Isotopic Assessment of Animal Origin

Gabe Bowen
gabe.bowen@utah.edu

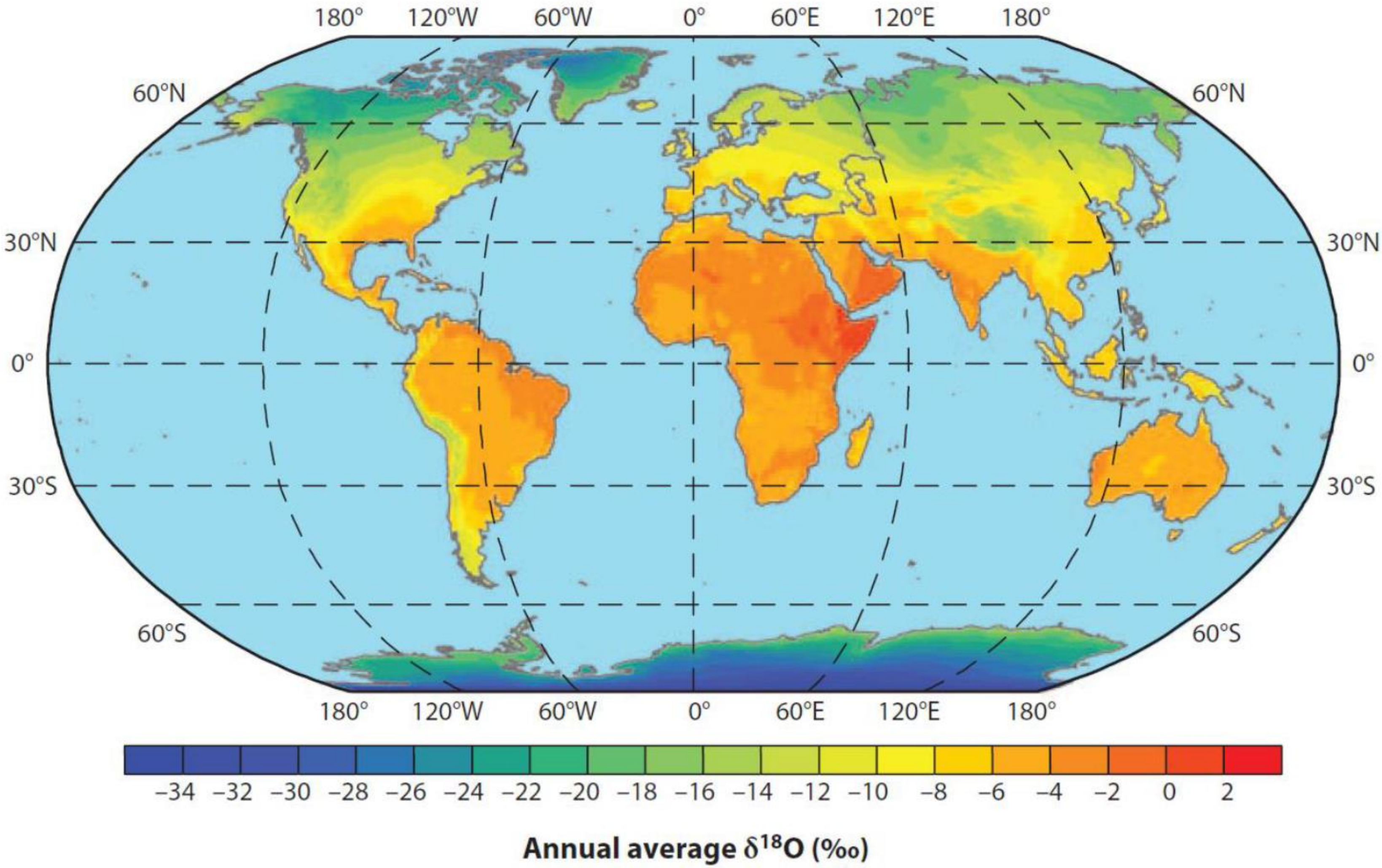
SPATIAL
University of Utah

THE UNIVERSITY OF UTAH

Migratory Origin Questions

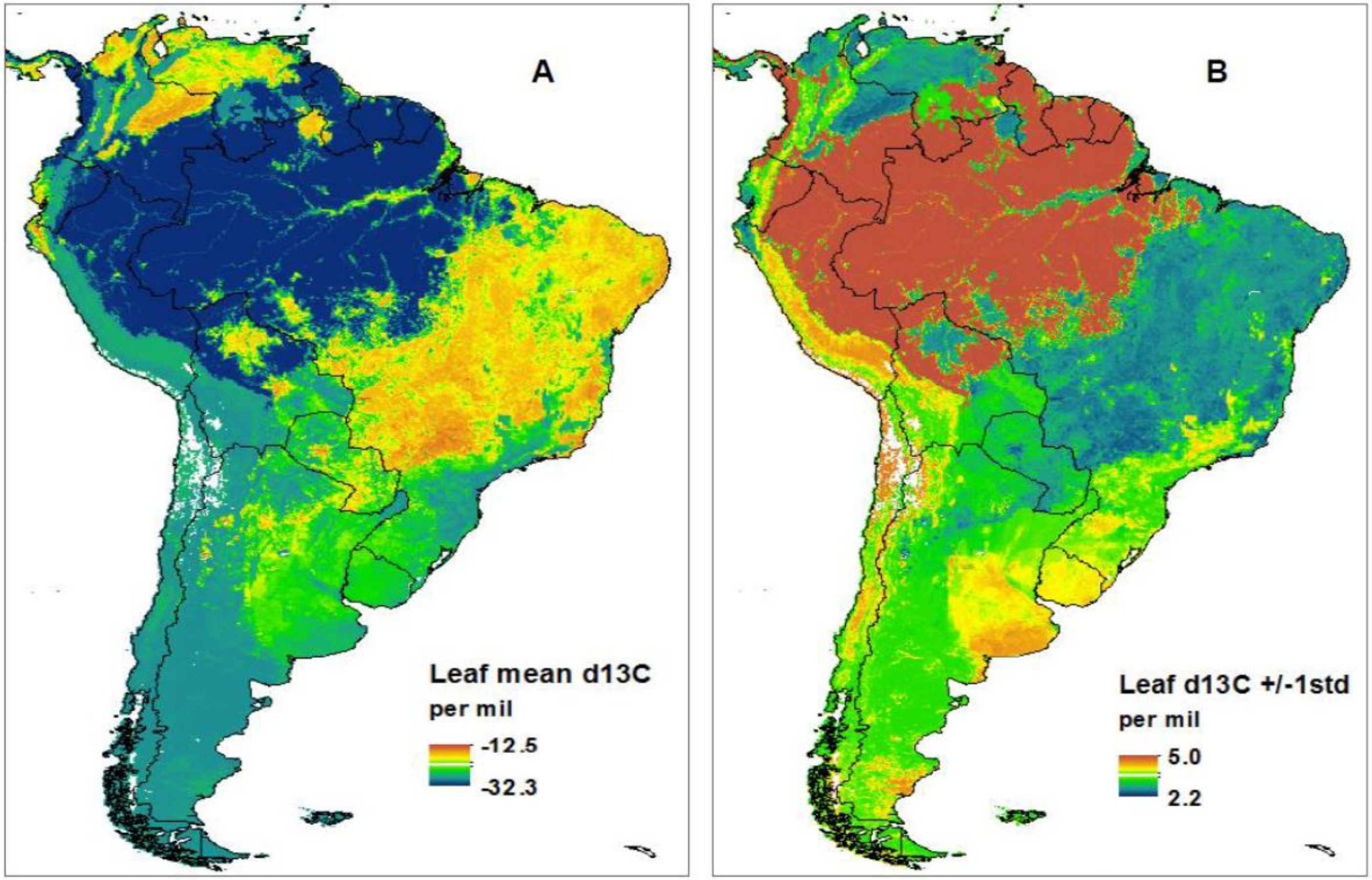
- Where did this individual breed?
- Where did this individual NOT breed?
- Which country did this individual come from?
- What is the most likely country of origin?
- Is this cheese from Parma?
- What is the pattern of pattern of connectivity in this population?

Useful Isotope Systems...



