Global Consensus Land Cover Data for Spatial Biodiversity Research

Mao-Ning Tuanmu & Walter Jetz

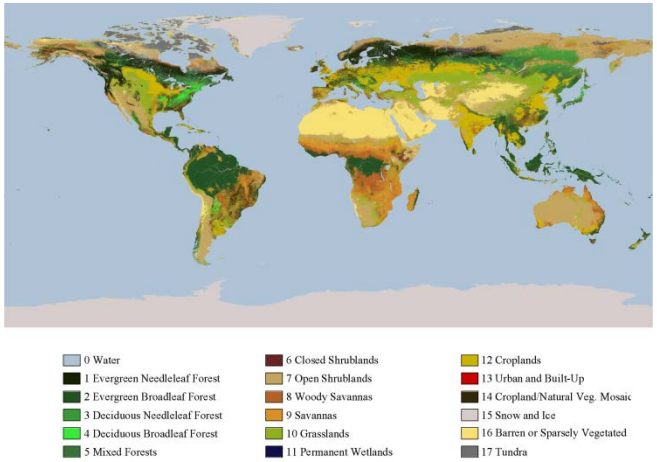
Department of Ecology and
Evolutionary Biology, Yale University

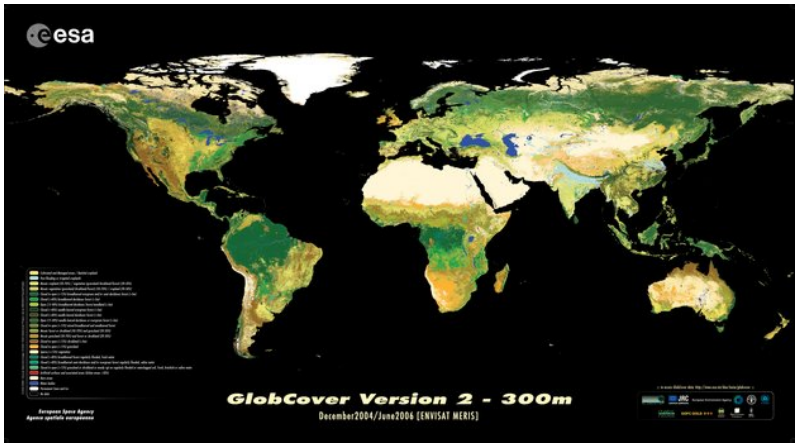
Land Cover

- Key role in the Earth system
  - Affecting material and energy flows
  - Connecting human and natural systems
- Land cover information
  - Dynamics of the Earth system
  - Resource management
  - Biodiversity and ecological processes

Remotely Sensed Land Cover

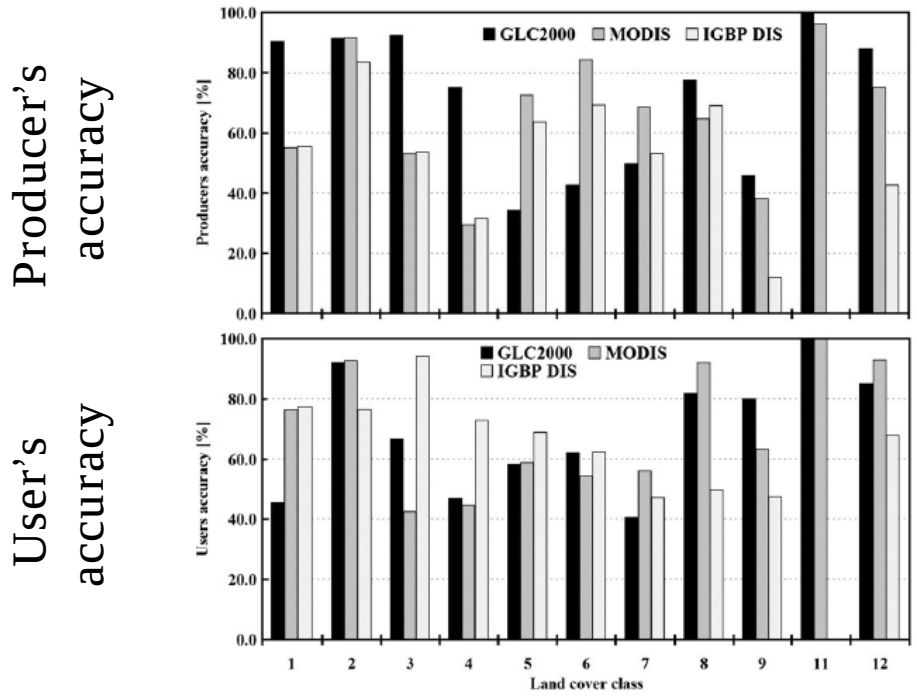
- Remote sensing
  - Spatially continuous and temporally consistent observations
- Global land cover products
  - IGBP DISCover, GLC2000, MODIS, GlobCover…





