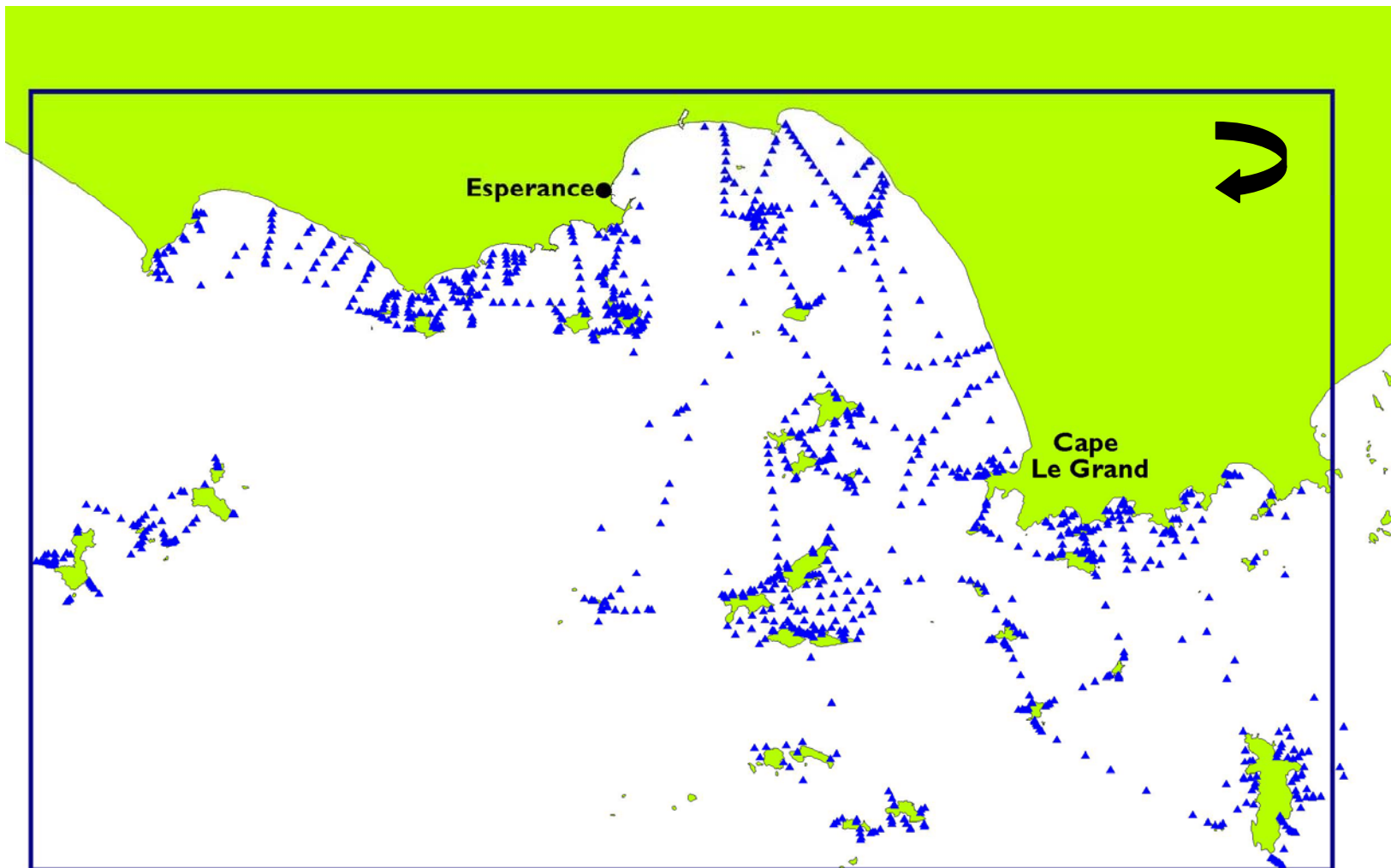


- **“Drop camera” survey (3000 camera drops)**

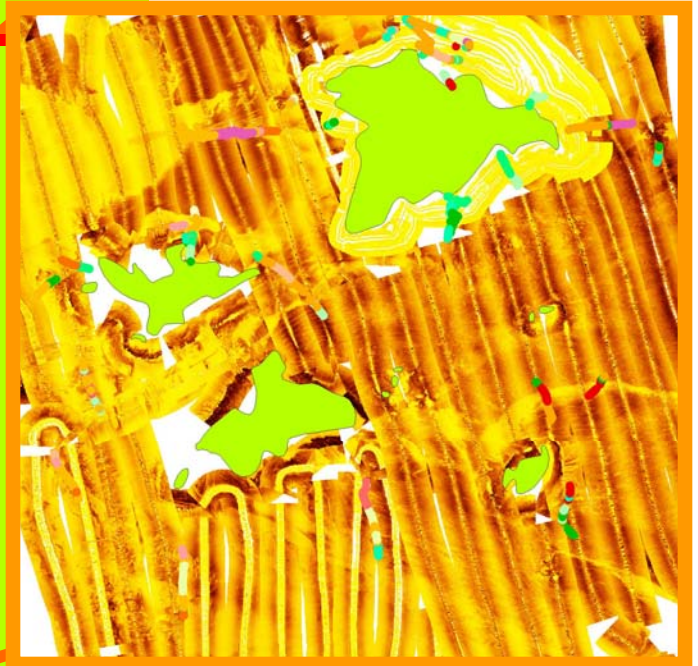
Or combinations of video and acoustic data

