# Testing the accuracy of species distribution models using species records from a new field survey

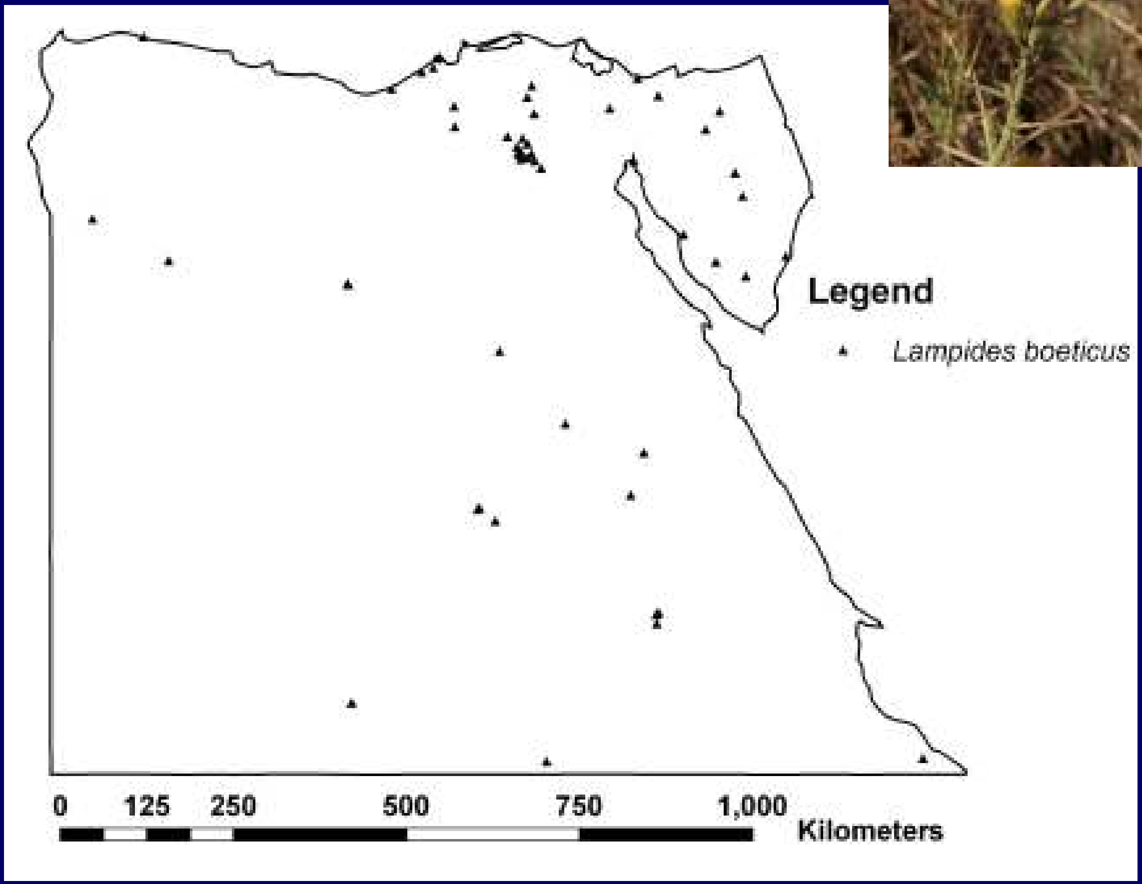
Tim Newbold, Tom Reader, Ahmed El-Gabbas, Wiebke Berg, Wael Shohdi, Samy Zalat, Sherif Baha-El-Din, Francis Gilbert

# Outline

- Species distribution models
- Museum data
- Evaluating distribution models
- Variation in model accuracy among species
- Egypt field survey
- Results

