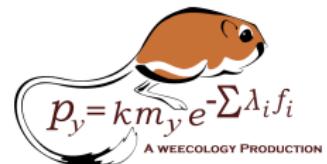


A DATA-INTENSIVE ASSESSMENT OF THE SPECIES-ABUNDANCE DISTRIBUTION.

Elita Baldridge



MACROECOLOGY

One approach to studying ecological patterns and processes.

- Data intensive.
- Large scales
 - Spatial
 - Temporal
 - Taxonomic
- Search for generality.

MACROECOLOGY

Criticisms of macroecology

- North American terrestrial bias.
- Lack of identification of pattern generating mechanisms.

MACROECOLOGY

Best practice recommendations

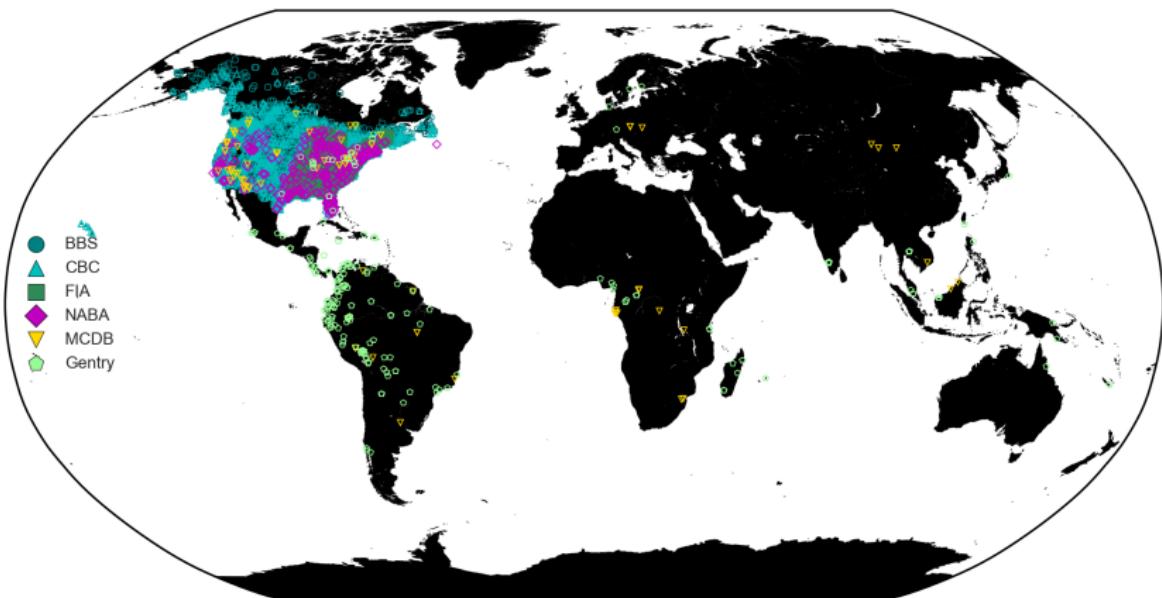
- Test patterns with multiple taxonomic groups/ecosystems.
- Simultaneous testing of competing models and model predictions with a consistent statistical approach.

THE RULES OF ECOINFORMATICS

Garbage in, garbage out.

- All data are good, not all data are appropriate.
- Fit the data to the question.

DATA



DATA

Major macroecological datasets

- Largely terrestrial
- Largely North American
- Many publicly available, some not.