- **Sidescan sonar surveys (substrate), Multibeam (depth, slope)**
- **Targeted towed video**

# Drop video surveys

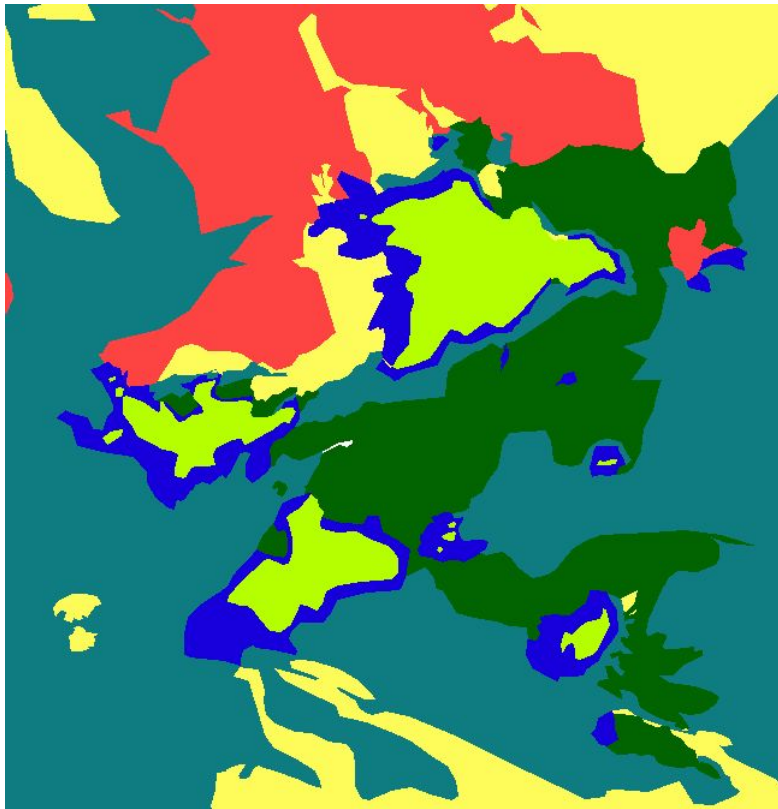


# Towed video to validate sidescan



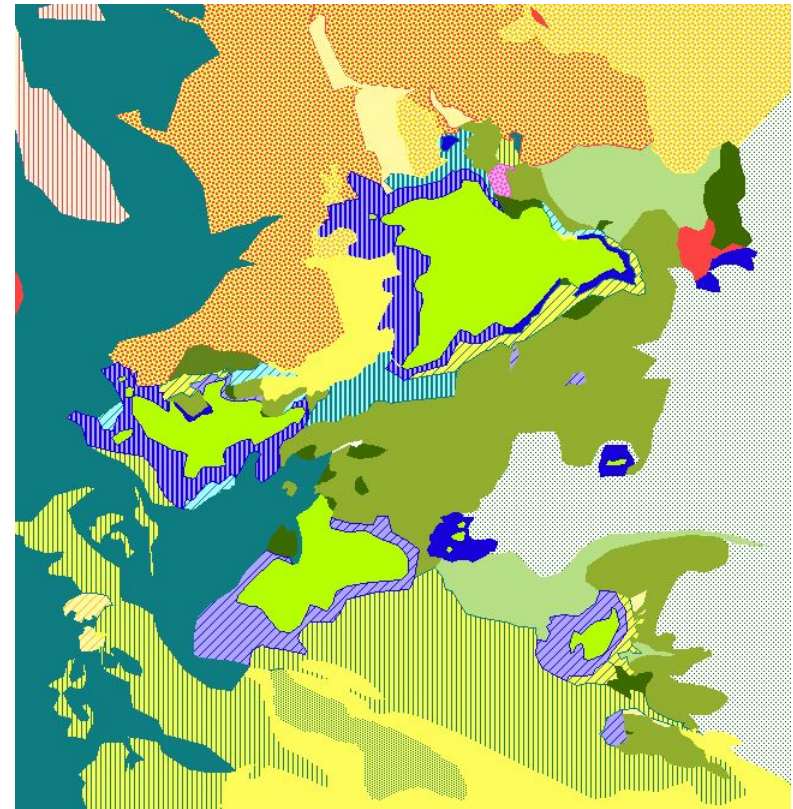
# Produce Habitat Maps

**Broad**  
(5 types)



5km

**Functional**  
(or detailed – 30 types)



# Habitat Classification - Broad Types



Seagrass



Low profile reef



High profile reef



Rhodoliths



Sand

# What we need to predict with



- Locational instances of habitat type (presence / absence data) \*
- Geophysical descriptors of habitat \*
  - eg. slope/relief
  - substrate
  - depth
  - measures of exposure
- A sensible, easy to use model that is not too complex to develop or understand