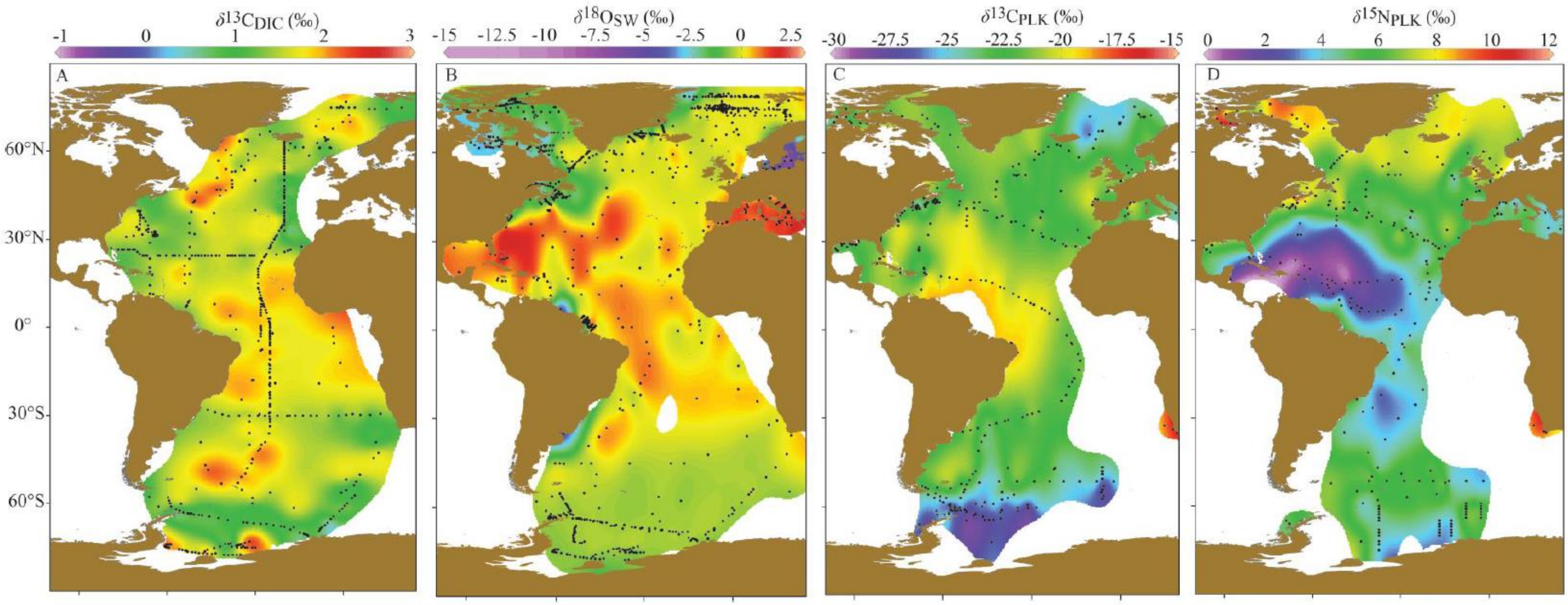
Useful Isotope Systems…

South American canopy leaf $\delta^{13}C$ predictions



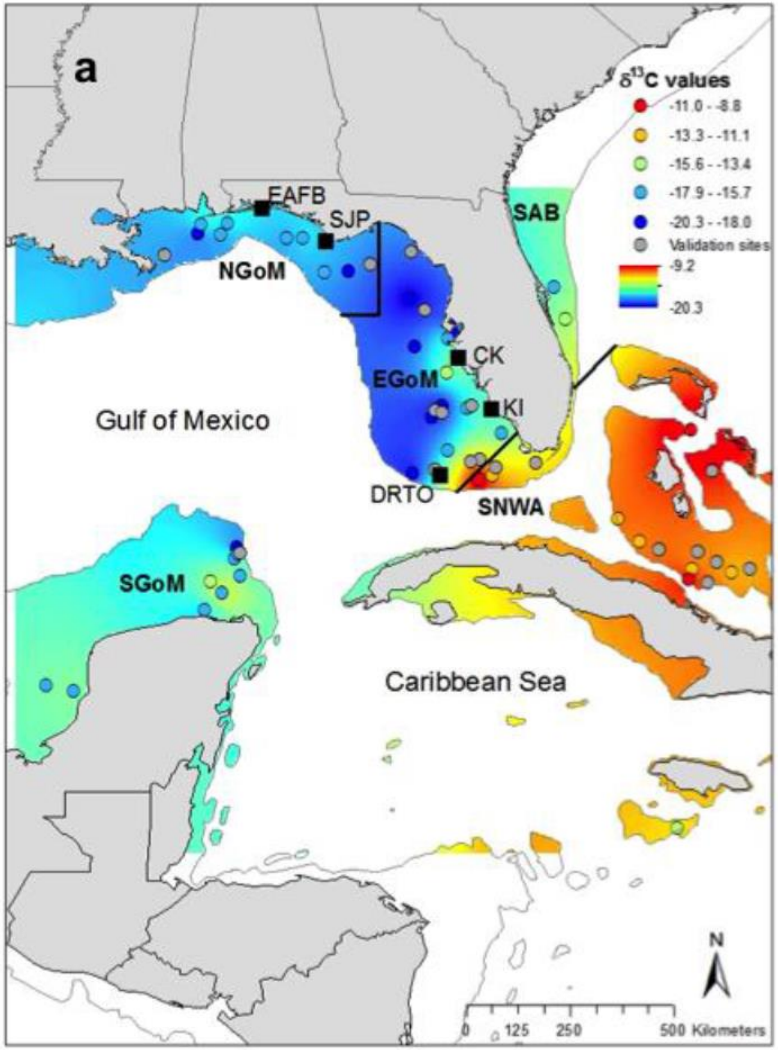
Powell, Yoo, and Still, "Vegetation and Soil Carbon-13 Isoscapes for South America: Integrating Remote Sensing and Ecosystem Isotope Measurements" *Ecosphere* (2012)

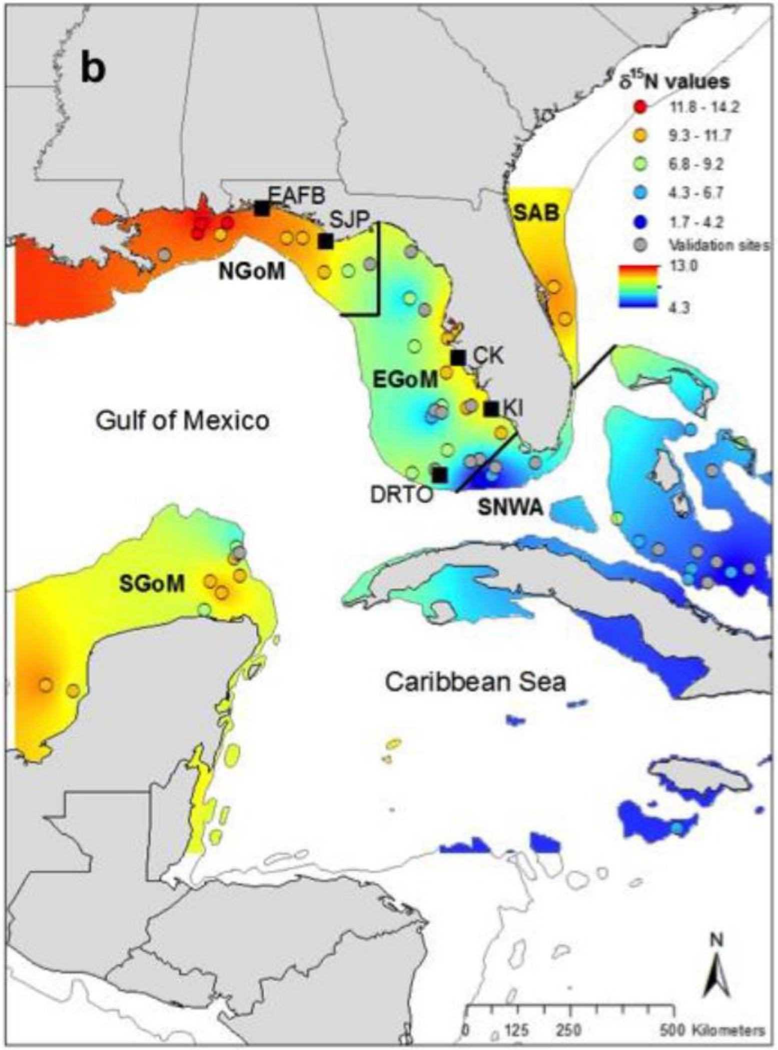
Useful Isotope Systems…



McMahon et al., 2013

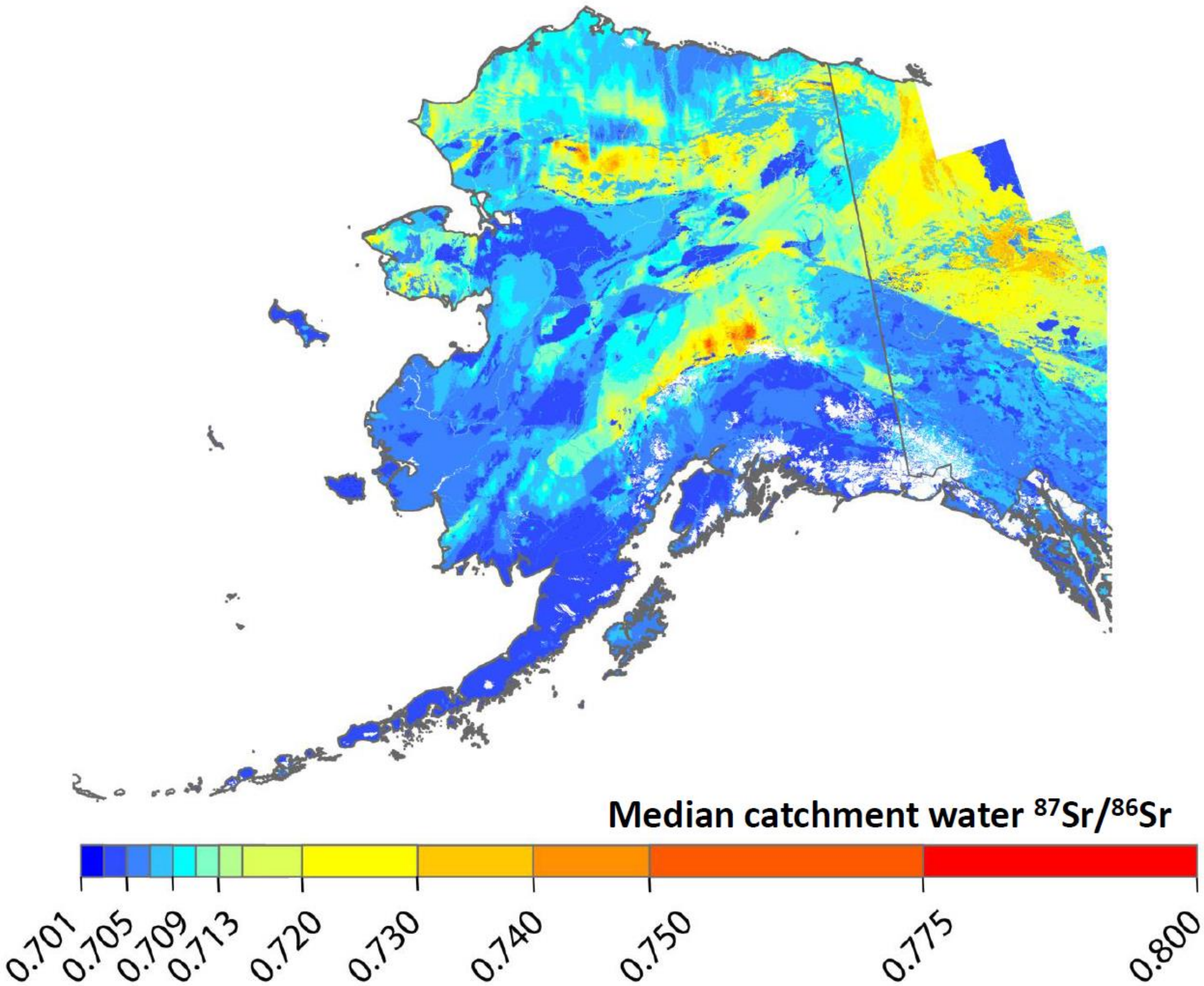
Useful Isotope Systems...





Vander Zanden et al., 2015

Useful Isotope Systems…



Bataille et al., 2014

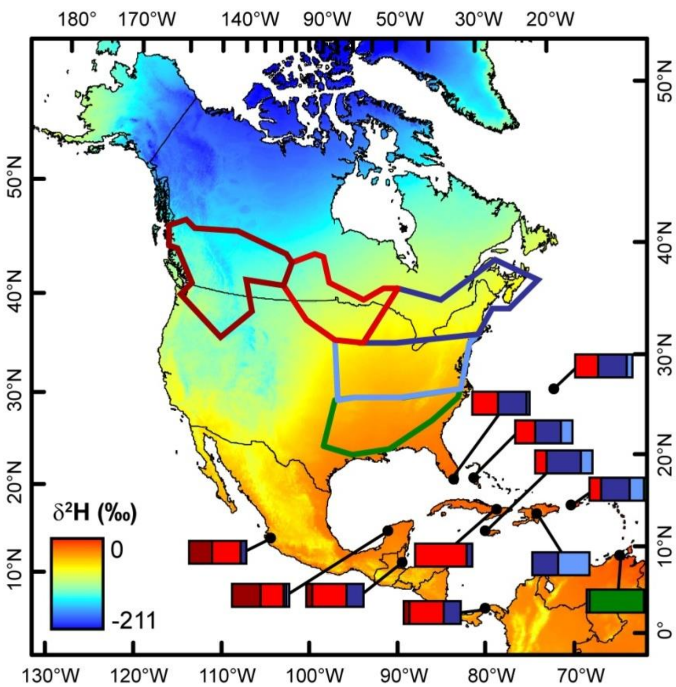
Bayesian Inference of Origin

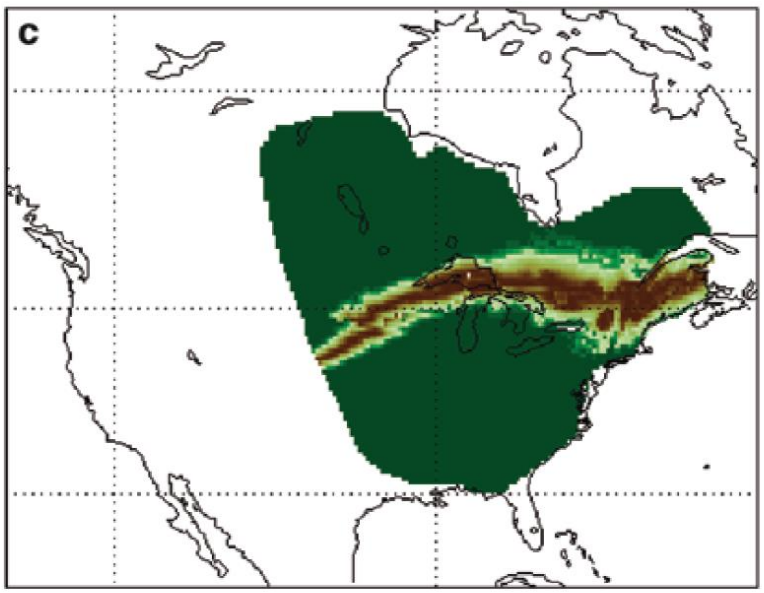
$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum P(B|A_j)P(A_j)}$$