Limitations of Existing Products

- Classification errors
  - Variation among land cover classes and products

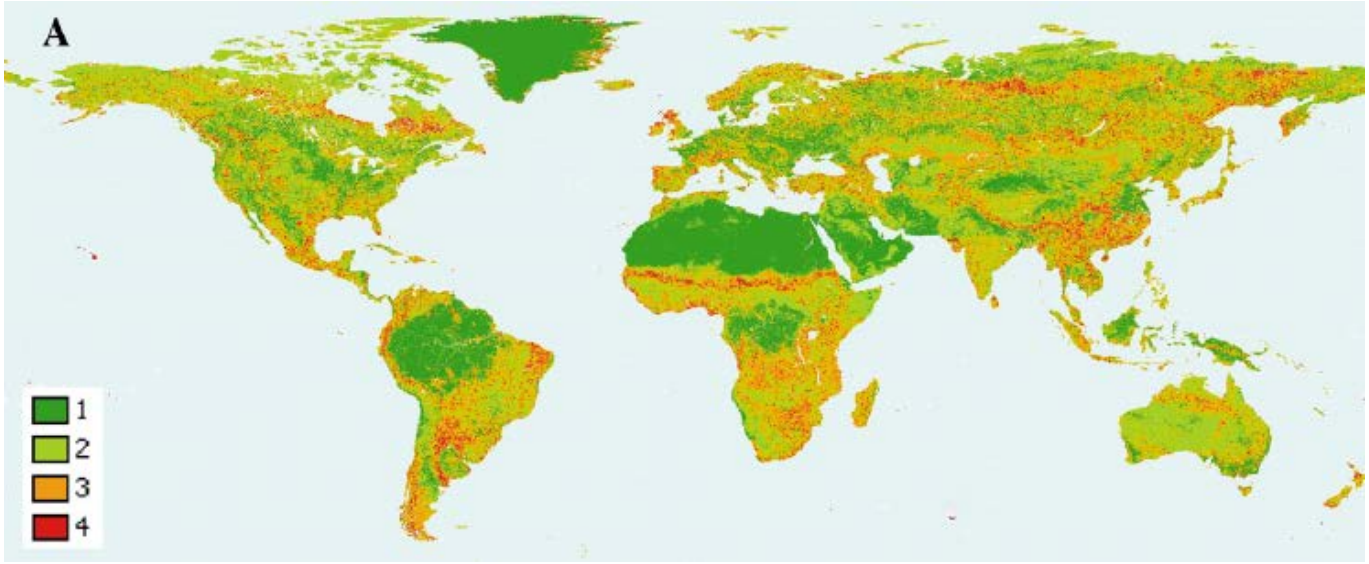


Land cover classes

Herold et al. 2008

Limitations of Existing Products

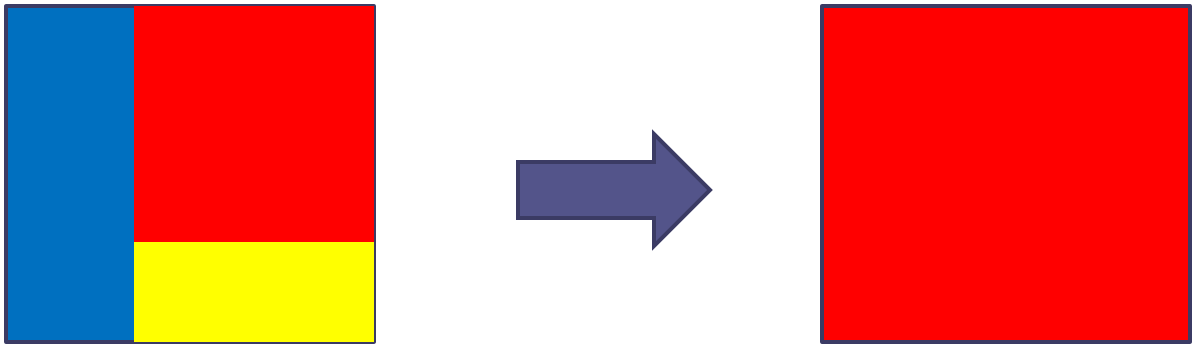
- Classification errors
- Among-product disagreements