\* Implies an underlying classification scheme

## Prediction methods applied to drop video data.

Need to classify what geophysical factors control the presence of a habitat....

Methods?

- Decision Trees
- **Generalised Linear Models (GLMs)**
- ALL have their own ASSUMPTIONS and LIMITATIONS about data sampling, spacing, no's of samples, abundance or presence/absence, categorical or continuous data, how to assess model accuracy, cryptic ways of displaying results

# Habitat Models

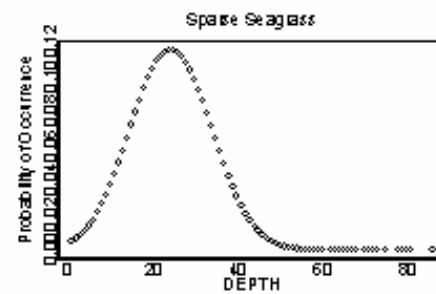
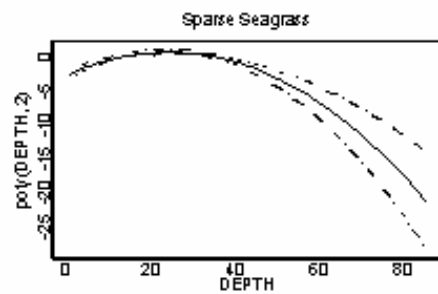
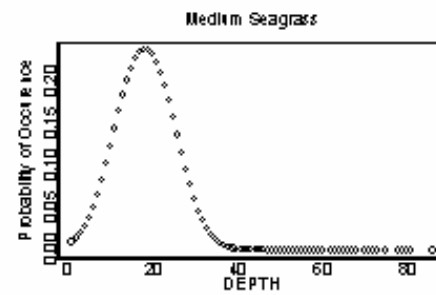
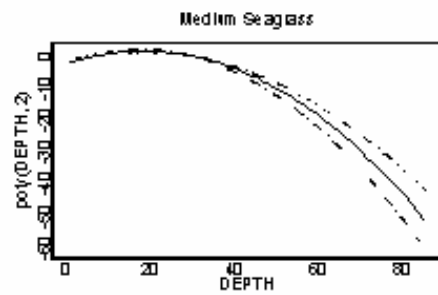
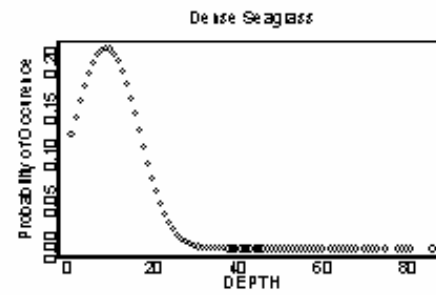
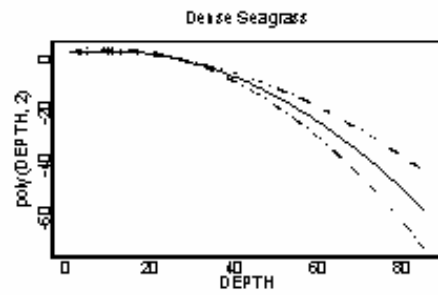
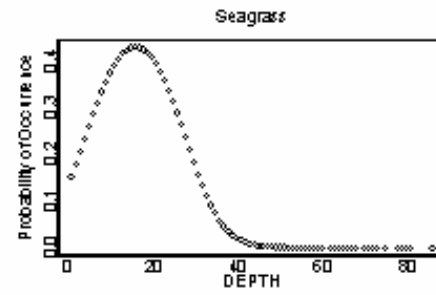
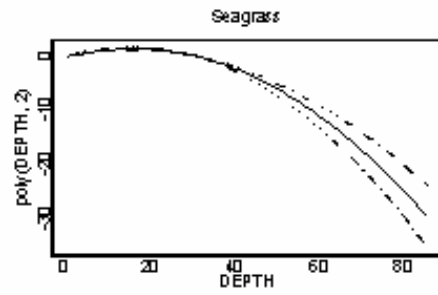