- The posterior probability of model *i* being the true model given some observations is a function of
  - The conditional probability of the observations given model *i*
  - The prior probability of model *i*
  - The probabilities associated with all other hypotheses

Defining Hypotheses

- Discrete (nominal)
- Continuous





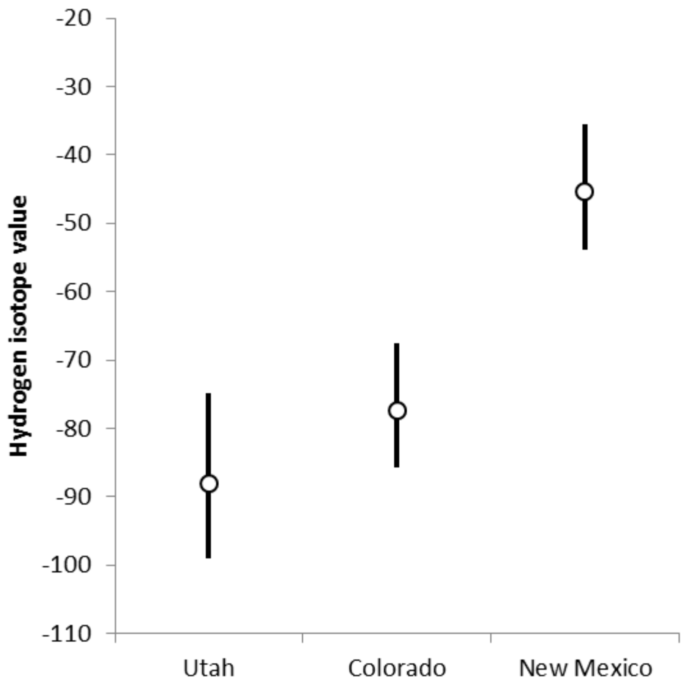
Norris et al., 2006; Wunder, 2010

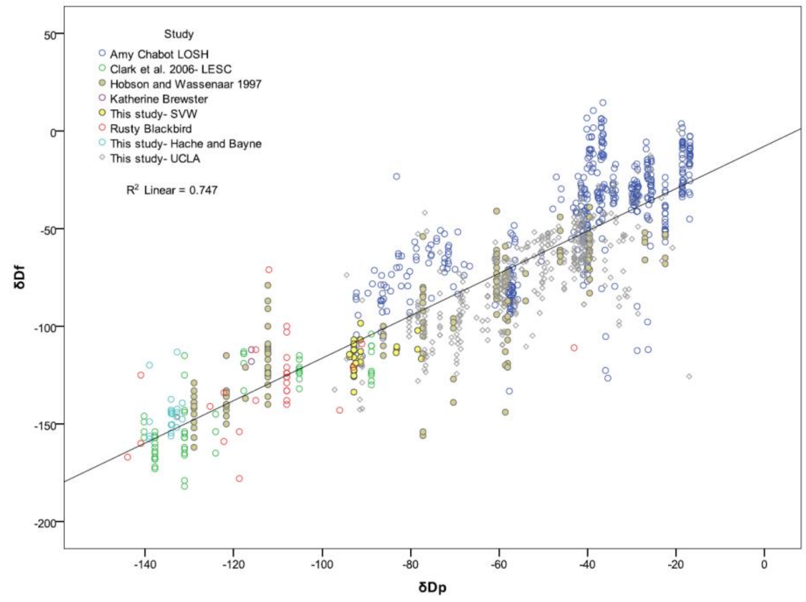
Defining Hypotheses

- Discrete (nominal)
  - Pose hypotheses in terms of discrete regions from outset
  - + analysis unit can reflect question (management unit, political boundaries)
  - + unit structures sampling needed to evaluate conditional probabilities
  - - units are sometimes ecologically unrealistic
  - - inflexible (granularity not suited to re-analysis)
- Continuous
  - Pose hypotheses in terms of large number of arbitrary (evenly distributed) locations
  - + preserves maximum information content
  - + conducive to reanalysis
  - - requires post-analysis summarization to answer ecological and management questions
  - - requires model-based evaluation of conditional probabilities

Evaluating Conditional Probabilities

- Sample-based
- Model-based



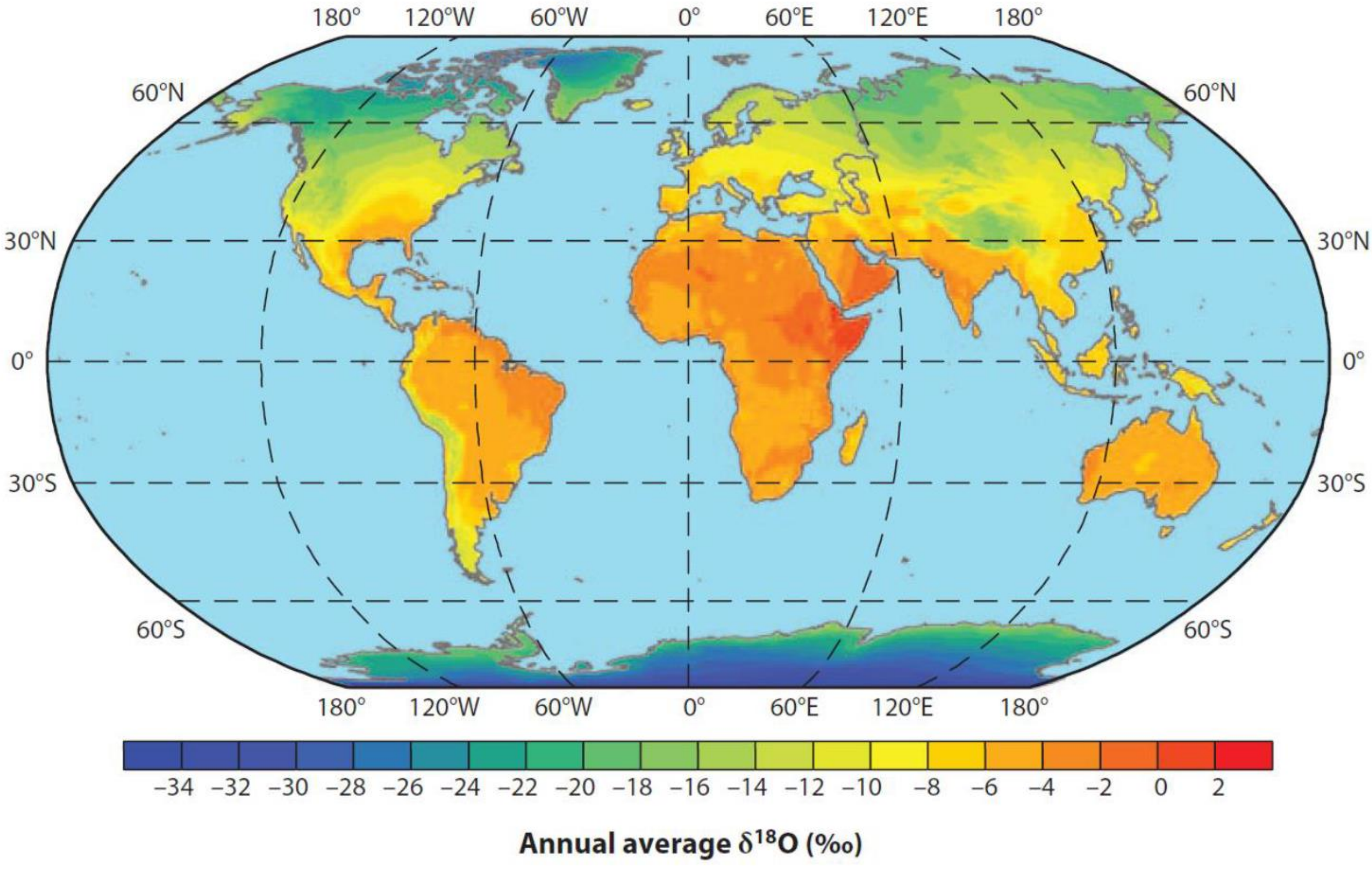


Hobson et al., 2012

Evaluating Conditional Probabilities

- Sample-based
  - Sample known-origin individuals to characterize the distribution of values
  - + simple estimation of distribution
  - - labor-intensive and expensive
  - - prohibitive for continuous analysis
- Model-based
  - Use existing data + information about system to estimate distribution
  - + cheap
  - + amenable to continuous analysis
  - - requires model (for full distribution)
  - - estimating uncertainty can be challenging and complex