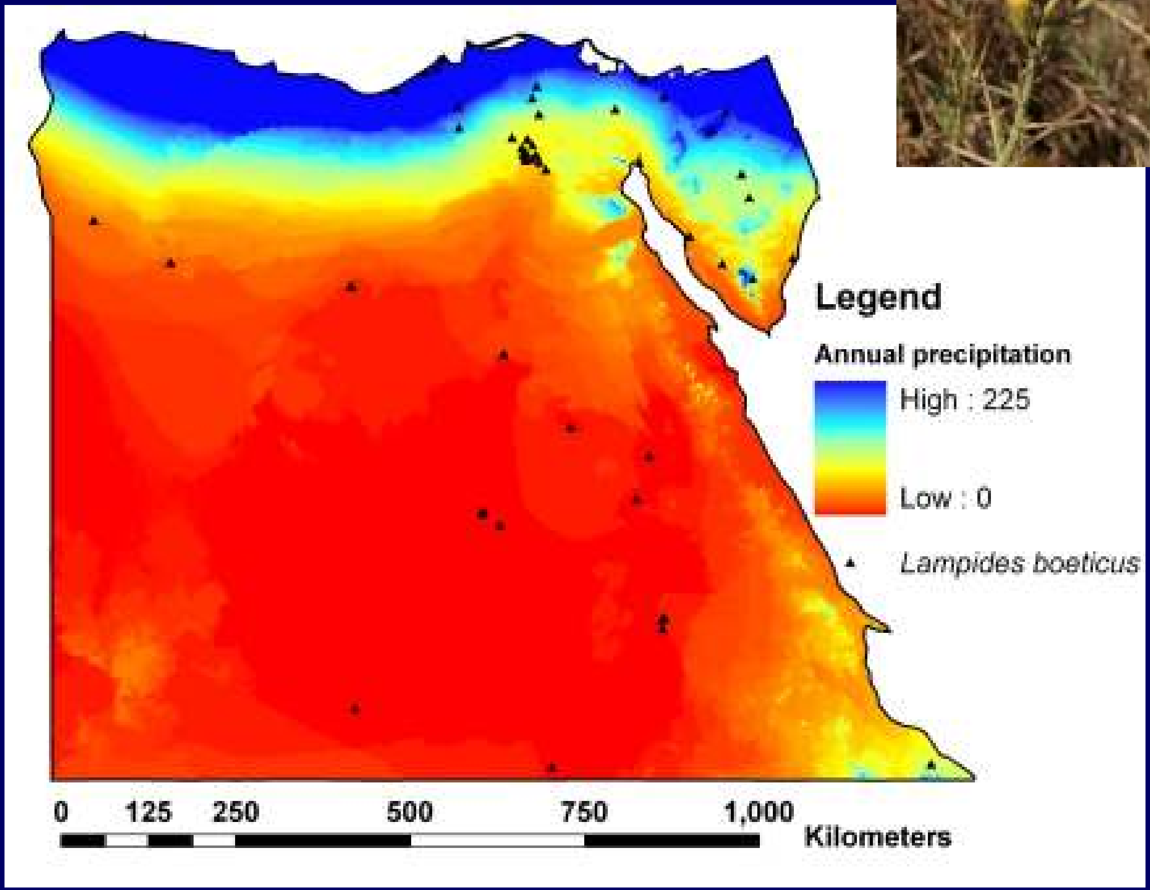
# Species distribution models

- Recorded occurrences



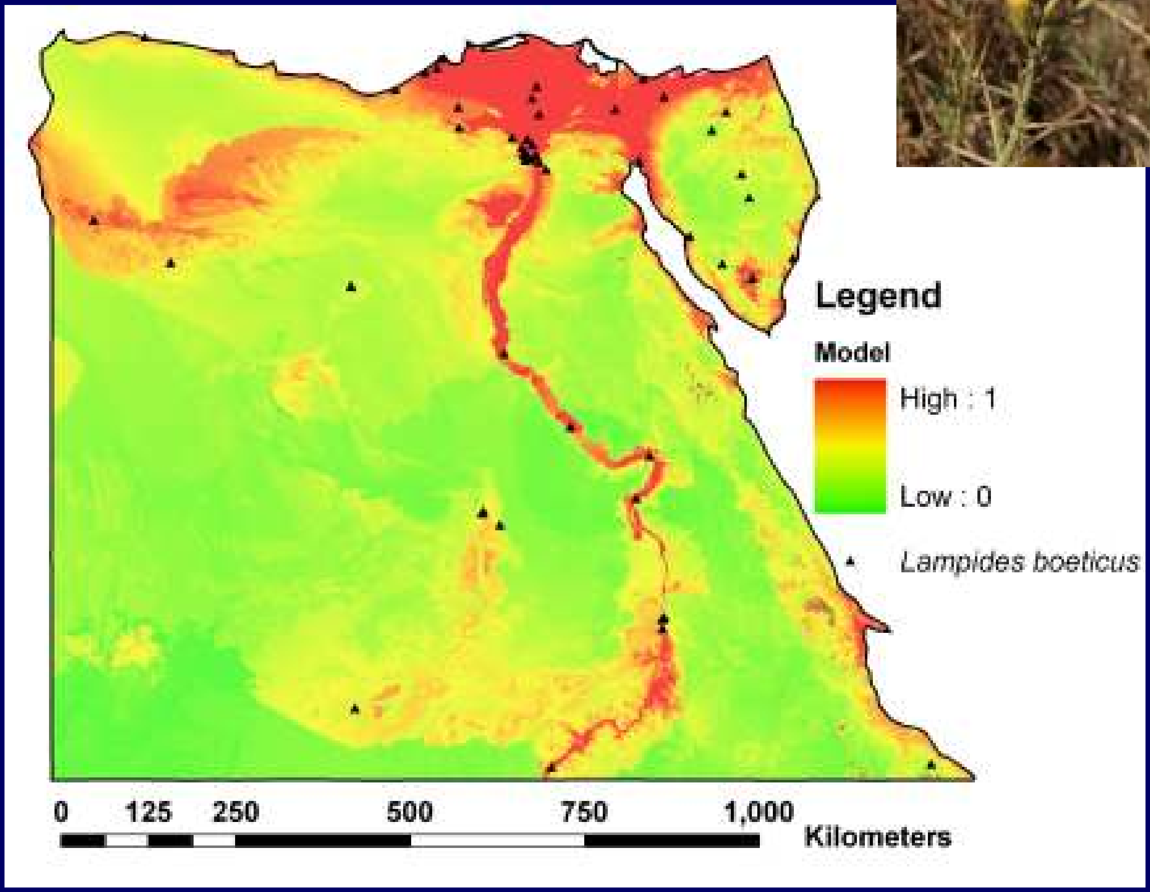

# Species distribution models

- Environmental variables




# Species distribution models

- Model prediction



# Applications of SDMs

- Conservation planning, e.g. protected areas (Thorn et al., 2009, *Div & Distns;* Newbold *et al., 2009, J Biogeog*)
- Finding new populations of species (Raxworthy et al., 2003, *Nature*)
- Predicting impacts of climate change & land-use change (Thomas et al., 2004, *Nature*)
- Ecological/evolutionary questions (Peterson et al., 1999, *Science*; Eaton et al., 2008, *Biol J Linnean Soc*)