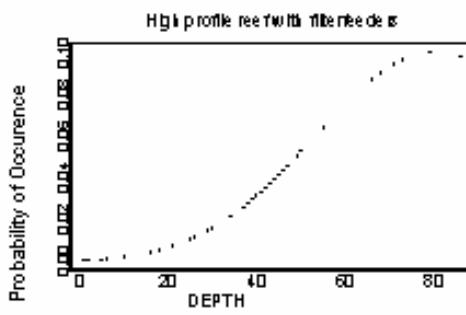
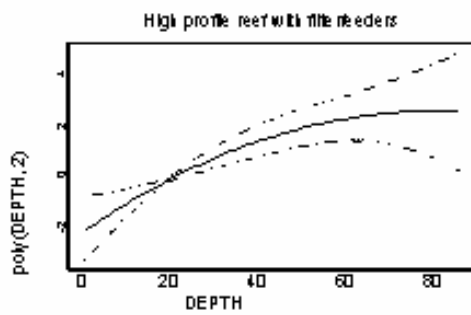
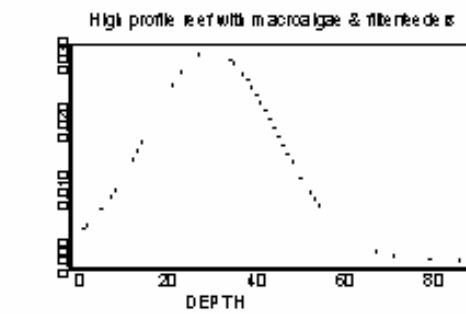
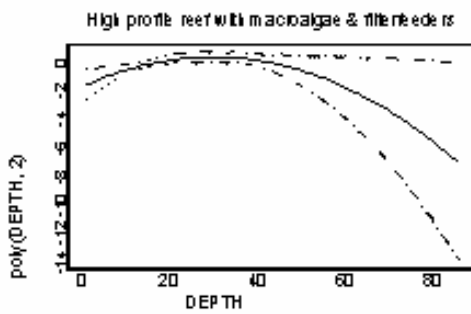
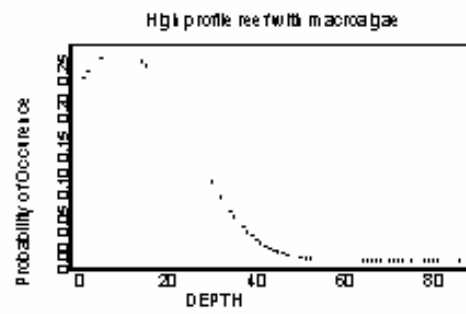
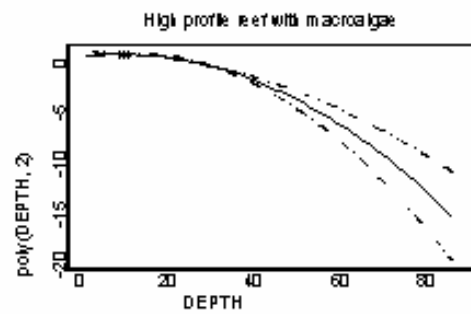
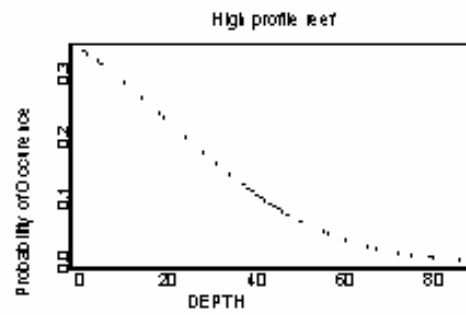
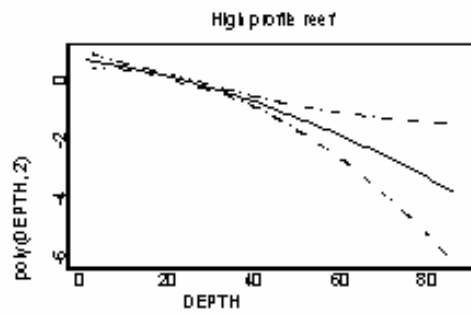
- **using video data** model both broad & functional habitat types with GLMS and DTs
- using available classified data on substrate, depth & relief, exposure indices (from oceanography data) and the habitat type
- Exposure data only available for Esperance Bay region
- **Two subsets** – all data, no exposure
  - **E Bay data with exposure**

# GLM Results

## GLMs

- Plots probability of occurrence for each factor
- Using stepwise model determines which factors are important in reducing model deviance
- can test against actual data to determine model accuracy
- Modelled results mapped as probabilities of occurrence.