Model-based Starting Point



From Environment to Tissue

You are what you eat…

From Environment to Tissue

You are what you eat…

+ 3‰

From Environment to Tissue

You are what you eat…

~~+ 3‰~~

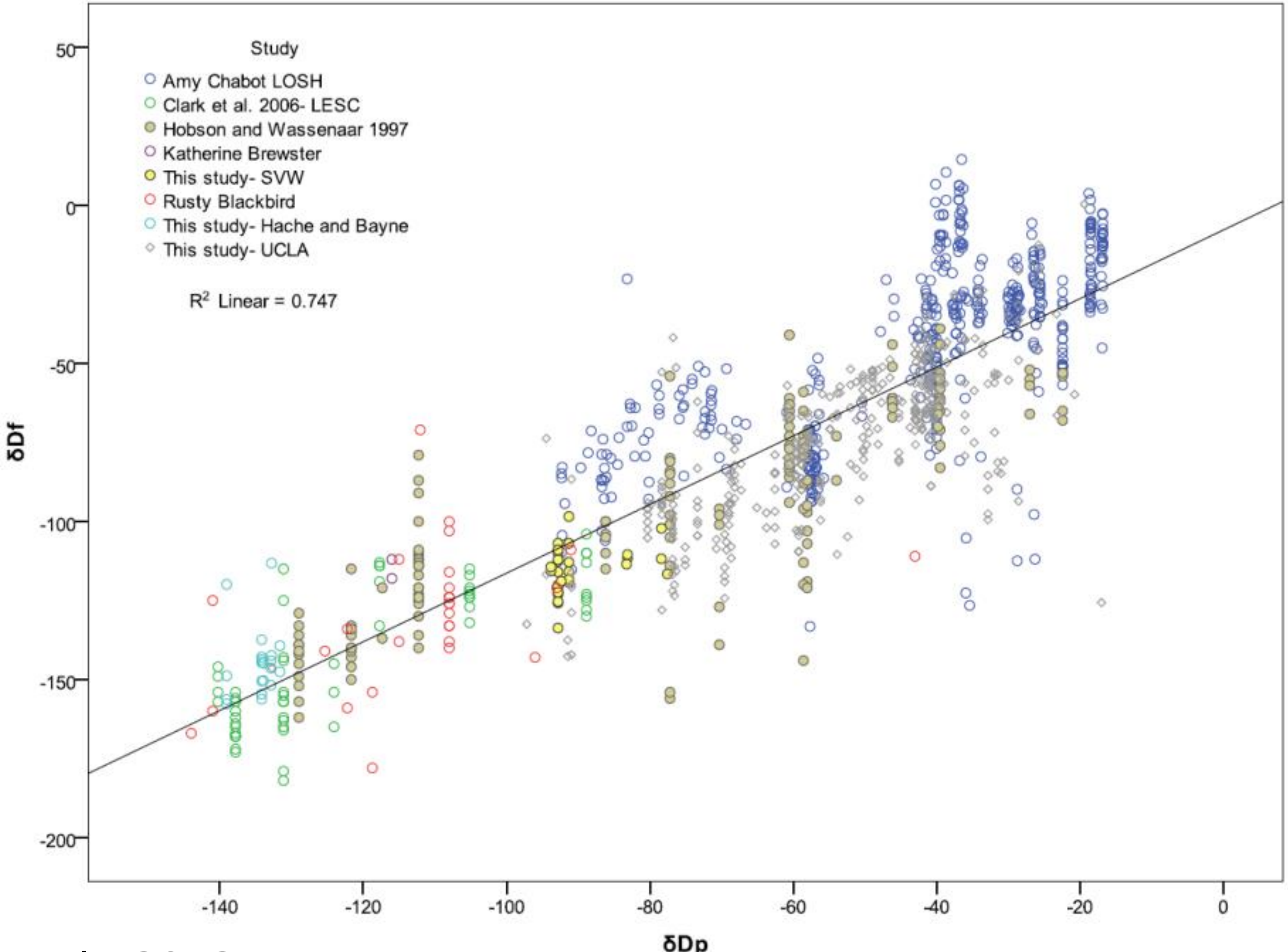
+ what you drink

+ what you breathe

- any fractionating losses

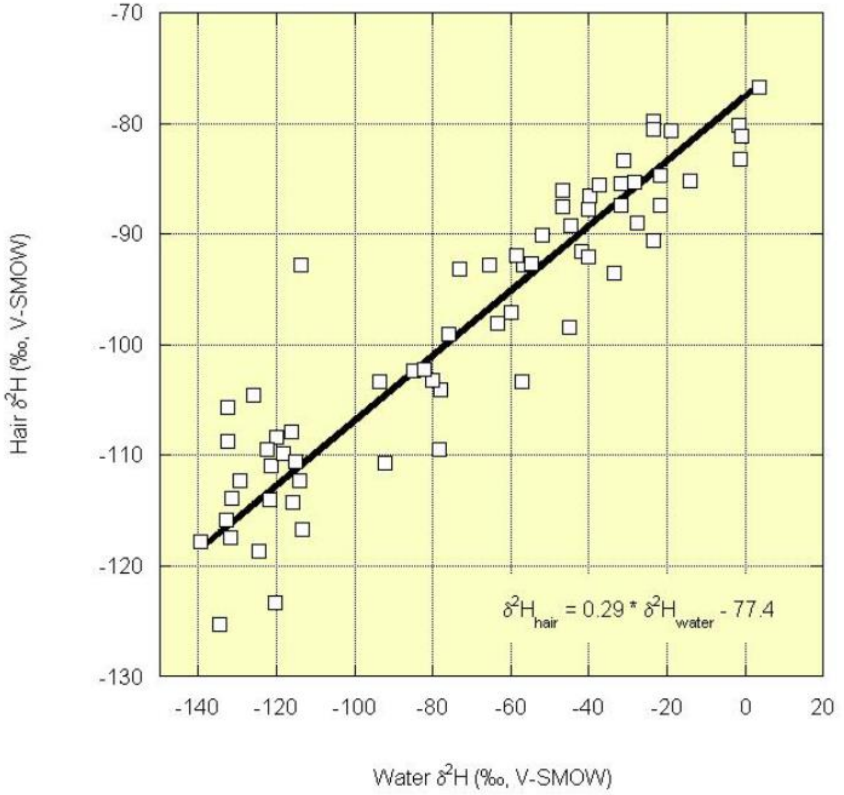
+/- any fractionation associated with tissue synthesis