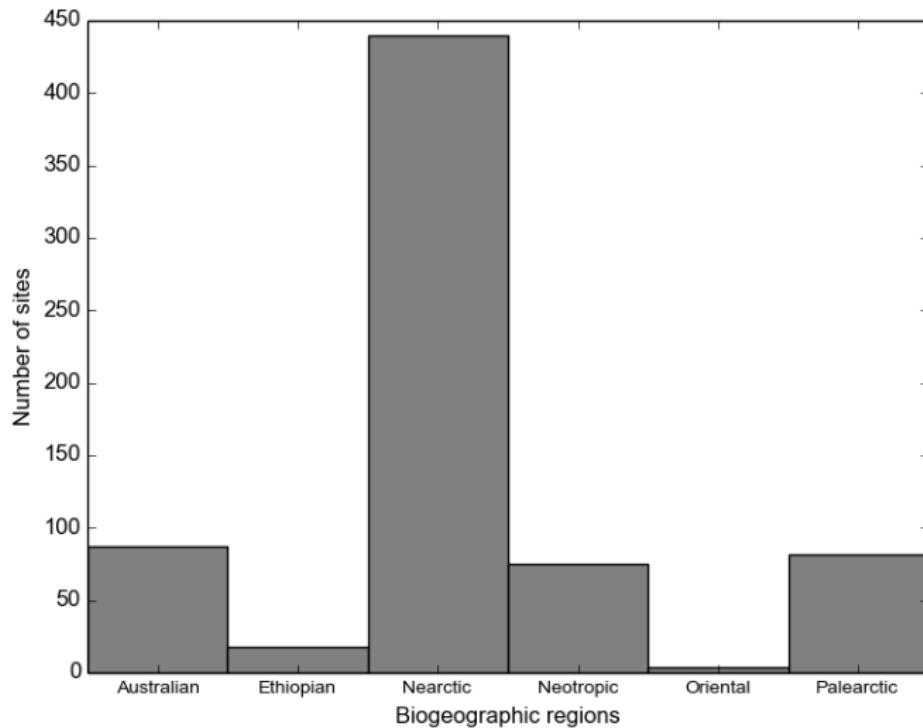
Lots of data in the literature.

DATA

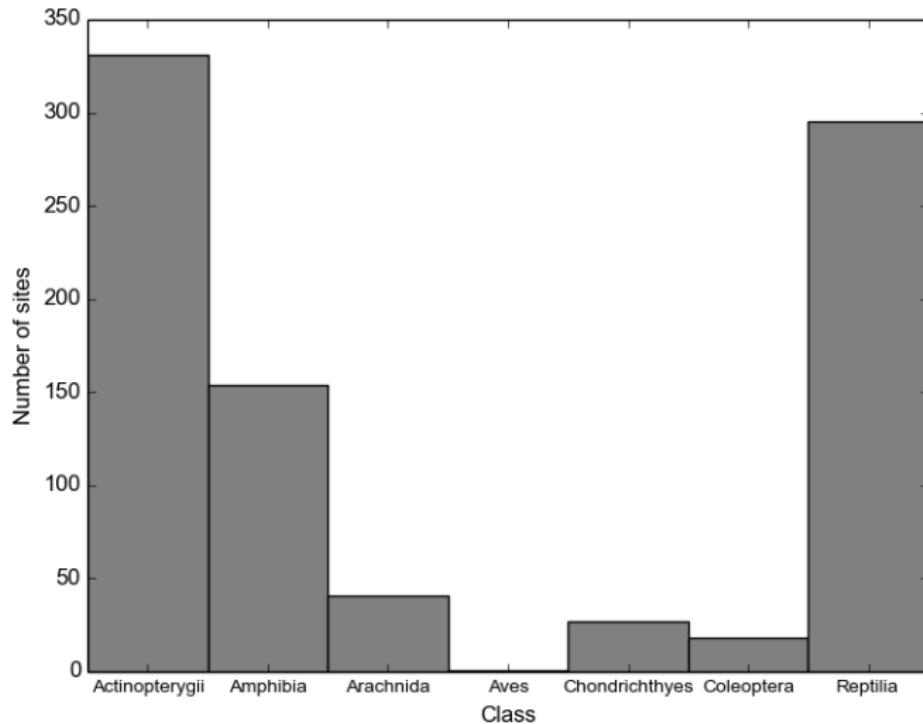
Variable name	Variable definitions
Class	Taxonomic class of species
Family	Taxonomic family of species
Genus	Taxonomic genus of species
Species	Specific epithet of species
Relative_abundance	Relative abundance of species
Abundance	Abundance of species
Collection_Year	Start of collecting
End_Collection	End of collecting
Site_Name	Name/description of site
Biogeographic_region	Biogeographic region
Site_notes	Additional site information

TABLE : List of variables collected.