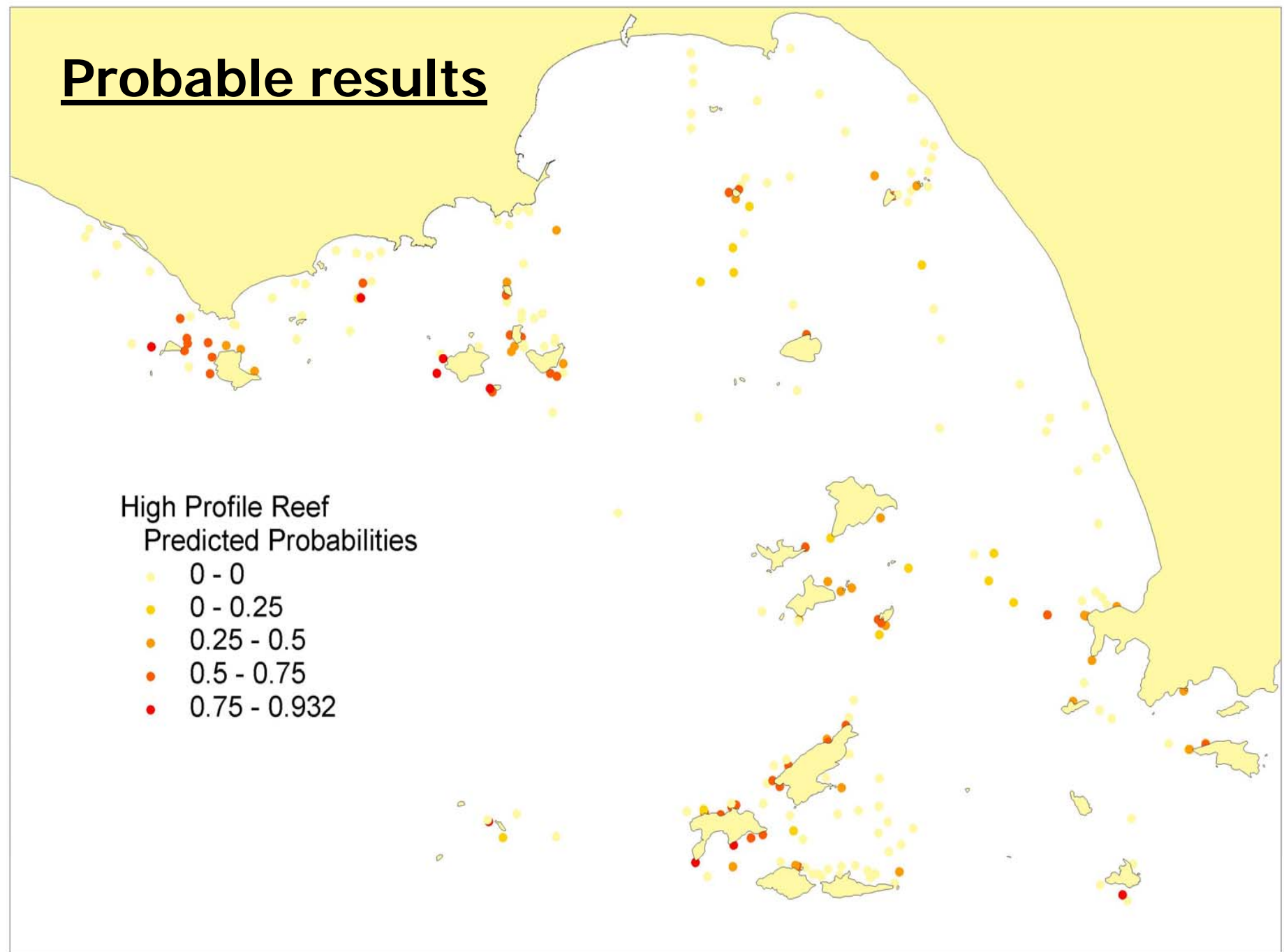
# Factors used

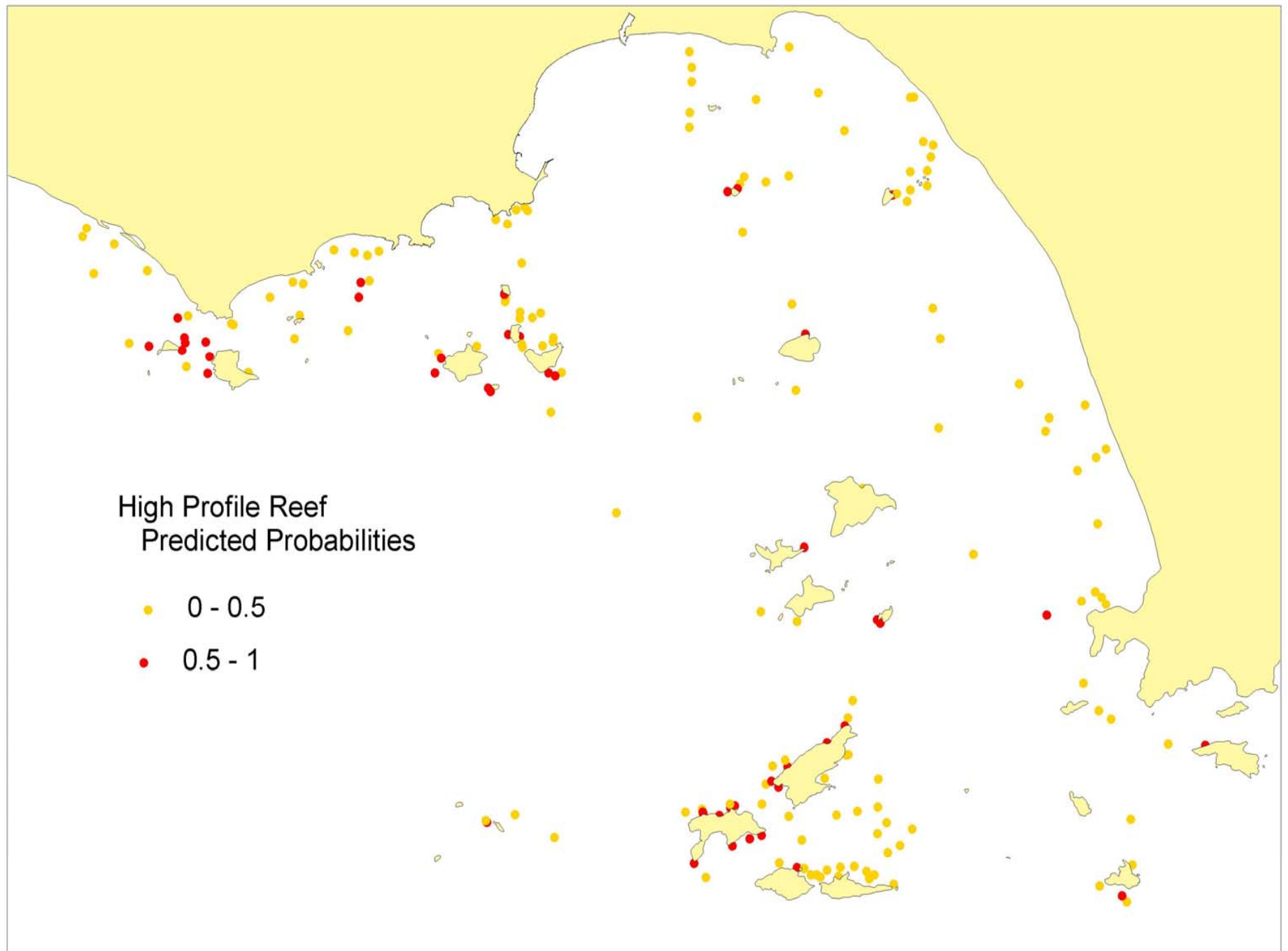
model	Dev (df)	AIC	Dsq	Final factors used
<b>High Profile Reef (Broad)</b>	286.06 (549)	304.0605	0.49	<b>DEPTH + HSMAX + SUB + ISLAND.DIS</b> all sig .05 except island dis 0.07
<b>High Profile Reef with Macroalgae</b>	257.52 (549)	275.5196	0.50	<b>DEPTH + HSMAX + SUB + ISLAND.DIS</b> all sig .05
<b>High Profile Reef with Macroalgae and Filterfeeders</b>	90.50 (549)	106.5056	0.30	<b>HSMAX + SSMAX + SUB</b> only ss.max & sub sig at .05 hsmax .15
<b>High Profile Reef with Filterfeeders</b>	12.32 (549)	34.31572	0.67	<b>ALL</b> only hs.max and island dis sig at .05 Depth,sub,ssmax not sig

# Probable results

High Profile Reef  
Predicted Probabilities

- 0 - 0
- 0 - 0.25
- 0.25 - 0.5
- 0.5 - 0.75
- 0.75 - 0.932





# Model accuracy

When you model, the final accuracy depends on the 'cut-off' value you apply to the predicted probabilities i.e.  $> 0.5$  better than chance.

Can't just use overall accuracy (0.5)

Actual predicted accuracy can be over optimised through the correct classification of absence data

Largely interested in how well it classifies presence....

# Model accuracy

Eg. HPR

		predicted	
		0	1
Actual	0	91.67%	8.33% (FP)
	1	40.43% (FN)	59.57%

Overall accuracy = 85% but true positive rate = 59.5%

# GLM considerations

- Sample size is important (broad OK, functional not always enough examples eg, hpr with filterfeeders lower probability)
- Substrate is very important in all the models
- Exposure is too... (HSMAX & SS MAX)  
(but restricted to applying models where there is enough depth data to calculate it though...)
- How we calculate accuracy affects our model interpretation

# Do the results make ecological sense?

- Functionally, some types classified well, others not so.
- The available data can only explain so much.
  - **other factors are needed** to train with
  - Are **more examples** of certain habitat types required?
- ‘Standard’ classification schemes may oversimplify complex ecologies
  - High accuracy with broad classes vs ecologically meaningful results
  - Making the data fit the scheme vs deriving scheme from data

# Problems with predicting

- enough examples or 'unique' habitats?
- How can you classify & then predict the locations of unique or endemic habitats?
  - Need 'full coverage' data to identify these habitats
  - or more data to describe them better - target for further study
- We can't classify and predict everything...