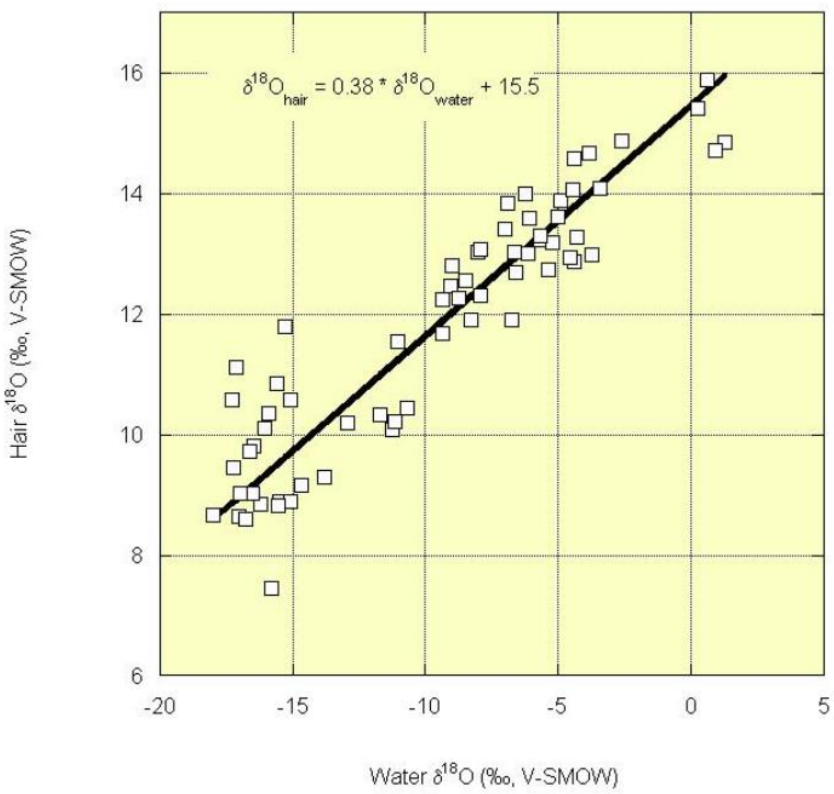
Empirical Calibration



Hobson et al., 2012

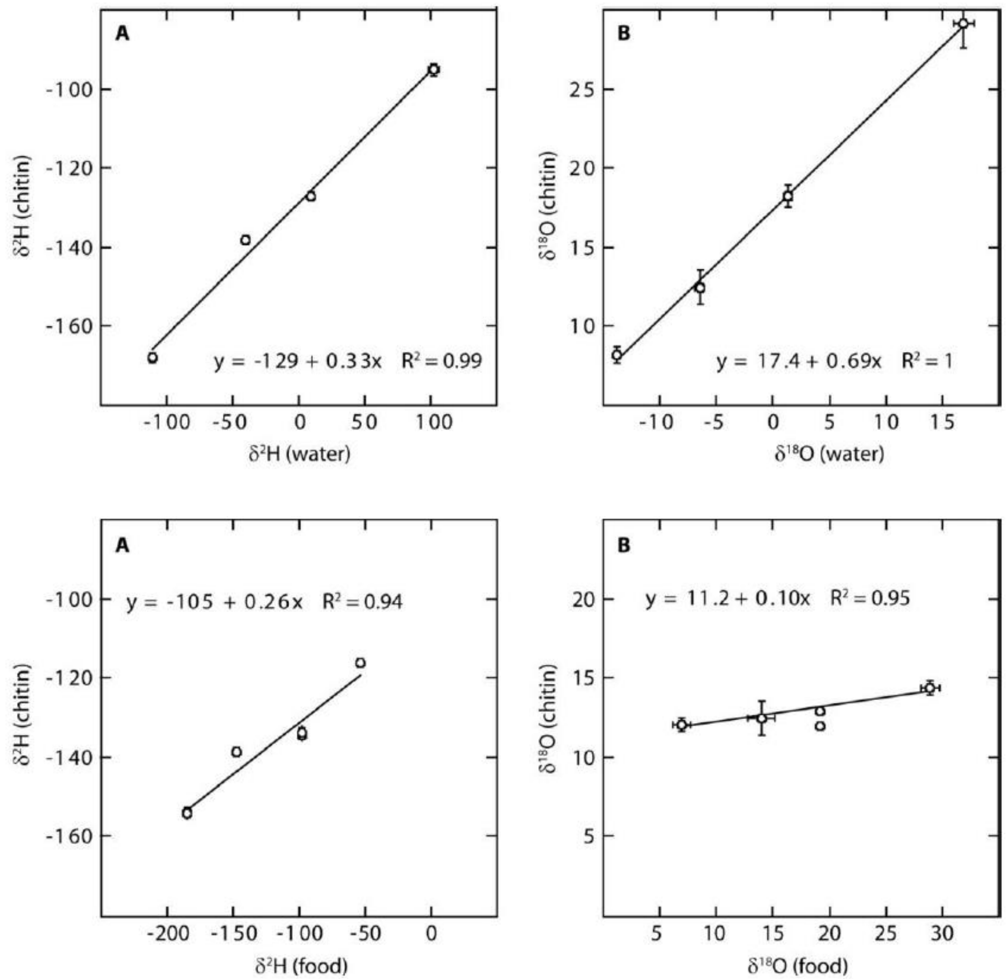
Empirical Calibration





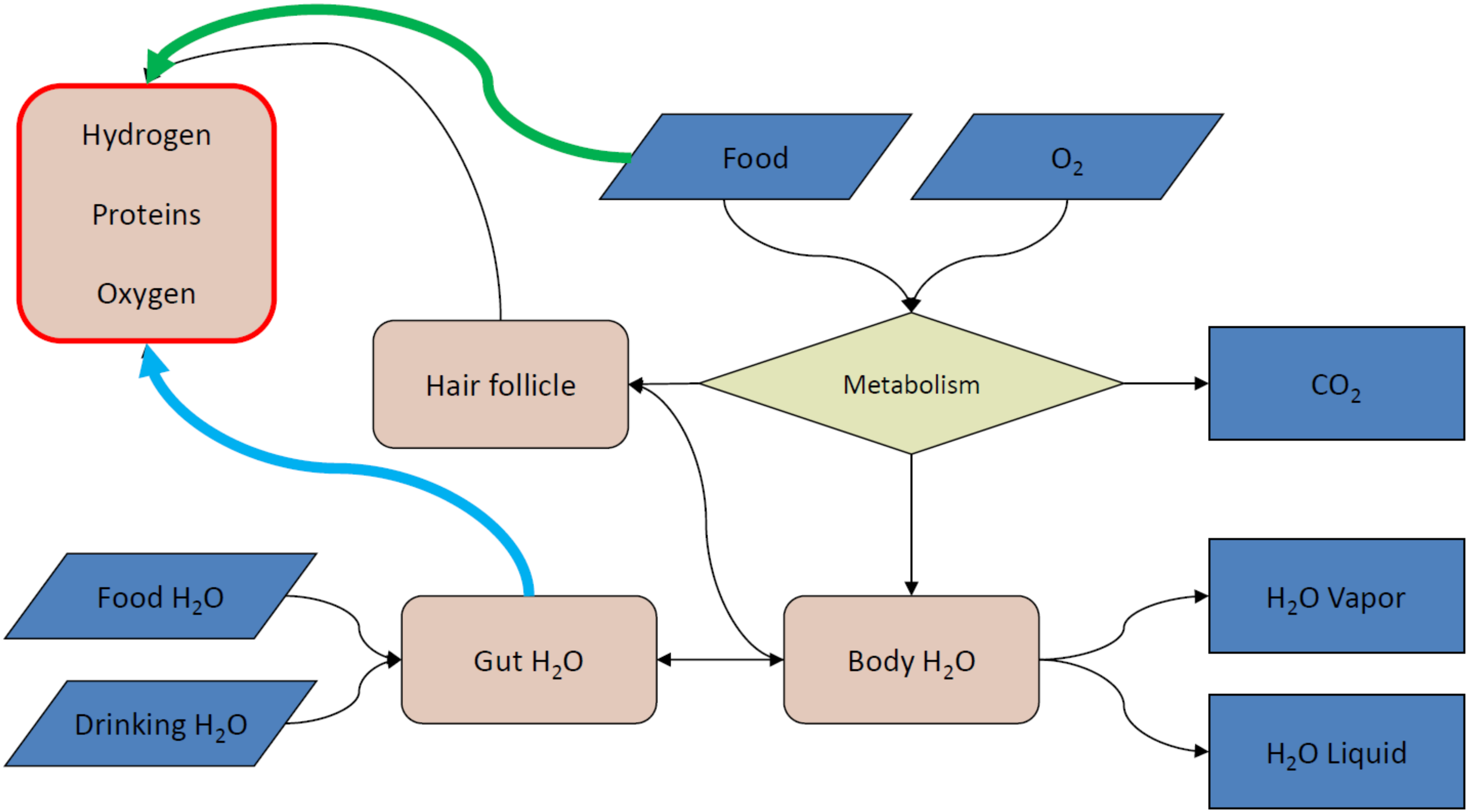
Ehleringer et al., 2008

Experimental Calibration



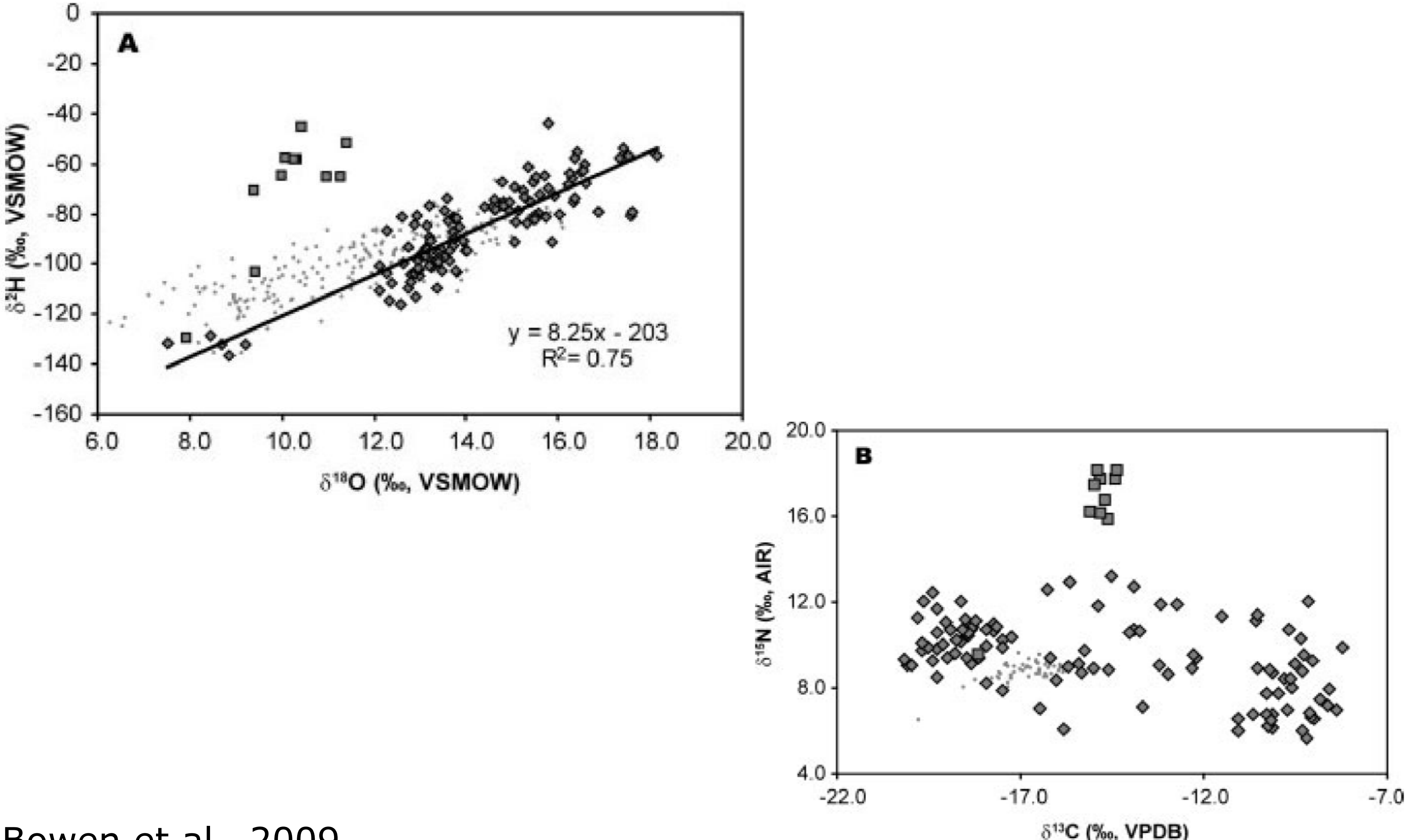
Nielson and Bowen, 2010

Theoretical Calibration Model



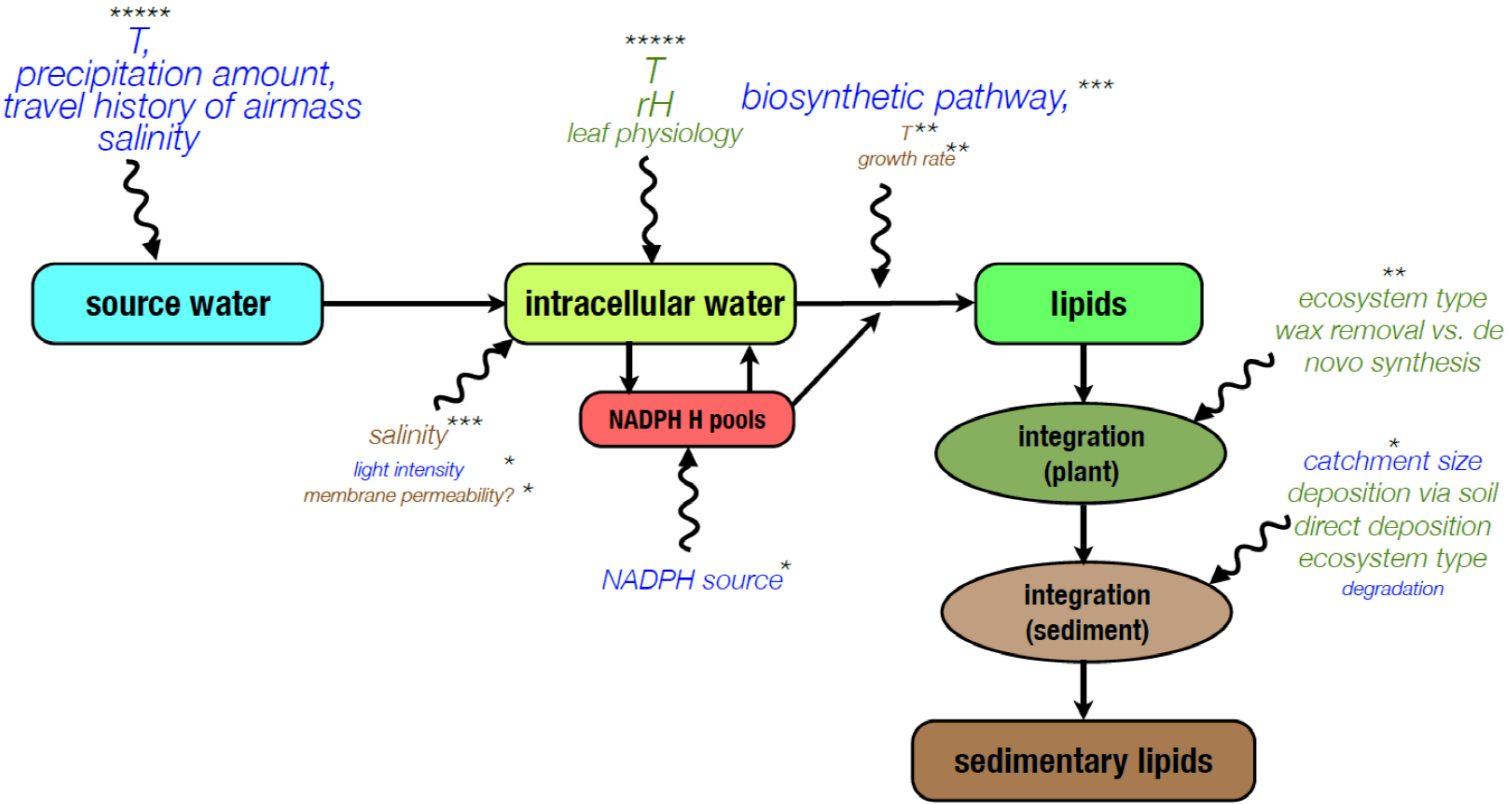
Ehleringer et al., 2008, Bowen et al., 2009

Theoretical Calibration Model



Bowen et al., 2009

Theoretical Calibration Model



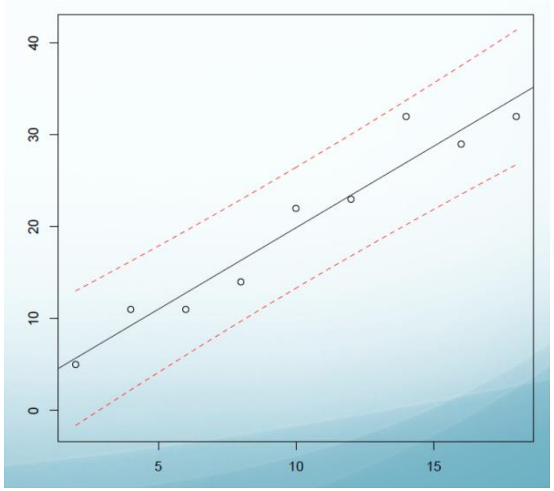
Sachse et al., 2012

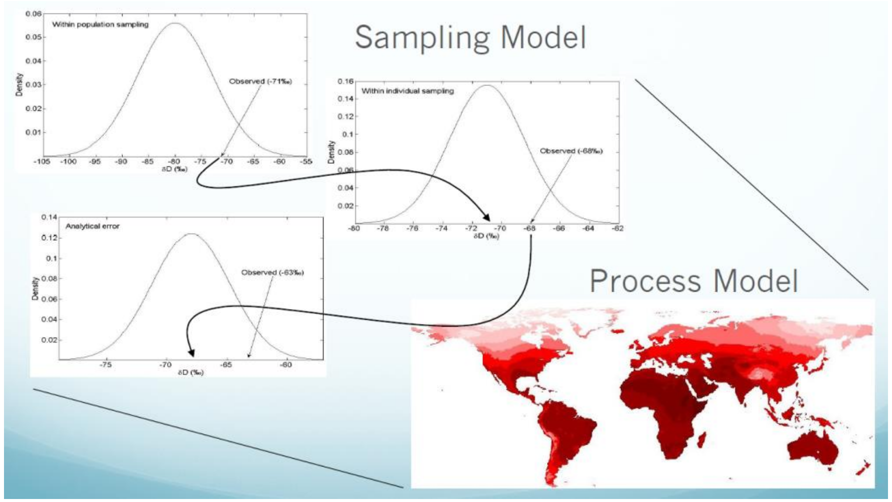
Evaluating Conditional Probabilities: Model Estimates of PDFs