NATIONAL GEOGRAPHIC NEWS
REPORTING YOUR WORLD DAILY
Tuesday, June 16, 2009

MAIN | ANIMAL NEWS | ANCIENT WORLD | ENVIRONMENT NEWS | CULTURES NEWS | SPACE/TECH NEWS | WEIRD | PHOTOS | VIDEO

**Computer Model May Identify Conservation Hot Spots**

Hilary Mayell
for National Geographic News
January 20, 2004

What do long-dead animals, satellite technology, and biologists counting lizards have in common? Together they may be able to help predict population distributions of animals threatened with extinction.

This is particularly important as the rate of habitat loss accelerates in regions where very little data is available.

LATEST PHOTO NEWS
LEOPARD PICTURES: Rare Snow Cats Caught by Camera Traps
LATEST VIDEO NEWS
New Terra-Cotta Warriors Dig Begins in China
Most Viewed News
ADVERTISEMENT
TRAVELER


Catherine Bennett in G2 | In G2 | In G2 | Plus Online & jobs

55p
Thursday
January 8 2004
Published in London
and Manchester
guardian.co.uk

**The Guardian**

**An unnatural disaster**

- Global warming to kill off 1m species
- Scientists shocked by results of research
- Third of life forms doomed by 2050

Paul Brown
Environment correspondent

Climate change over the next 50 years is expected to drive a quarter of land animals and plants into extinction, according to the first comprehensive study into the effect of higher temperatures on the natural world.

The sheer scale of the disaster facing the planet shocked those involved in the research. They estimate that more than 1 million species will be lost by 2050.

The results are described as "terrifying" by Chris Thomas, professor of conservation biology at Leeds University, who is lead author of the research from four continents published today in the magazine Nature.

Much of that loss – more than one in 10 of all plants and animals – is already irre-

tralia, Central and South America, and South Africa, showed that species living in mountainous areas had a greater chance of survival because they could simply move uphill to get cooler.

Those in flatter areas such as Brazil, Mexico and Australia, were more vulnerable, faced with the impossible task of moving thousands of miles to find suitable conditions.

Birds, which had the greatest chance of escape, could in theory move to a more suitable climate but the trees and other habitat they needed for survival could not keep pace and all would die.

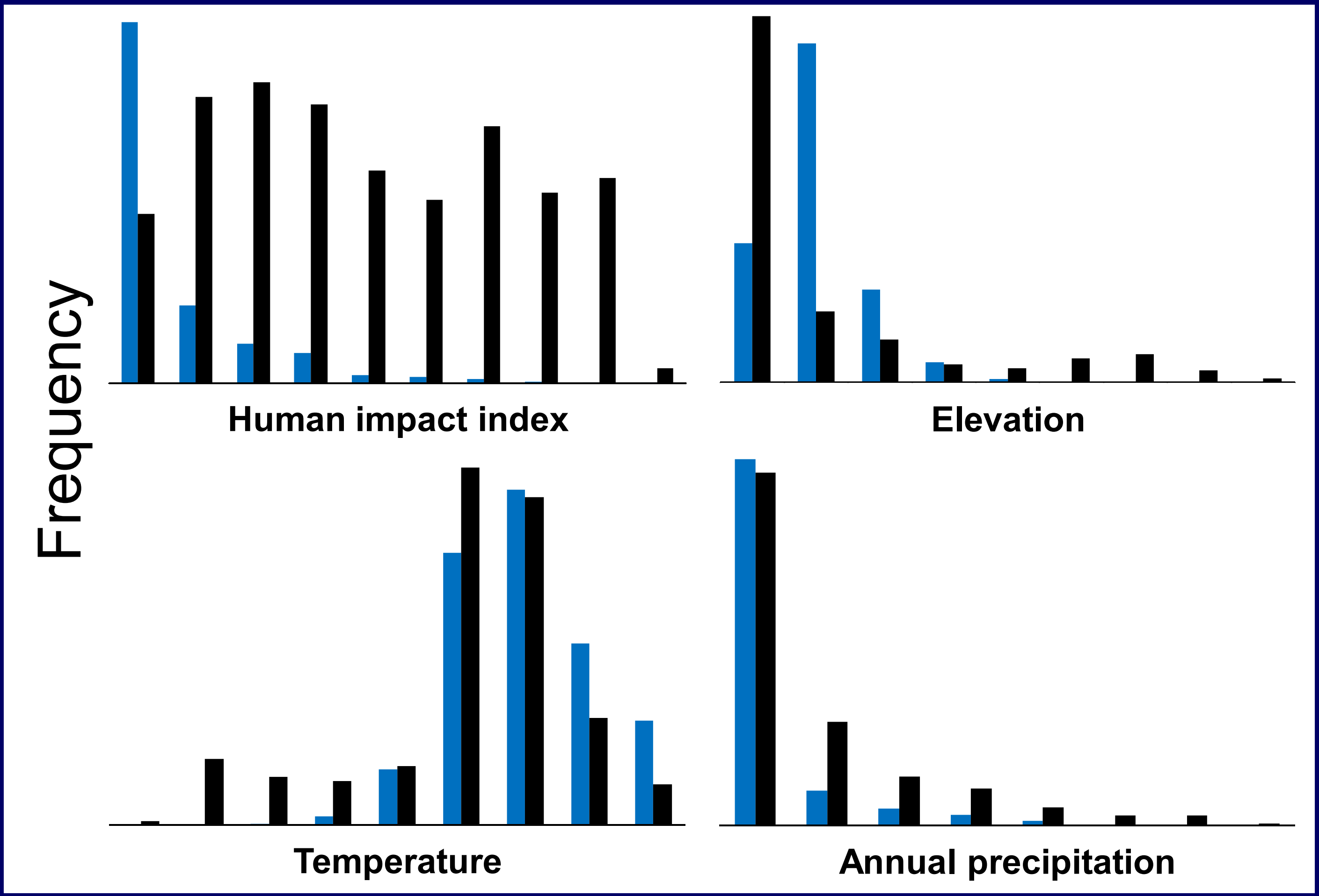
Professor Thomas said: "When scientists set about research they hope to come up with definite results, but what we found we wish we had not. It was far, far worse than we thought, and what we have discovered may even be an