# Spatial probability maps ?

What about when we want to apply these methods to a larger area? Is it practical?

## Implications

Even if we can classify certain habitats do we have the necessary data to predict their occurrence spatially?

- depth & relief – **produces fine distinctions**, if too coarse will lose necessary detail
- models **strongly reliant on substrate**
- Many samples needed for each class to train & test the model

# Conclusions

- Can't explain everything – may need other factors or more examples to explain differences
- But helps focus in on what we don't know
- Apply similar techniques using multibeam and sidescan data
  - high res depth, relief, substrate
  - **BUT WAY UP DATA NEEDS OF PREDICTION WITH ACTUAL COSTS OF SURVEYING.**
- Compare to other modelling techniques eg. Decision Trees
  - GLMs not as self explanatory as DTs in defining the range / bounds of factors that define one habitat from another

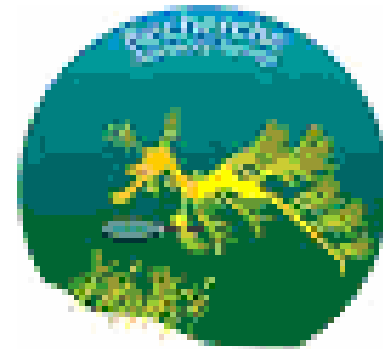
Thank you – Any questions?



THE UNIVERSITY OF  
WESTERN AUSTRALIA



FISHERIES  
RESEARCH &  
DEVELOPMENT  
CORPORATION



Esperance  
community &  
commercial  
fisherman



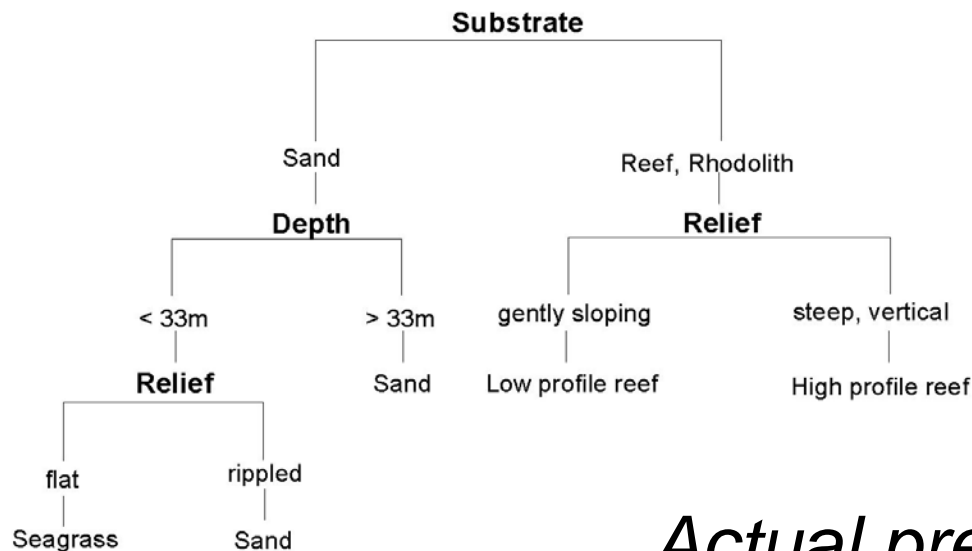
Coastal CRC

[www.marine.uwa.edu.au/recherche/](http://www.marine.uwa.edu.au/recherche/)