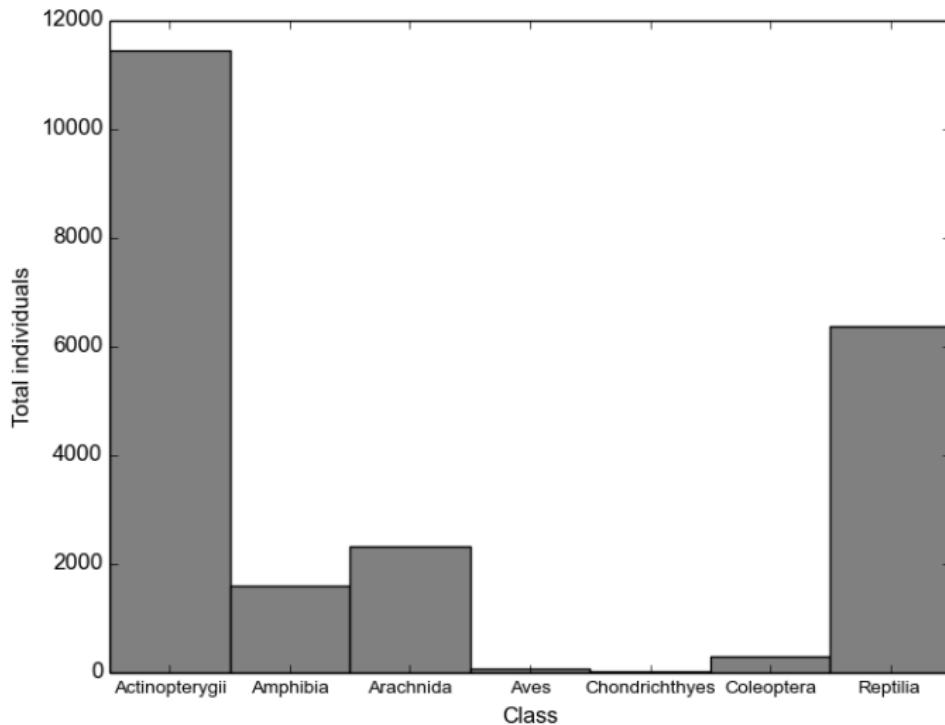
DATA



DATA



DATA



DATA

The final community abundance database
is publicly available and importable
through the EcoData Retriever.
(<http://www.ecodataretriever.org/>)

COMMONNESS & RARITY

The species abundance distribution:

- Describes the distribution of commonness & rarity of species.
- One of the most fundamental and ubiquitous patterns in ecology.
- Exhibits a hollow curve distribution.
 - Many rare species.
 - Few common species.
- Many forms of the species abundance distribution (SAD).

FORMS OF THE SAD

Model classes:

- Purely statistical
- Branching process
- Population dynamics
- Niche partitioning
- Maximum entropy
- Feasible set/combinatorics

SAD COMPARISONS

Most comparisons of the different models:

- Use only a small subset of available models (typically two).
- Focus on a single ecosystem or taxonomic group
- Fail to use the most appropriate statistical methods.

SAD COMPARISONS

Selected five models from four classes for comparison.

Model class	Form of the distribution
Purely statistical	Logseries, Poisson lognormal
Branching process	Zipf
Population dynamics	Negative binomial
Niche partitioning	Geometric

TABLE : After B.J. McGill et al. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology letters 10: 995-1015.

SAD COMPARISONS

Analysis:

- Model fitting with maximum likelihood estimation.
- Likelihood based model selection to compare the fits of the different models.
- Model comparison with corrected Aikaike Information Criterion (AICc) weights.

SAD COMPARISONS

Computational tools:

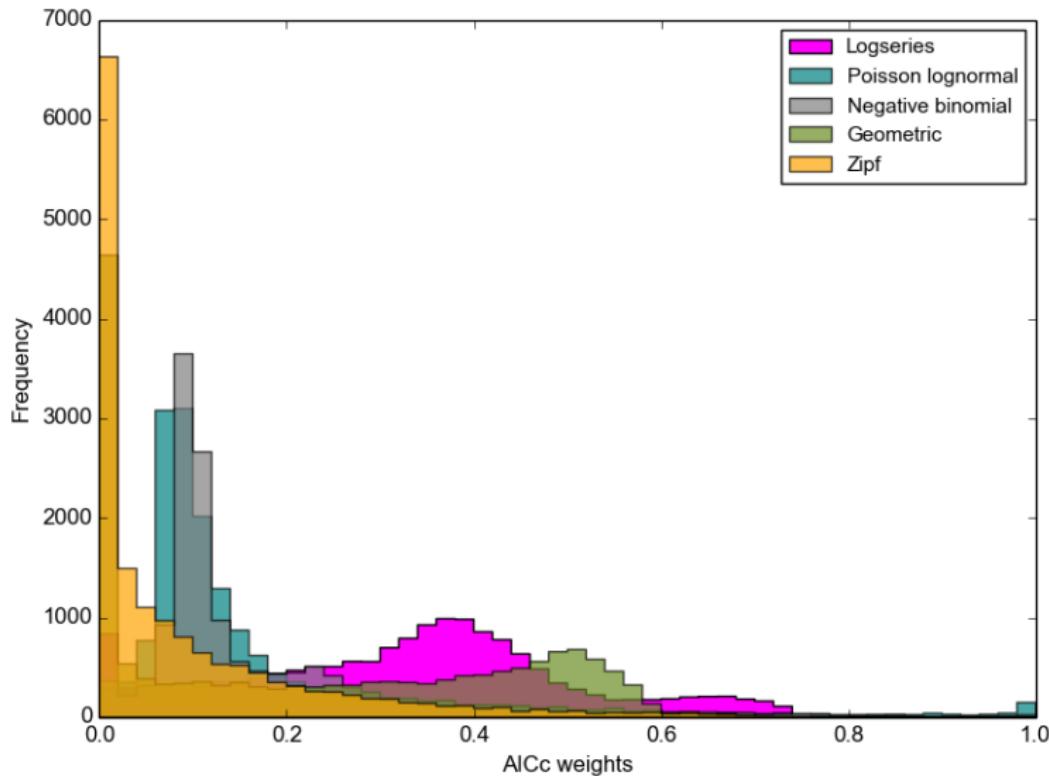
- Model fitting, log-likelihood, and AICc calculations performed with macroecotools Python package.
(<https://github.com/weecology/macroecotools>)
- All of the analysis code and the majority of the data is publicly available.
(<https://github.com/weecology/sad-comparison>)

SAD COMPARISONS

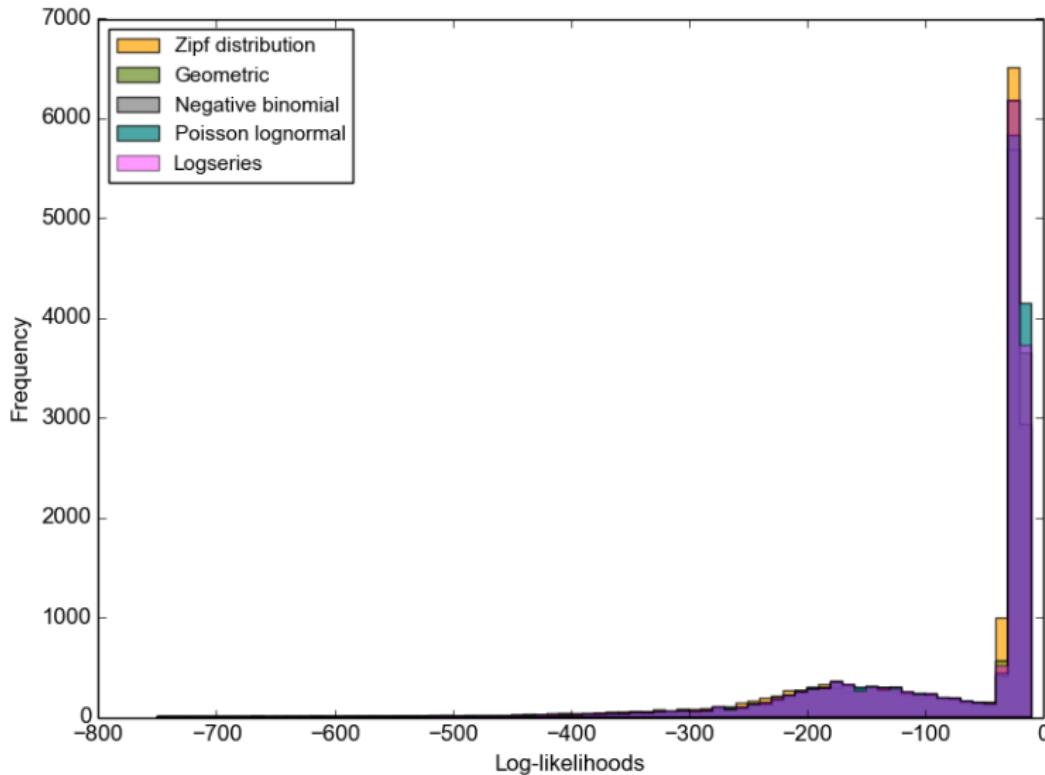
Dataset	Dataset code	Availability	Total sites
Breeding Bird Survey	BBS	Public	2769
Christmas Bird Count	CBC	Private	1999
Gentry's Forest Transects	Gentry	Public	10355
Forest Inventory Analysis	FIA	Public	220
Mammal Community Database	MCDB	Public	103
N. American Butterfly Count	NABA	Private	400
Actinopterygii, this dissertation	Actinopterygii	Public	161
Reptilia, this dissertation	Reptilia	Public	138
Amphibia, this dissertation	Amphibia	Public	43
Coleoptera, this dissertation	Coleoptera	Public	5
Arachnida, this dissertation	Arachnida	Public	25