# Museum data

- Valuable source of species records
- Errors
- Biases:
  - Spatial
  - Temporal
  - Taxonomic
- Environmental bias → Poor distribution models
- BioMAP data for Egypt

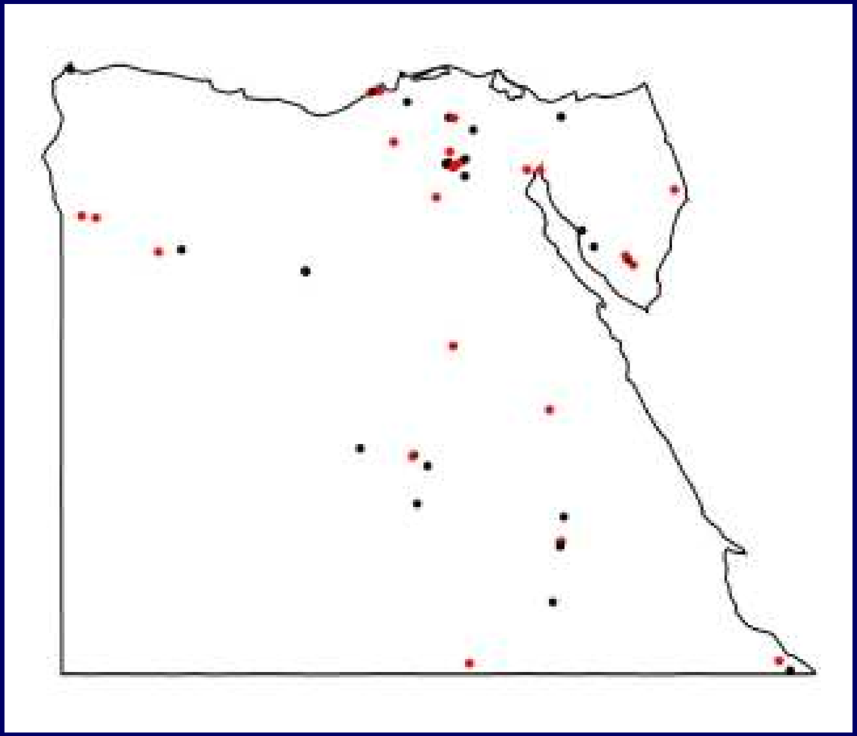
# Bias in BioMAP butterflies data

Black bars = butterfly sites; blue bars = all grid cells



# Evaluating distribution models

- Common practice is to divide sightings of species for model development and evaluation
- Gives over-optimistic estimates of accuracy with biased data
- Better to collect new data

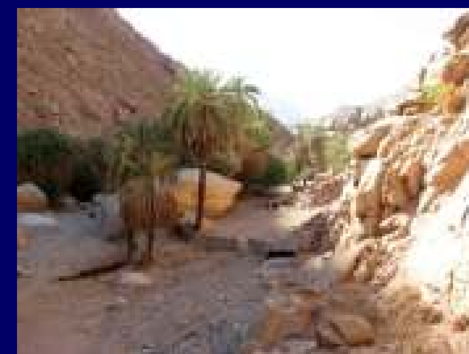


# Methods

- Developed distribution models for Egyptian reptile & amphibian (n = 20), butterfly (n = 10) and mammals (n = 4)
- Environmental variables - temperature, rainfall, elevation, habitat
- Compared 2 methods of model evaluation: 1) Dividing BioMAP records into 2 halves; 2) Collecting new data
- AUC statistic