- Estimation of mean values expected for each hypothesis is not enough, we must describe the complete probability density function
- Sample-based approach allows straight-forward characterization of distribution of expected values associated with a given hypothesis
- Estimation of distribution is more challenging with model-based approach
  - Challenge grows with complexity of model
  - Opportunity to learn grows with complexity of model

Model Estimates of PDFs

- Aggregate estimation
- Hierarchical estimation

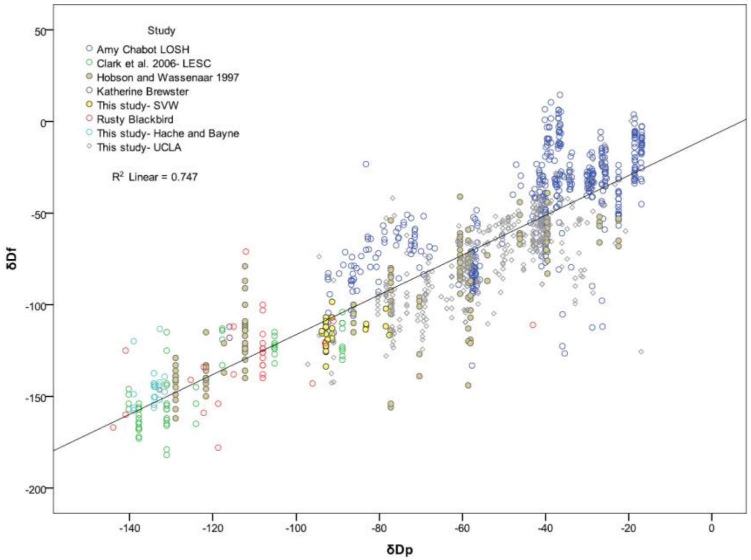


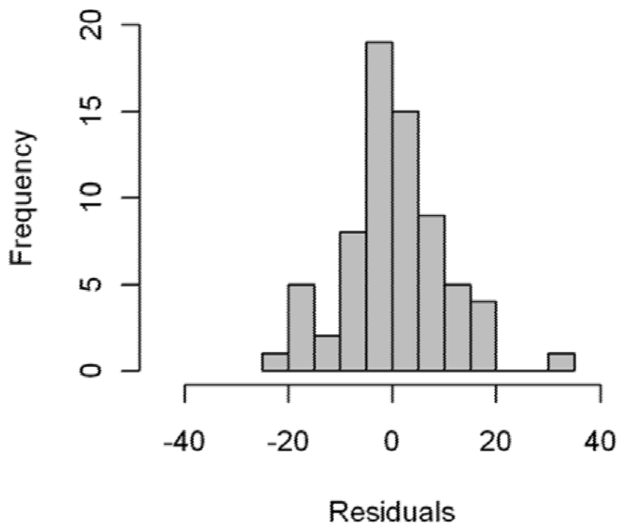


*In both cases we are often assuming parametric distributions for simplicity

Aggregate Estimation

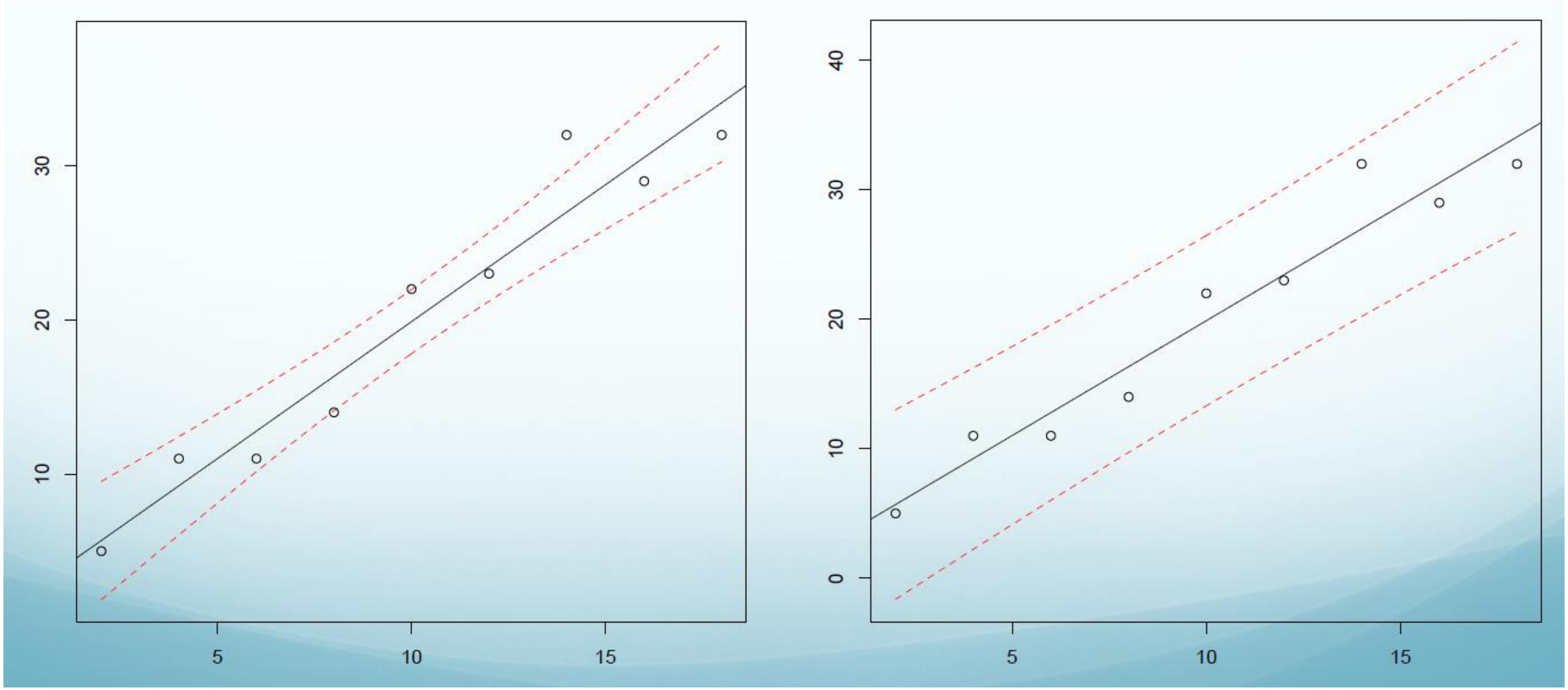
- Use field data to evaluate the variability in tissue measurements associated with repeat sampling
- Method 1: Use distribution of residuals from tissue isotope calibration relationship





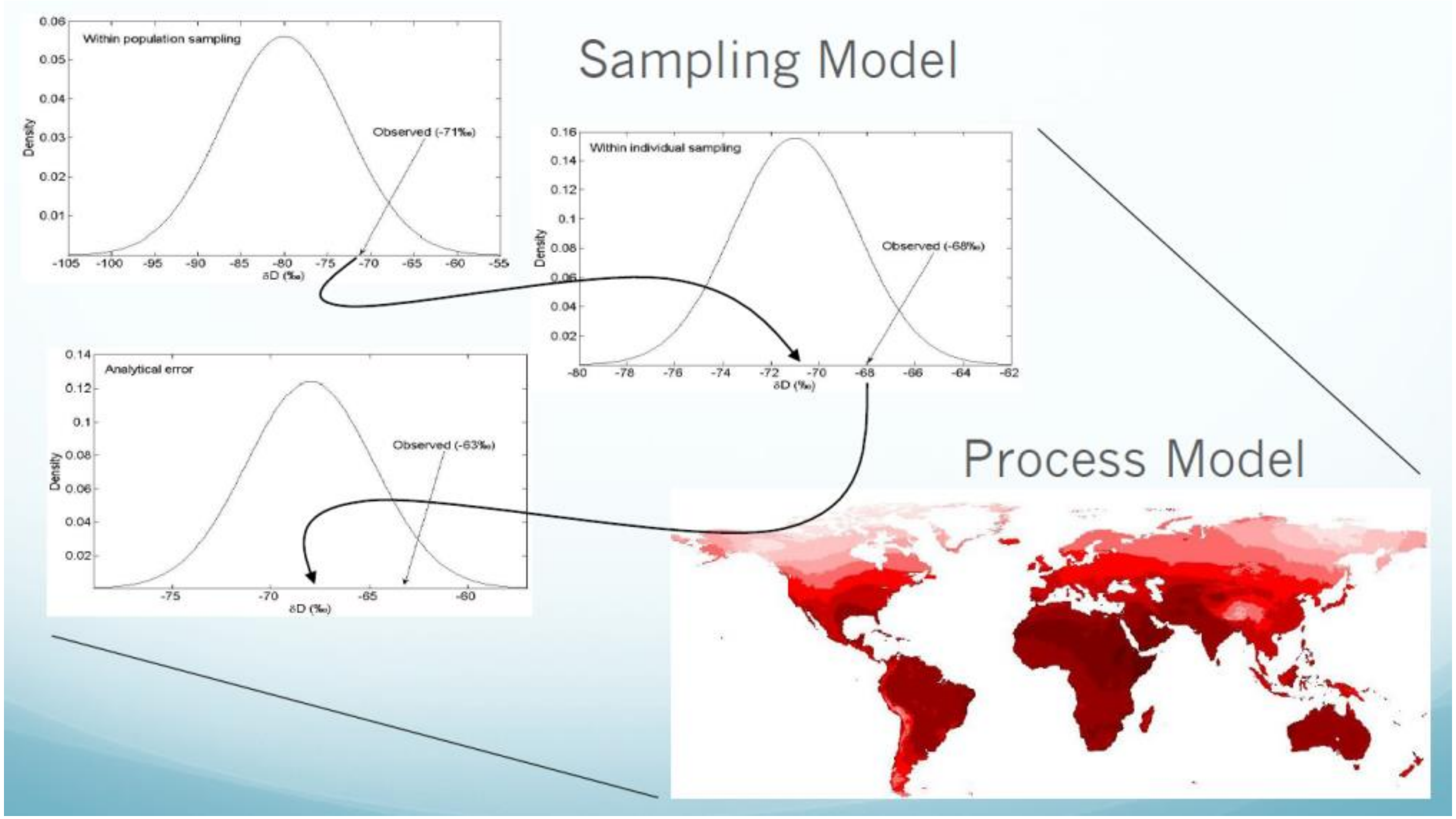
Aggregate Estimation

- Method 2: Use prediction intervals for tissue isotope calibration relationship
- Use sample statistics, not population statistics!