# Decision Tree Results

Decision Tree classifies broad types with **high overall accuracy (~80%)**

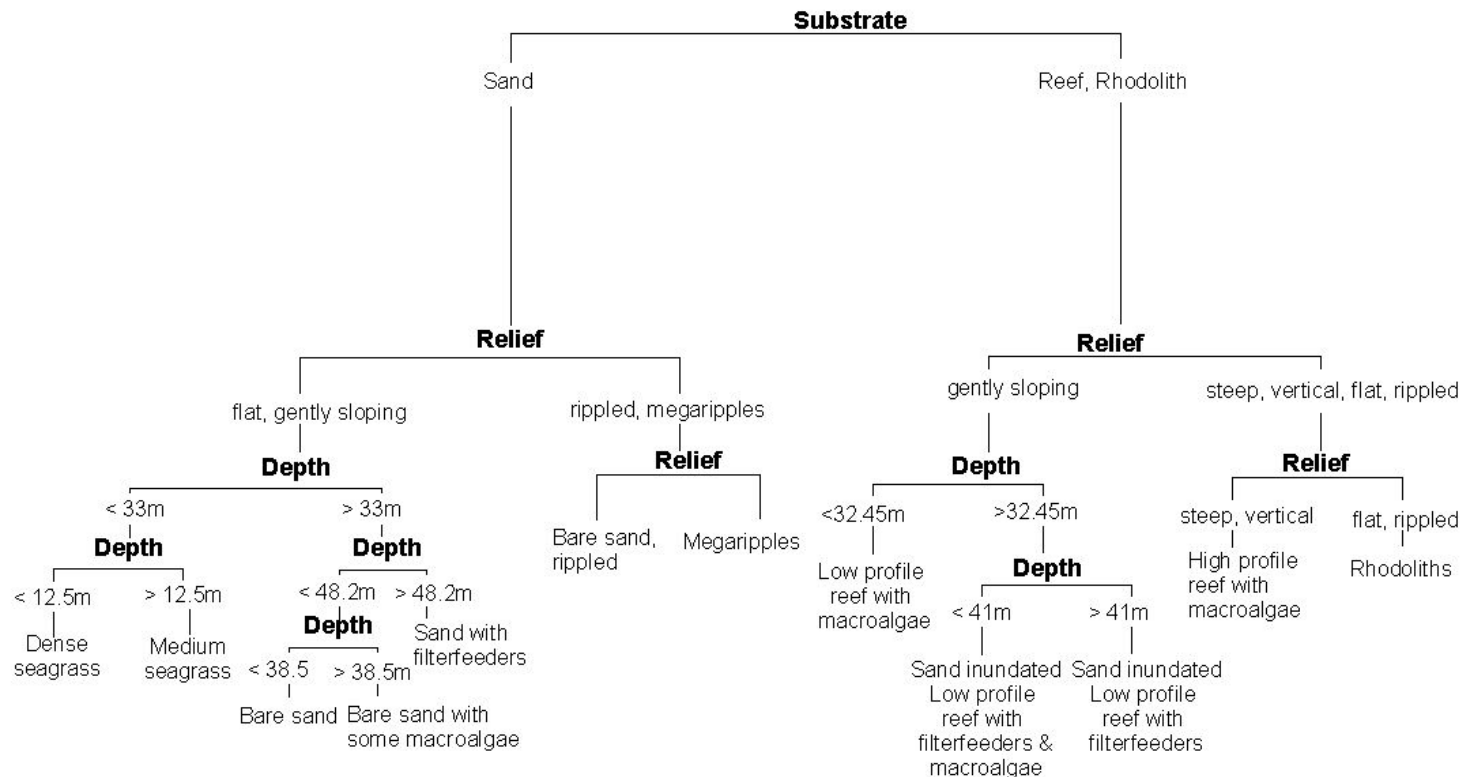
**BUT** very generalised patterns.... No surprises.



*Actual predicted accuracy ~ 58%*

# Functional Habitat Types

Lower accuracy (~50%) but useful ecological detail for some functional types.



For example – Low profile reefs

# Functional Habitat - Low Profile Reef



**Macroalgae  
species**  
( $<32.45\text{m}$ )

**Macroalgae  
& filterfeeders**  
( $>32.45\text{m}$   
and  $<41\text{m}$ )

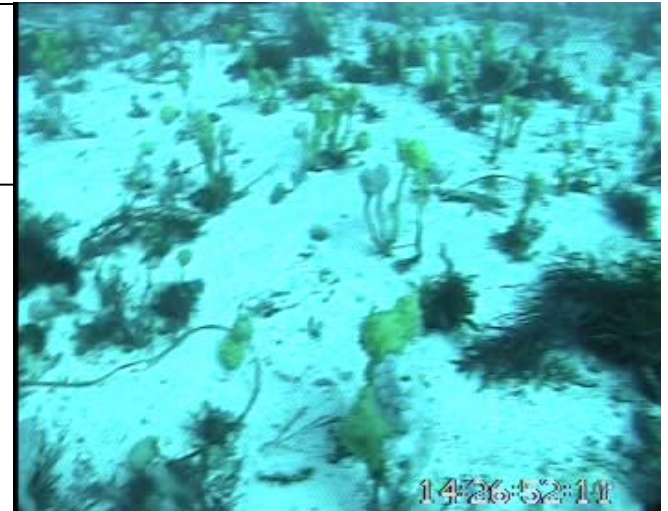


## Different functional types



**Sand inundated  
with filterfeeders**  
( $>41\text{m}$ )

**Filterfeeders**  
( $>41\text{m}$ )



# Functional Habitat - High Profile Reef



**with macroalgae**



## Different functional types



**with filterfeeders**