# Methods

- Field surveys May-July 2007 and 2008
- Impossible to survey randomly
- Sampled as many habitats as possible
- 21 sites
- 4 walking transects at each site



# Model accuracy

- Species detectability:
  - Species missed
  - Less complete distribution data
- Outcome depends on probability of occurrence (Ψ) and probability of detection (p)
- Modelled Ψ and p using maximum likelihood

| | |
|---|---|
| 1001 | 1100 |
| 1100 | 0001 |
| 1000 | 0101 |
| 0000 | 1010 |
| 1111 | 1000 |
| 1110 | 0001 |
| 1000 | 0011 |
| 1100 | 1001 |
| 1001 | 1111 |
| 1111 | 1010 |

# Model accuracy

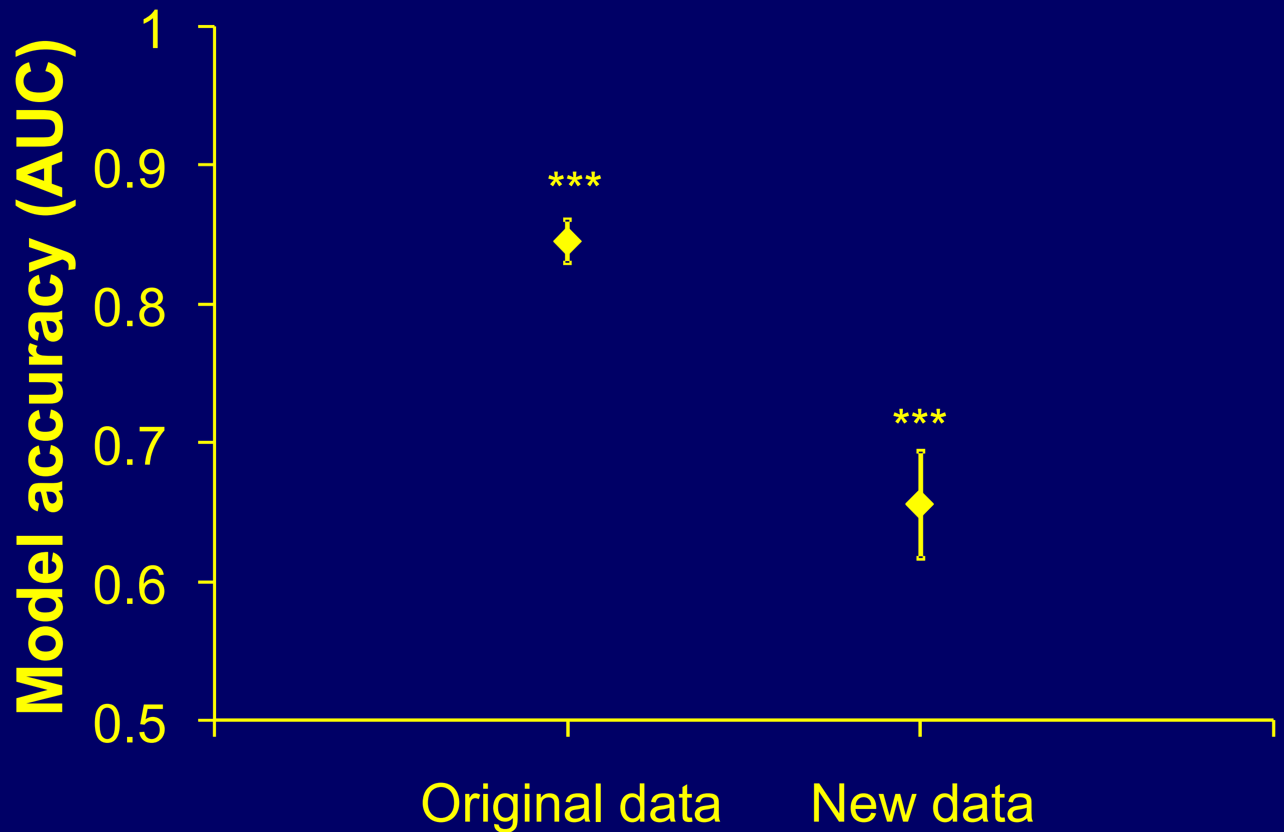
- Species characteristics:
  - Niche breadth
  - Range size
  - Migratory behaviour
  - Mobility
- Tested the effect of range size and mobility
- For tests of other characteristics, see Newbold et al., 2009, Biodiv & Conserv

# Results

- Detection probabilities ranged from < 0.001 to c. 0.75
- Snakes, mammals and migrant butterflies had low detectability
- Lizards, most butterflies and the Dorcas gazelle (faeces and tracks) were highly detectable