Hierarchical Estimation

- Build estimate from variance associated with individual model levels



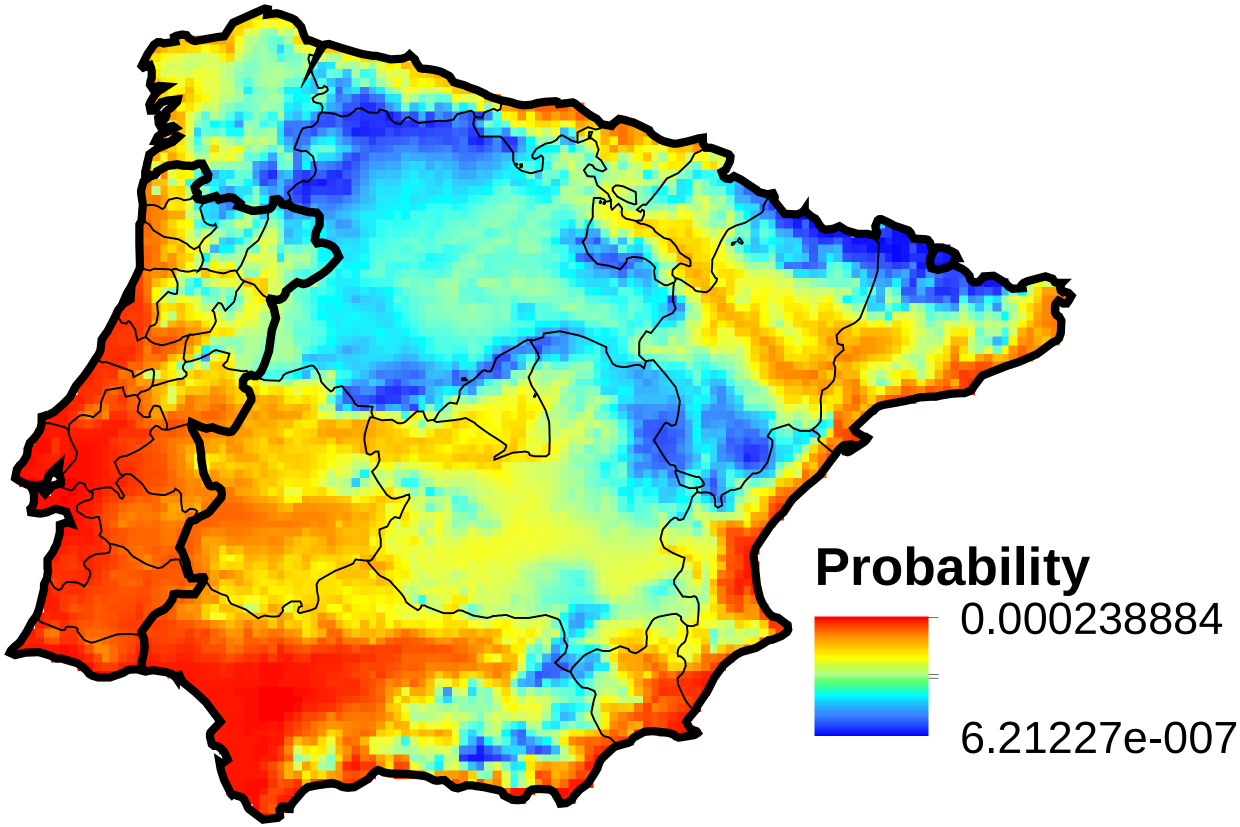
Wunder, 2010

Simple Semi-Parametric Empirical Bayesian Assignment

$$P(\delta_s|A_i) = \frac{1}{\sqrt{2\pi{\sigma_i}^2}} e^{\left(-\frac{(\delta_s-\mu_i)^2}{2{\sigma_i}^2}\right)}$$

- Assumes normally distributed PDF for sample values at a given location
- Aggregate or model-based estimate of within-site variance

Example Result



Incorporating Priors

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum P(B|A_j)P(A_j)}$$

- In most cases we have some form of prior information
  - Range maps
  - Population density
  - Band/recapture
- Easy to impose any of these on our continuous analysis IF we can represent the prior probability at each grid cell in our map area

Incorporating Priors

Chabot et al., 2012

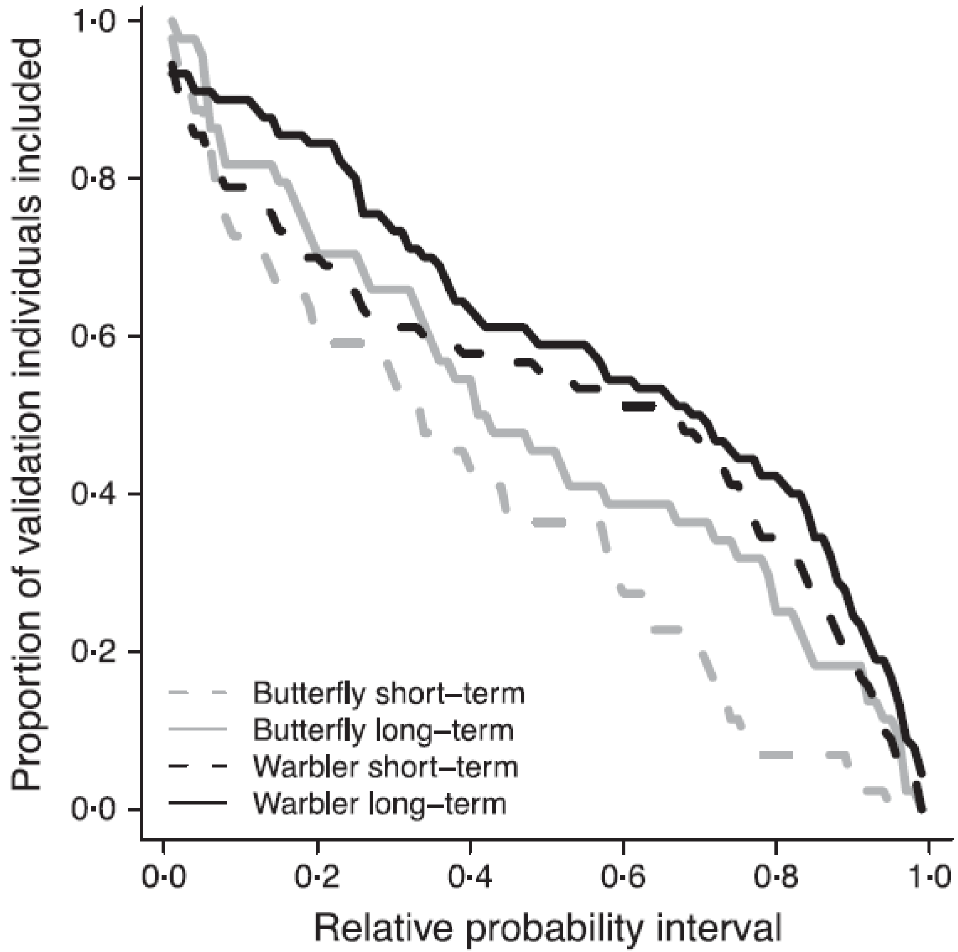
Interpreting Continuous Results

- Often (usually) we want to aggregate results to answer specific ecological or management questions
- Lots of flexibility to develop metrics suited to the question, but no single ‘right’ answer

- Evaluating accuracy and precision
- Comparing hypotheses (likelihood ratios)
- Binary assignment (yes/no)
- Working with multiple individuals