IGBP DISCover, U of Maryland, GLC2000 and MODIS; Herold et al. 2008

Limitations of Existing Products

- Classification errors
- Among-product disagreements
- Categorical data – False absences of minor land cover classes



Goal

- To generate a **harmonized** set of 1-km resolution land cover product that provides **scale-integrated** and **accuracy-weighted** consensus land cover information on a **continuous scale**

Specific Objectives

- To generate a consensus dataset from four existing global land cover products
- To evaluate the ability of the consensus and the four input products to **capture sub-pixel land cover information**
- To examine the utility of these products for **modeling species distributions**

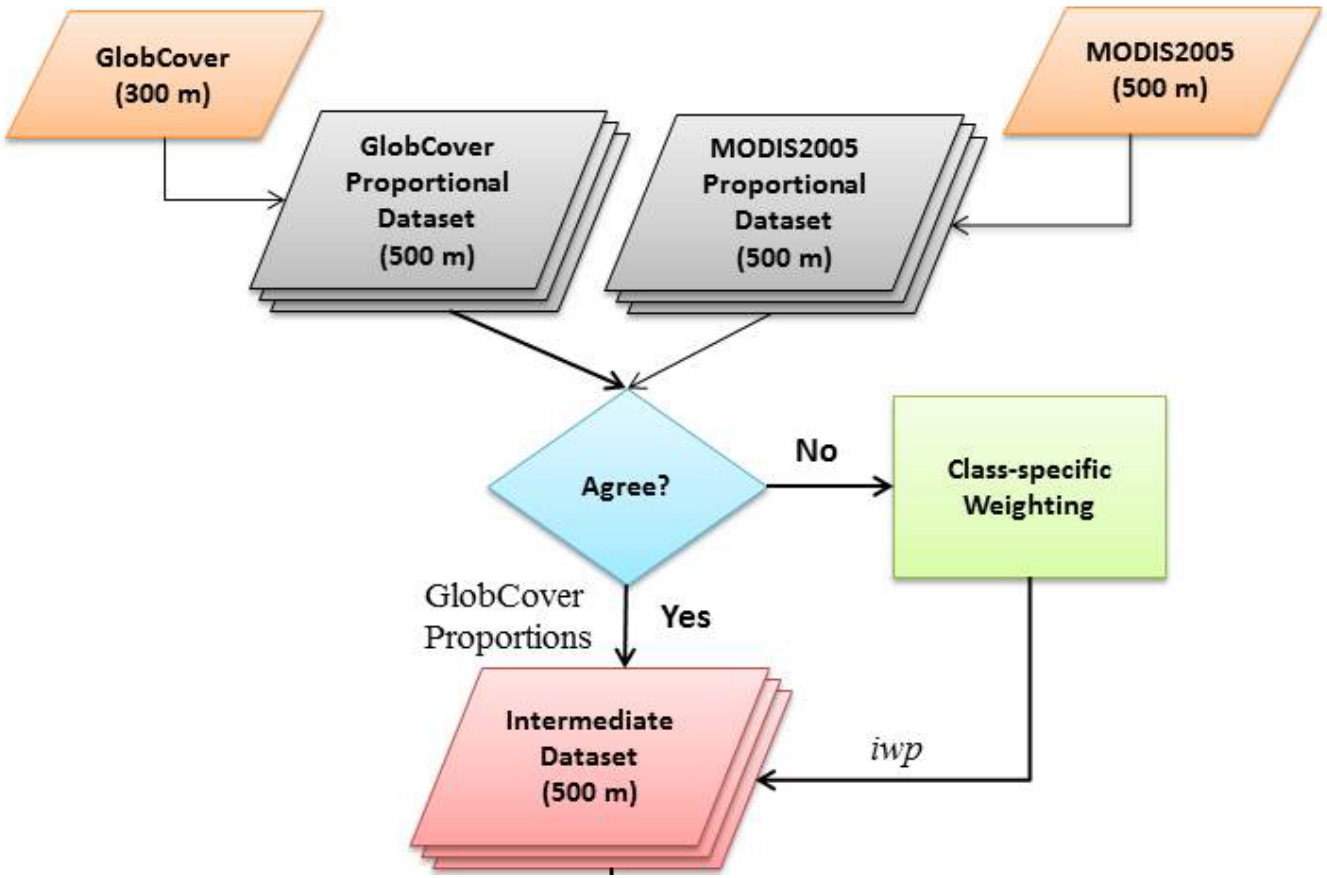
Four Global Land Cover Products

	DISCover	GLC2000	MODIS2005	GlobCover
Sensor	AVHRR	VEGETATION	MODIS	MERIS
Satellite	NOAA	SPOT	Aqua, Terra	ENVISAT
Image Acquisition Time	Apr 1992 - Mar 1993	Nov 1999 - Dec 2000	2005	Dec 2004 - Jun 2006
Input Data	Monthly NDVI composites	Diverse composites of reflectance in four spectral bands, NDVI and/or derived metrics	32-day composites and annual metrics of nadir BRDF-adjusted reflectance in bands 1 - 7, EVI and LST	Bi-monthly surface reflectance composites of 13 spectral bands