# Results

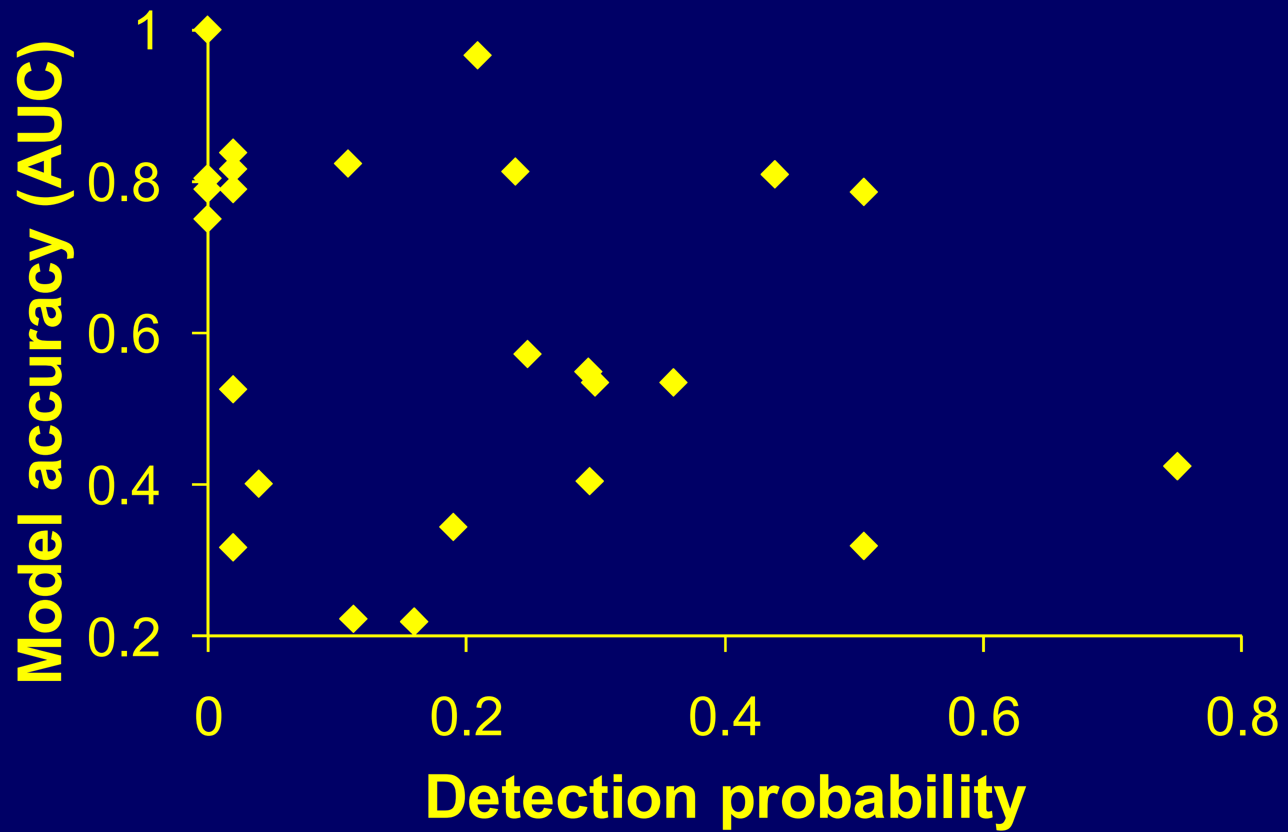
- Estimates of model accuracy lower using new data
- But both estimates were sig. better than random