TABLE : Datasets used for species-abundance distribution comparisons.
Datasets marked as Private were obtained through data requests to the providers
resulting in Memorandums of Understanding governing data use.

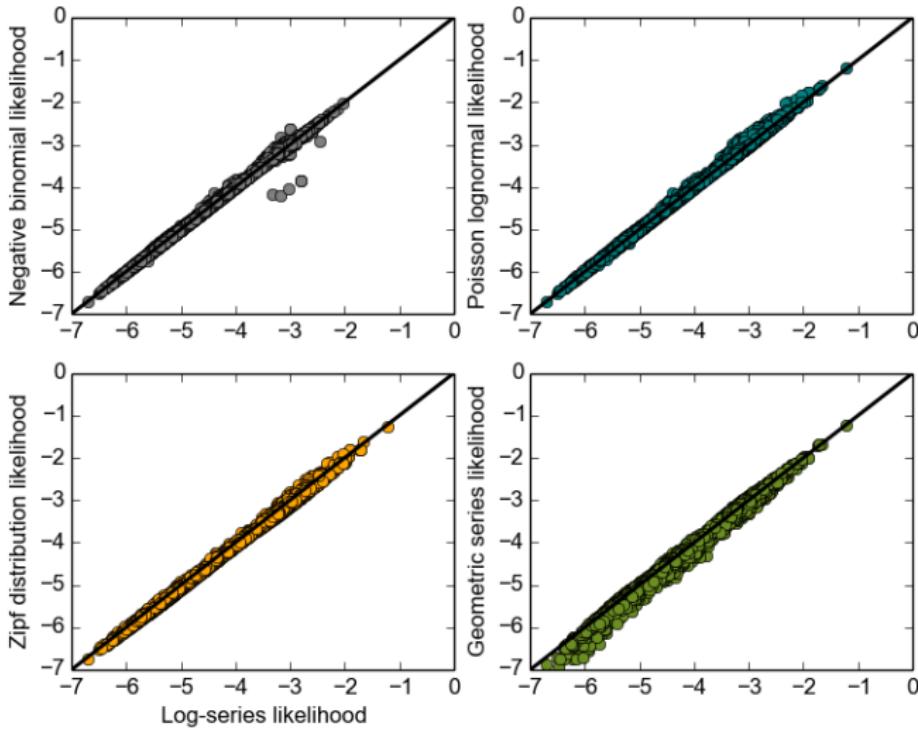
SAD COMPARISONS



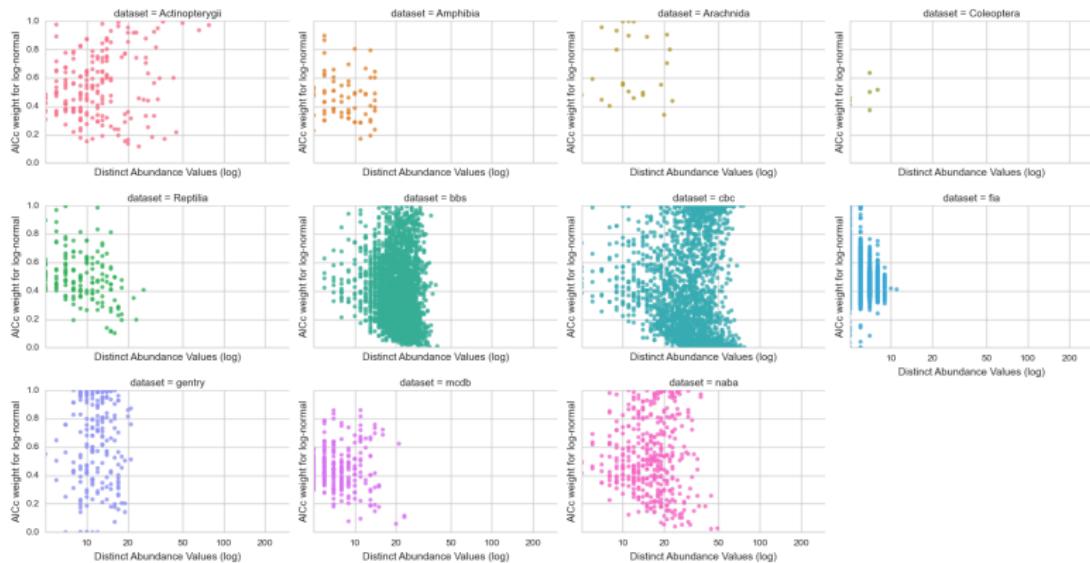
SAD COMPARISONS



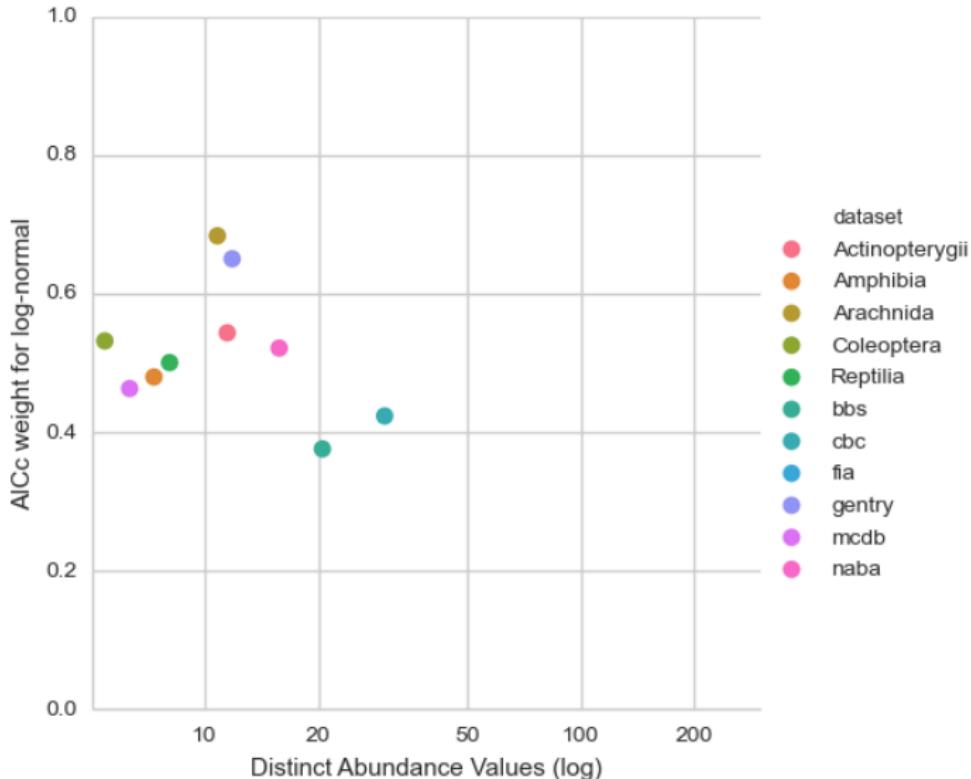
SAD COMPARISONS



NEUTRAL ANALYSIS



NEUTRAL ANALYSIS



ACKNOWLEDGEMENTS

Funding sources:

- USU Department of Biology
- Intellectual Ventures private funding to Morgan Ernest
- National Science Foundation CAREER Grant to Ethan White
- Gordon & Betty Moore Foundation's Data-Driven Discovery Initiative Grant to Ethan White.
- USU Graduate School Dissertation Fellowship

ACKNOWLEDGEMENTS

Weecologists past, present, & future



(especially Xiao Xiao & Ken Locey (creator of the whiteboard))

ACKNOWLEDGEMENTS

Dr. Thomas Price & USU Student Health Center