Classification Technique	Unsupervised classification	Flexible classification depending on the responsible institutions	Supervised classification decision tree	Per-pixel supervised and unsupervised classification; Per-cluster unsupervised classification
Processing Sequence	Continent-by-continent	Region-by-region	Global	Region-by-region
Classification Scheme	IGBP; 17 classes	LCCS-based; 22 classes	IGBP; 17 classes	LCCS-based; 22 classes
Spatial Resolution	1 km	1 km	500 m	300 m
Overall Accuracy	66.9% (Scepan, 1999)	68.8% (Mayaux etal., 2006)	75% (Friedl et al., 2010)	73.1% (Bicheron et al., 2008)

Generalized Classification Scheme

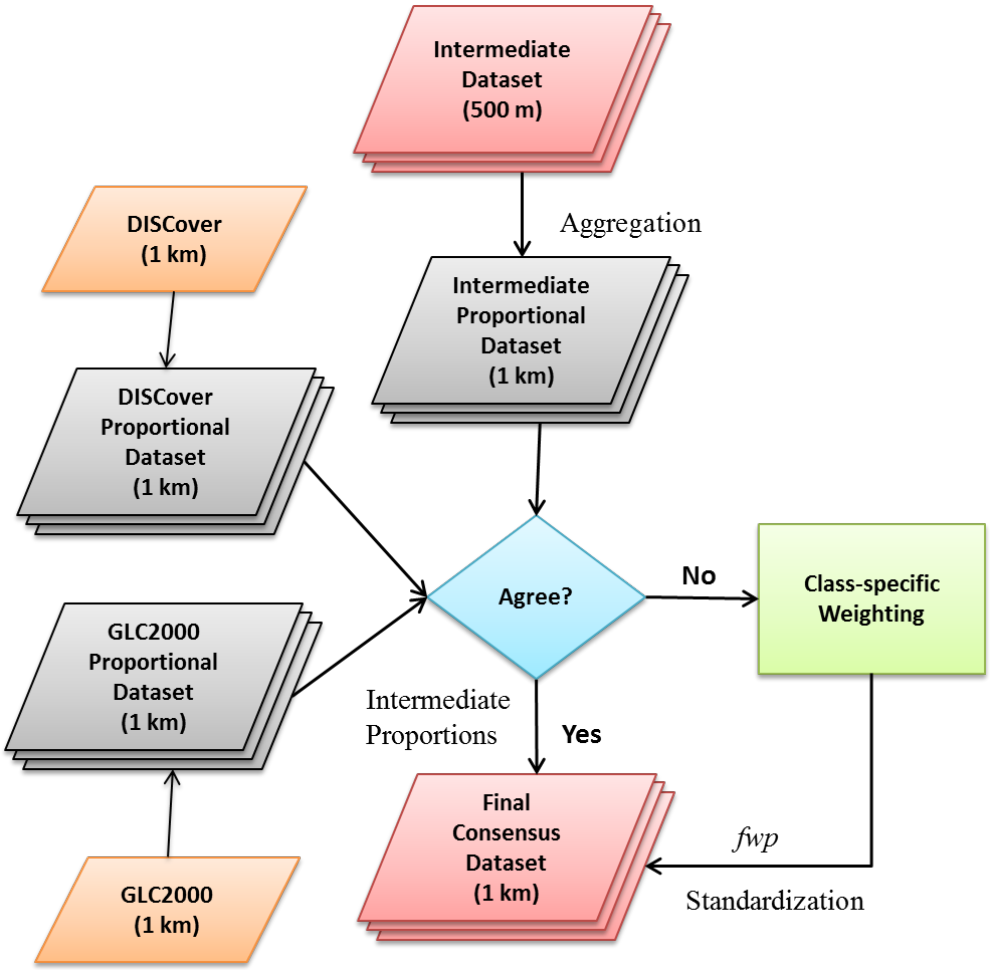
- Harmonization of classification schemes
  - 12 generalized classes (Herold et al. 2008)
    - Evergreen/deciduous needleleaf trees
    - Deciduous broadleaf trees
    - Evergreen broadleaf trees
    - Mixed/other trees
    - Shrubs
    - Herbaceous vegetation
    - Cultivated/managed vegetation
    - Regularly flooded vegetation
    - Urban/built-up areas
    - Snow/ice
    - Barren areas
    - Open water