***One-sample t-test: sig. better than random (AUC = 0.5) p < 0.001

# Results

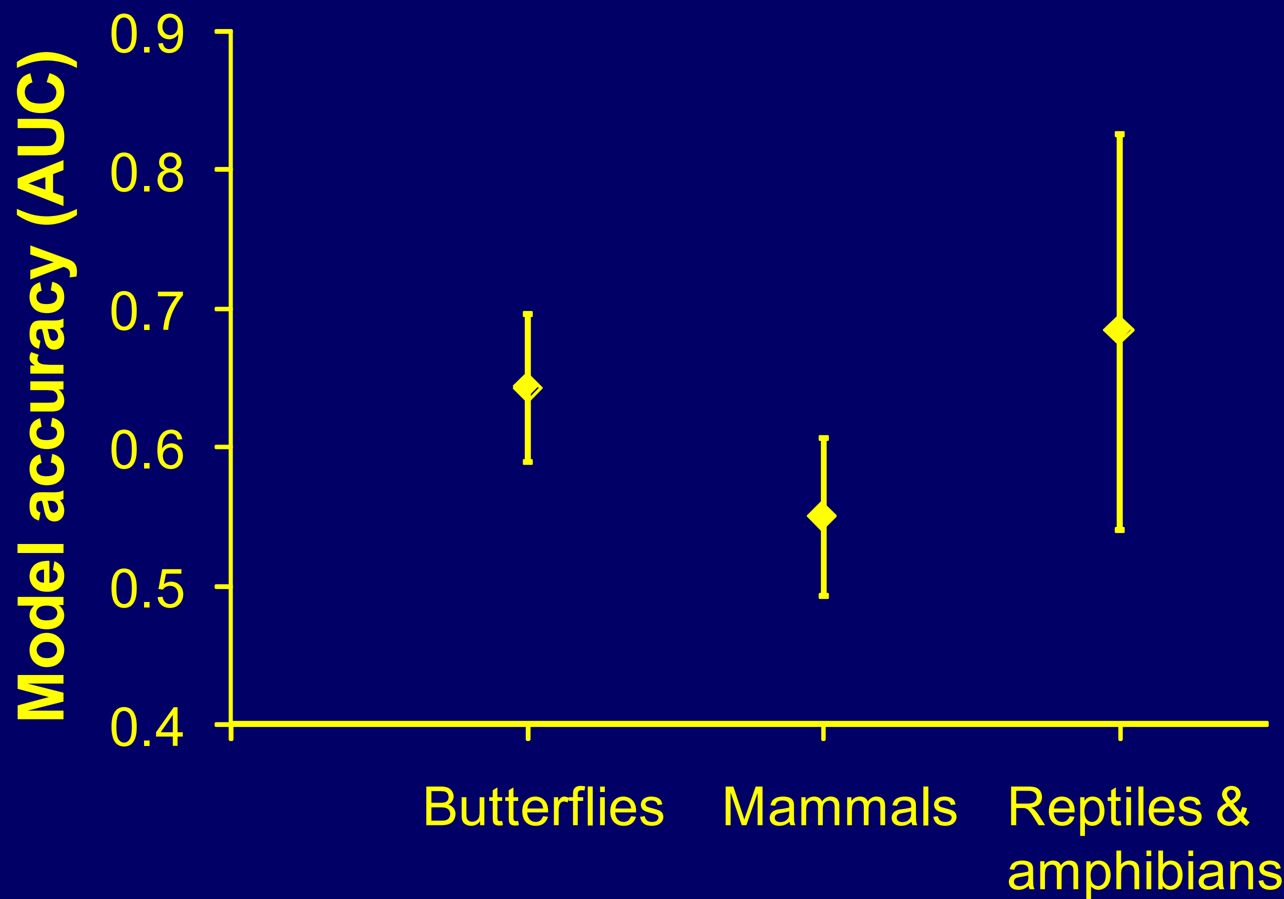
- Model accuracy was not related to detection probability




Spearman's rank correlation: $r_s = -0.294$, $n = 25$, $P = 0.154$

# Results

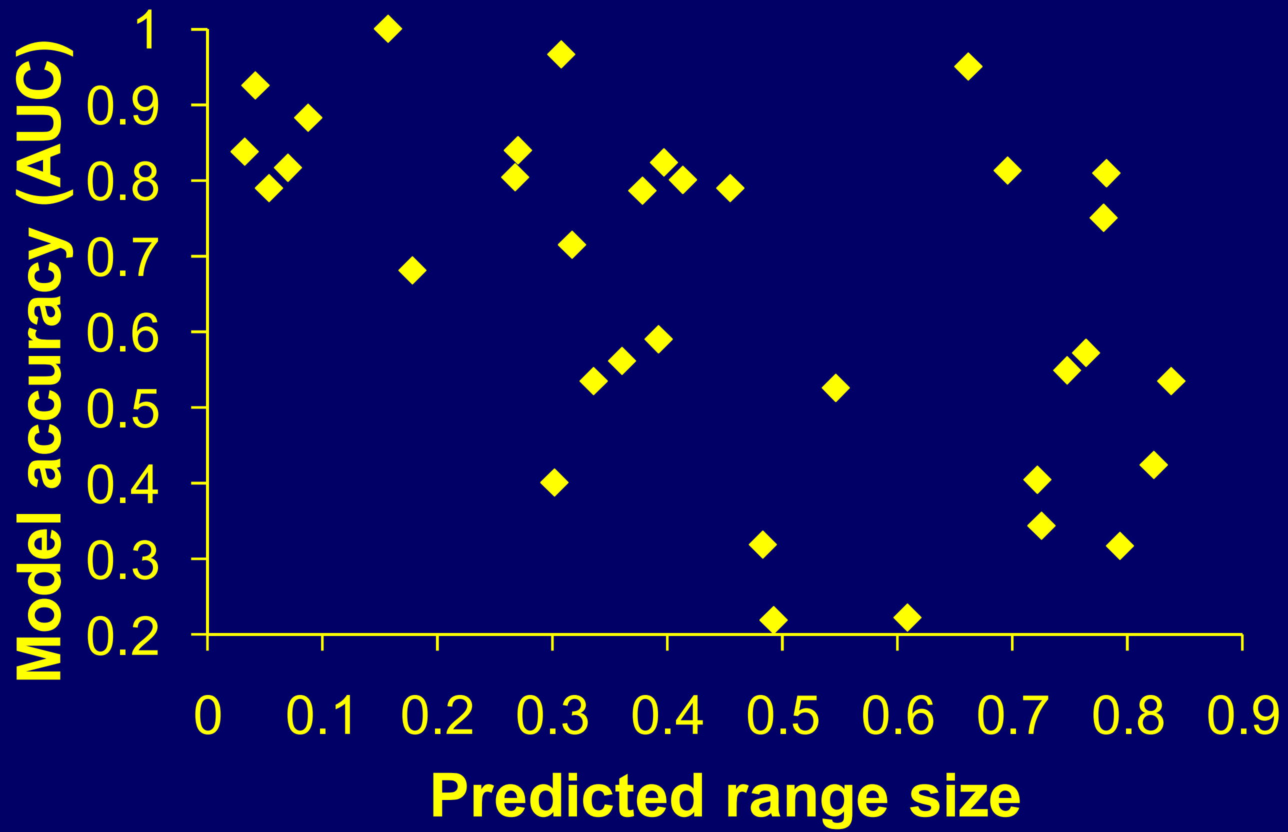
- Model accuracy didn't vary among taxonomic groups