Product Integration



$$iwp_i = 0.5 \cdot (\sum_j p_{Gj} \cdot w_{Gij} + \sum_j p_{Mj} \cdot w_{Mij})$$

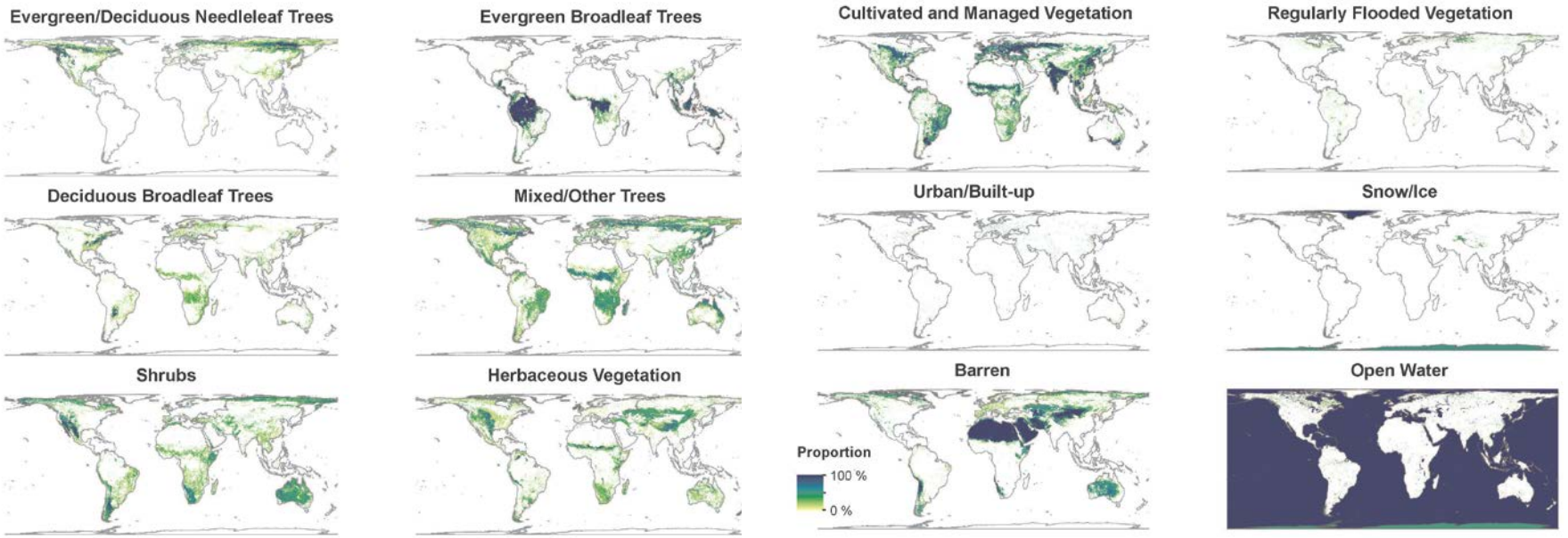
Product Integration



$$fwp_i = 0.5 \cdot [iwp_i + 0.5 \cdot (\sum_j p_{Dj} \cdot w_{Dij} + \sum_j p_{Lj} \cdot w_{Lij})]$$