Sum of AIC weights = 0.172

# Results

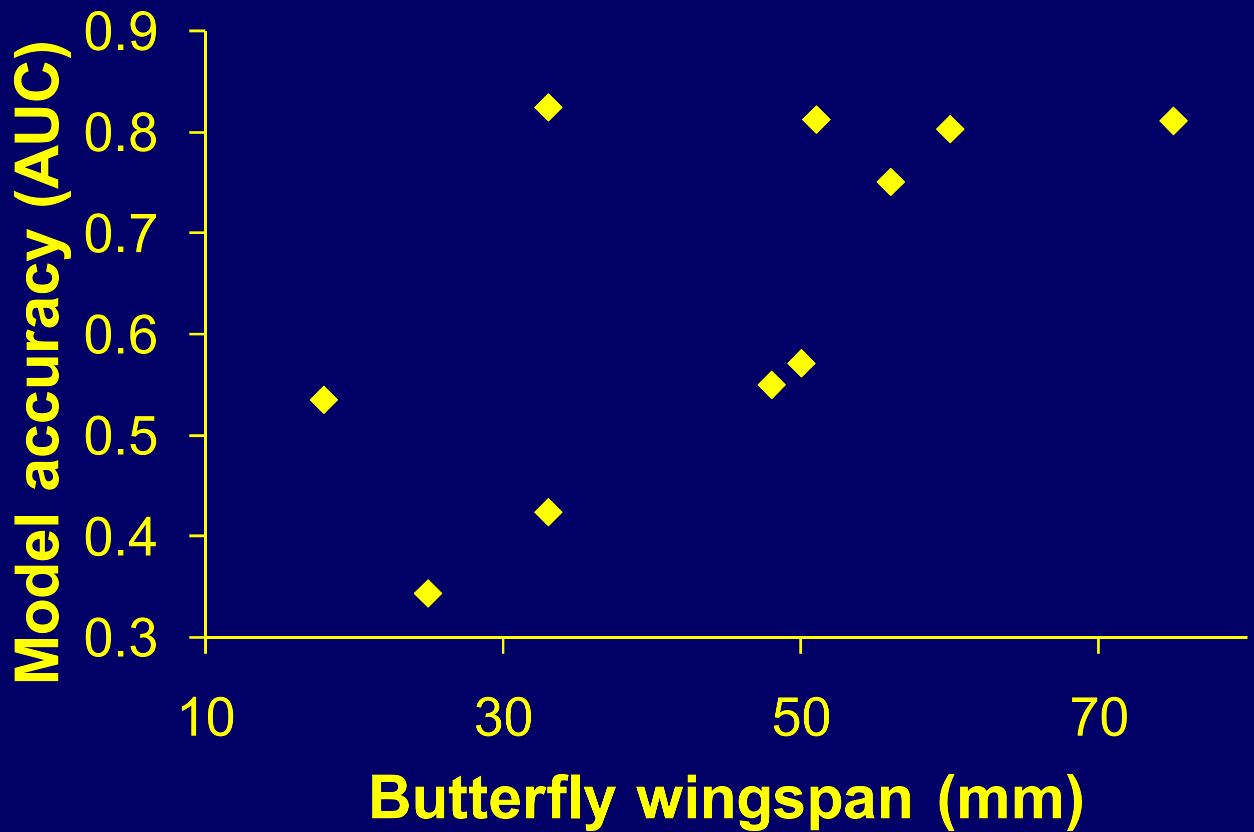
- Species with larger ranges had less accurate models




Sum of AIC weights = 0.952

# Results

- Larger butterflies had more accurate models
- Detectability or mobility?




Pearson’s correlation test: $r = 0.652$, $n = 10$, $p = 0.041$

# Conclusions

- Species distribution models were generally very accurate
- Important to collect new field data to validate models
- Model accuracy not related to detectability
- But did vary among species
- Reveals differences among species in response to environment

# Acknowledgements


- Rashed Refaey, Ahmed Refaey & other field guides
- BioMAP staff
- Behavioural ecology group, Nottingham
- NERC
- Nottingham University


The University of Nottingham

NATURAL ENVIRONMENT RESEARCH COUNCIL

# Methods

$$L = \left[\psi^{n.} \prod_{t=1}^{t=4} p^{n_t} (1-p)^{n.-n_t}\right] \times \left[\psi \prod_{t=1}^{t=4} (1-p) + (1-\psi)\right]^{N-n.}$$

- Ψ = probability of occurrence
- p = probability of detection
- t = transect number
- $n_t$ = number of sites at which the species was detected on transect t
- n. = number of sites at which species was recorded on one transect
- N = total number of sites

- Ψ and p were estimated by maximum likelihood