Land Cover Proportions

- 12 data layers

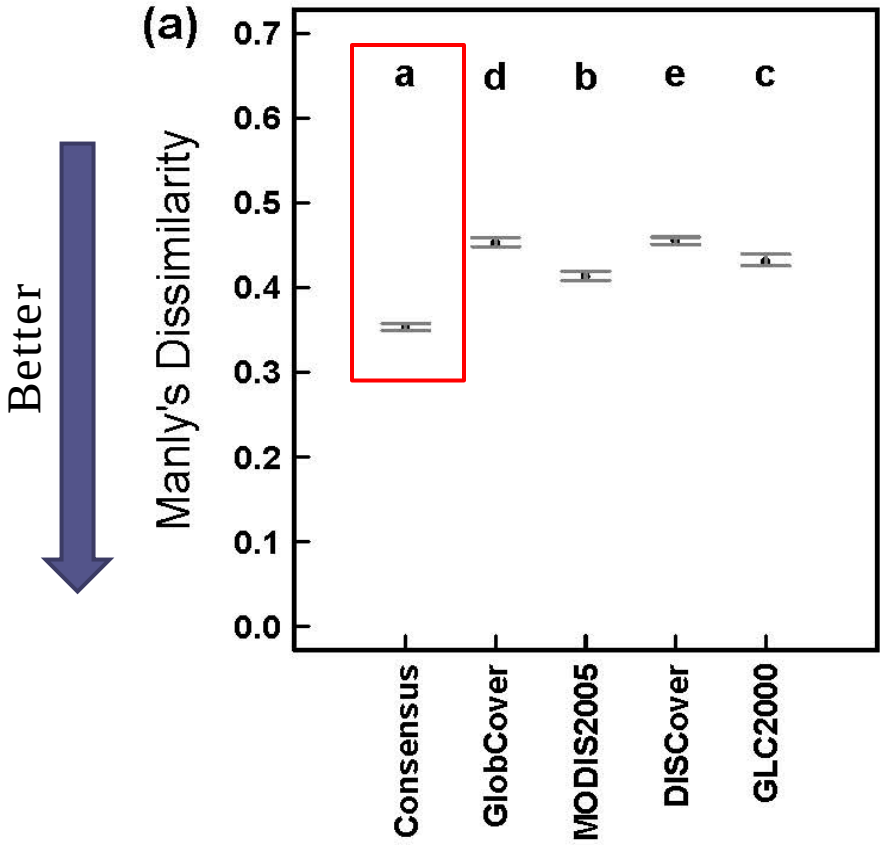


Ability to Capture Sub-pixel Information

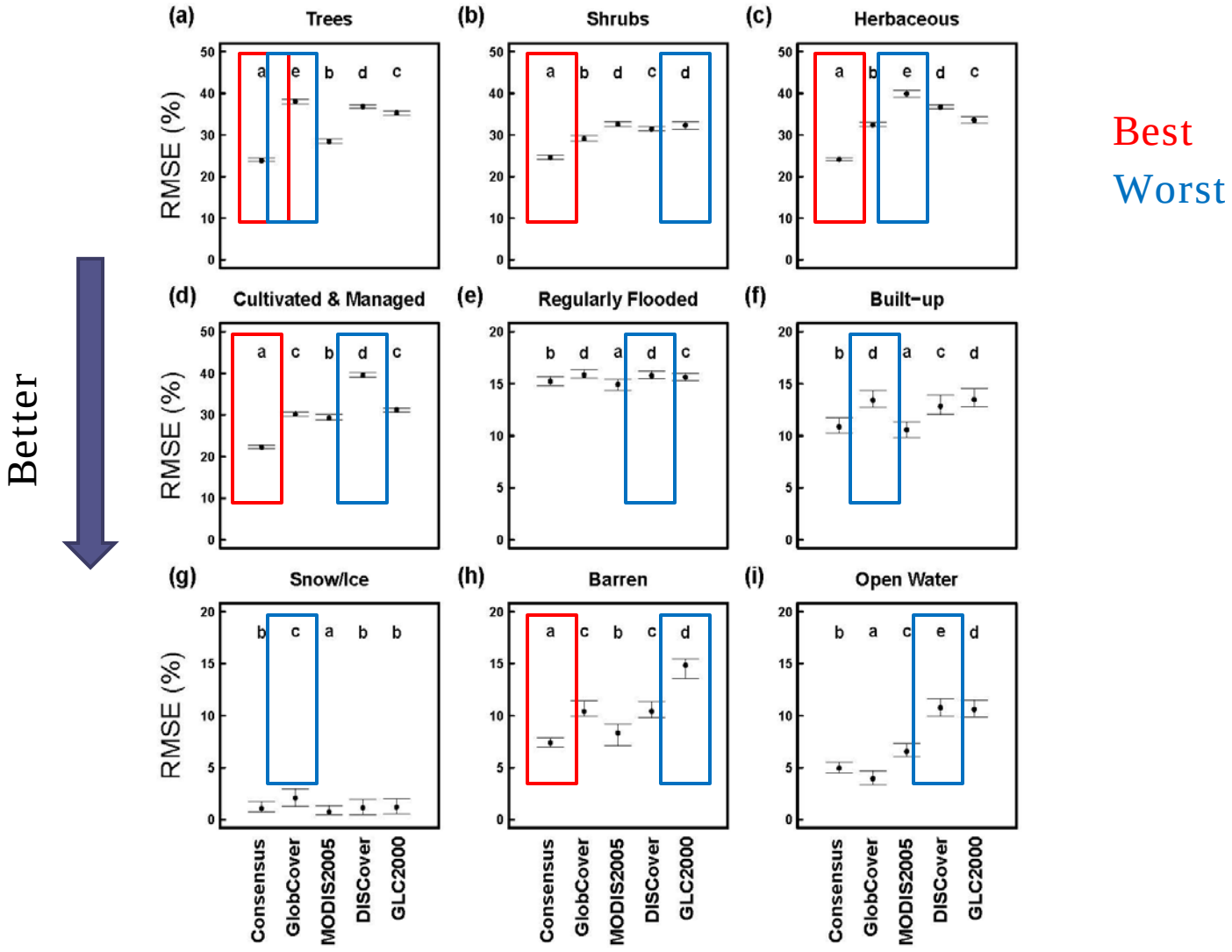
- 30-m NLCD2006 for the conterminous US (validation data)
- Calculating land cover prevalence at the 1-km resolution
- 9 land cover classes (Classes 1-4 were aggregated into a “trees” class)

Overall Dissimilarity

- Manly's dissimilarity
- $\frac{1}{2}\sum_i |E_i - V_i|$
- E_i and V_i are proportions of class i in the evaluated and validation datasets
- 20 sets of randomly selected 10,000 1-km pixels



Dissimilarity for Individual Classes



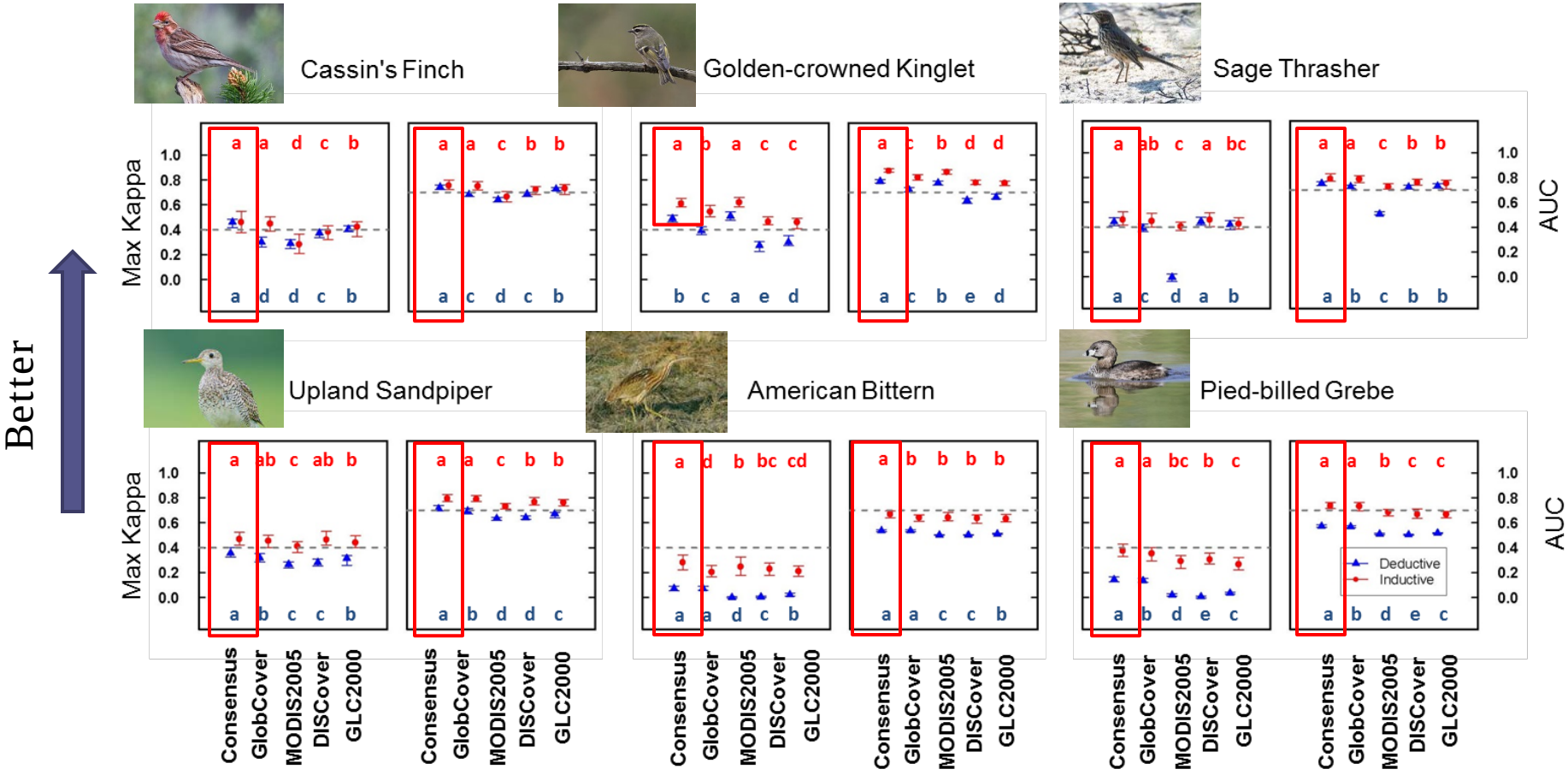
Best

Worst

Utility for Species Distribution Modeling

- Six avian habitat specialists
- Presence/absence data from North American Breeding Bird Survey (BBS)
- Deductive models
  - Occurrence probability = prevalence of all required land cover classes
- Inductive models
  - Logistic regression
  - Correlating species occurrence with the prevalence of all land cover classes

Model Accuracy



Summary

- The consensus dataset retained strengths and reduced weaknesses of individual input products
- The consensus dataset has the better ability to capture sub-pixel land cover information
- The consensus dataset improves the accuracy of both deductive and inductive models for predicting species distributions

Discussion

- Advantages
  - Reduction of errors and uncertainties
  - Proportional land cover information
- Low temporal resolution
  - Imagery from 1992 - 2006
  - Complementary information from old products
  - Matching temporal resolutions of biodiversity data
- Beyond the four products
- Beyond biodiversity research

Acknowledgement

- Tien Ming Lee
- Colleagues in Jetz lab
- Colleagues in the Environment and Organisms Working Group at NCEAS
- NASA, NSF, NCEAS