Known-origin Tests

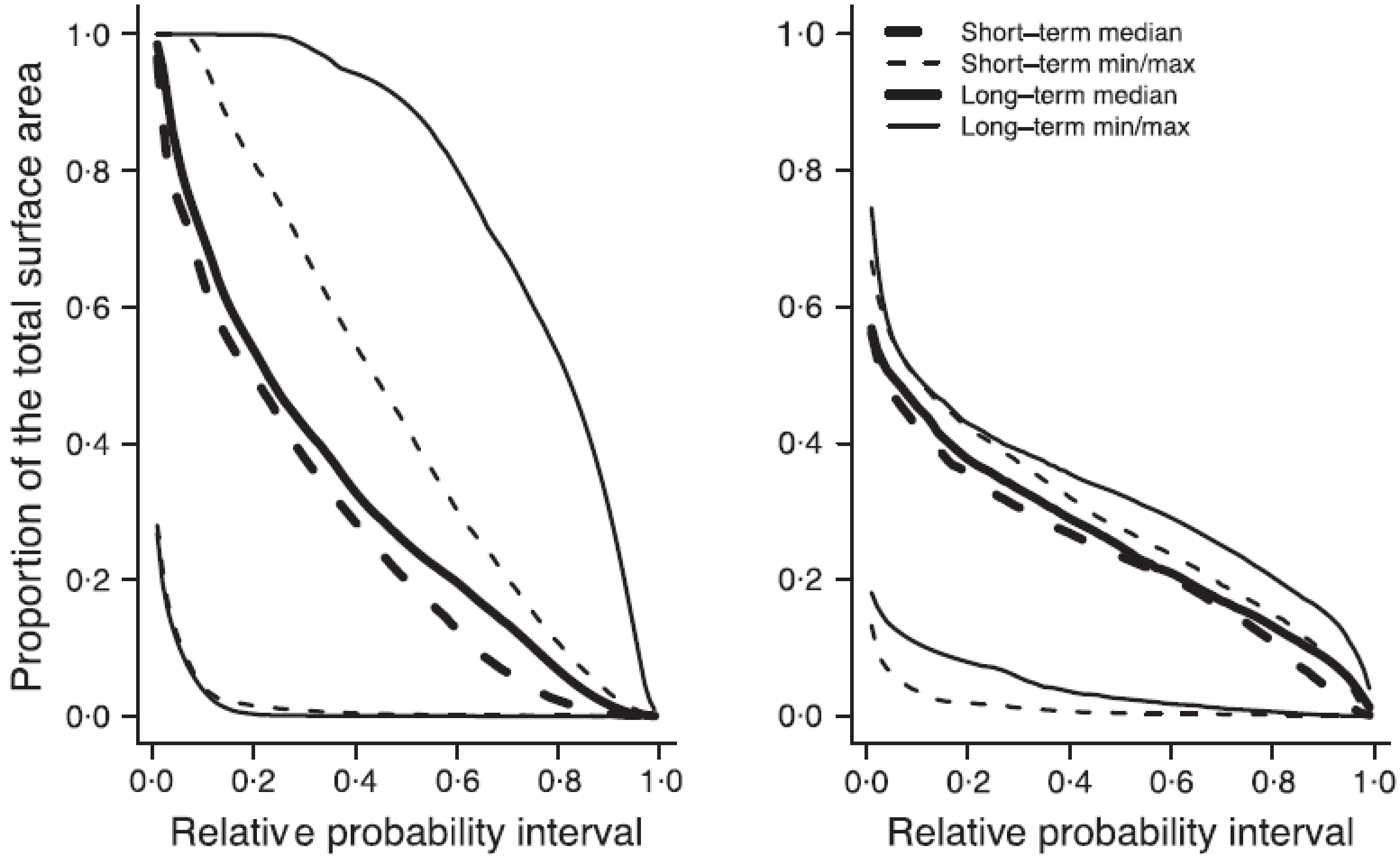
- Accuracy



Vander Zanden et al., 2014

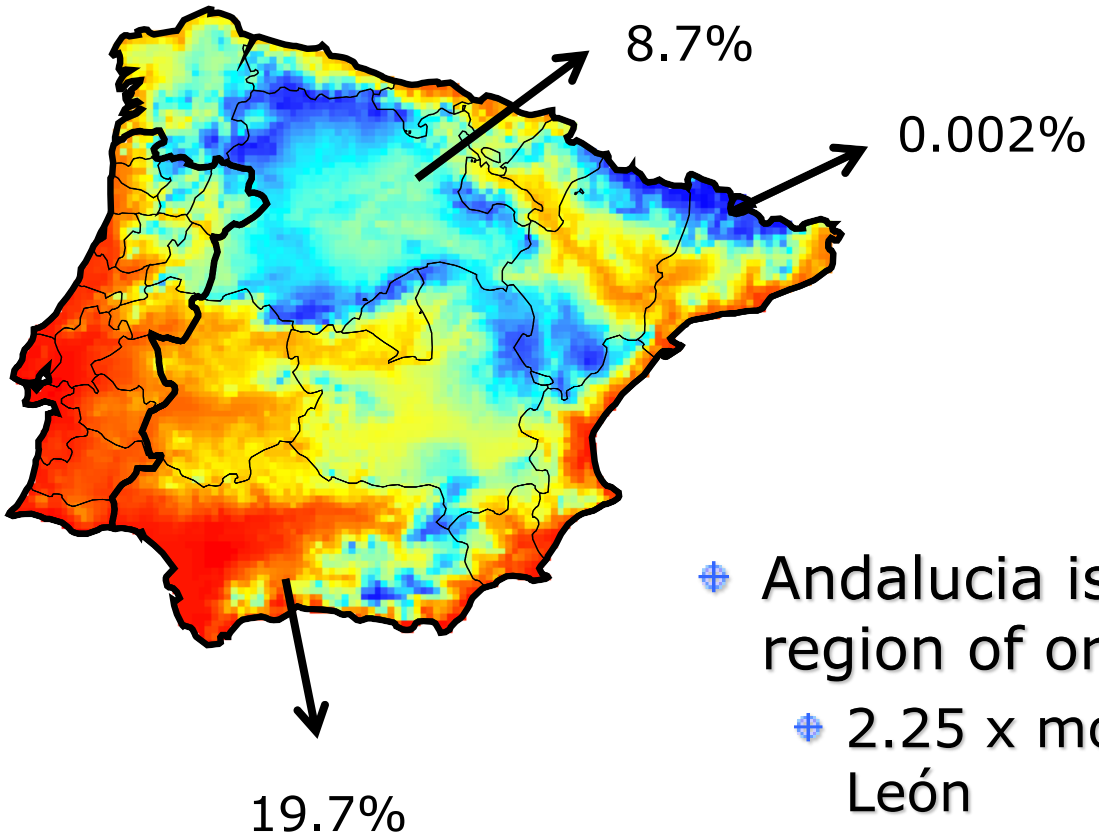
Known-origin Tests

- Precision



Vander Zanden et al., 2014

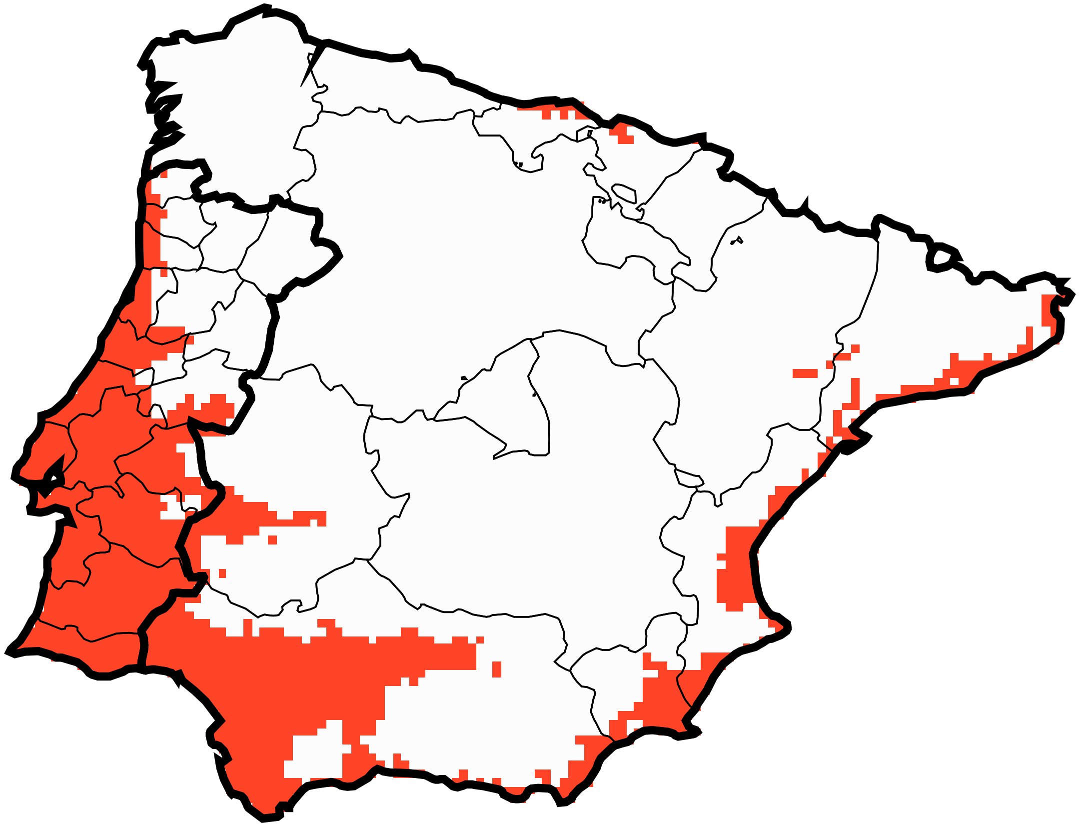
Comparing Hypotheses



- Andalucia is the most likely region of origin
  - 2.25 x more likely than Castilla y León
  - 7,600 x more likely than Andorra

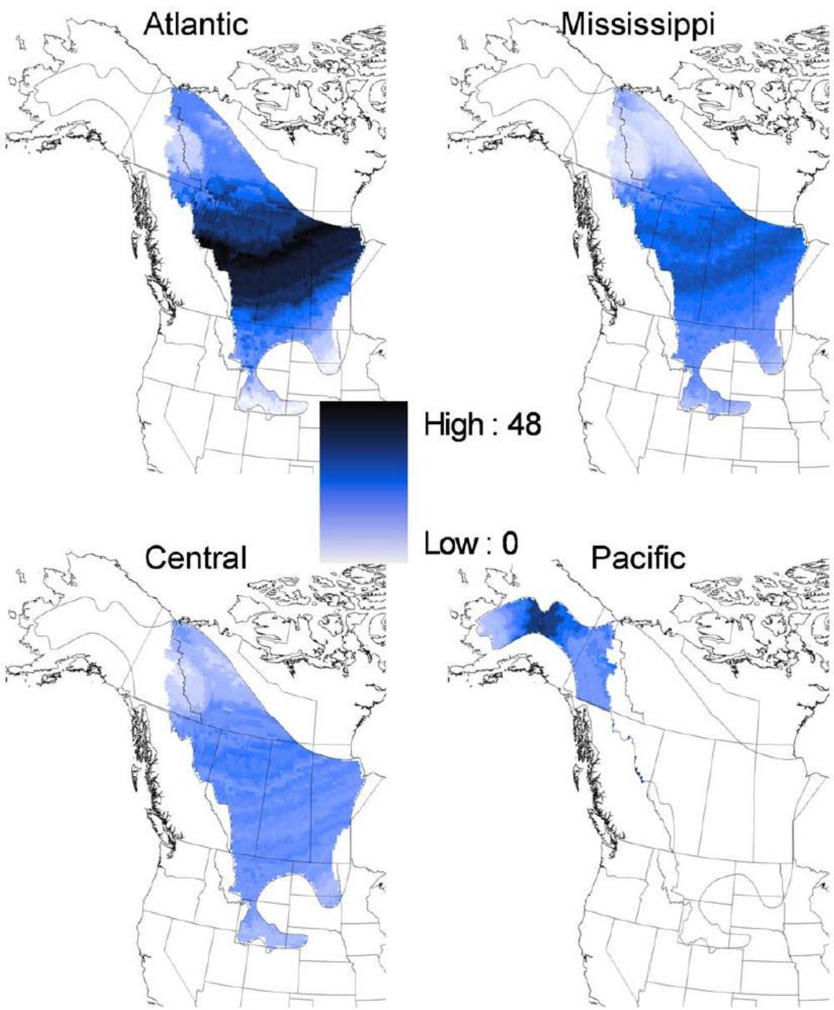
Binary Assignment

- Choose threshold, mask area of 'possible' origin



Multiple Individuals

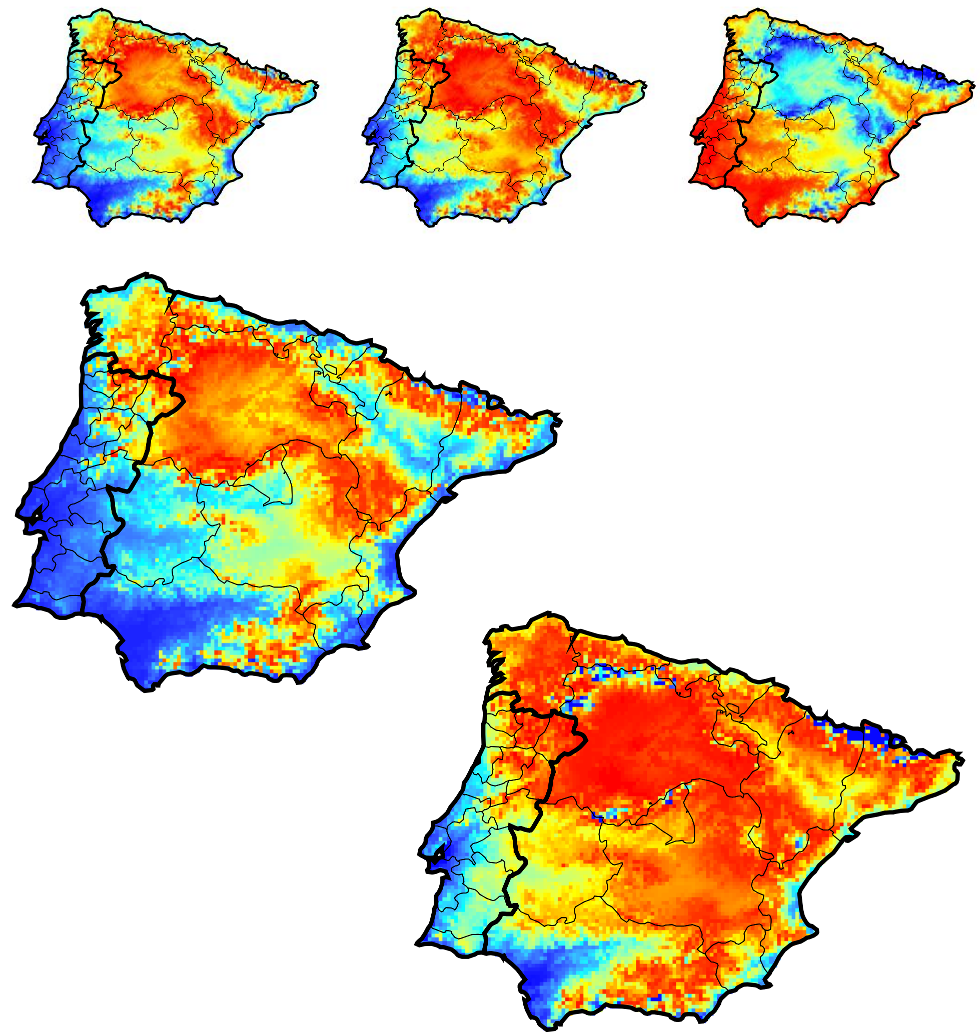
- Most studies (hopefully!) have >1 sample
- How do we summarize information from individuals?
- Binary assignment -> summation



Hobson et al., 2009

Multiple Individuals

- Joint probability
- Probability that BOTH samples are from a location (intersection):
  - $P(A \cap B) = P(A) \times P(B)$
- Probability that ANY sample is from a location (union):
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



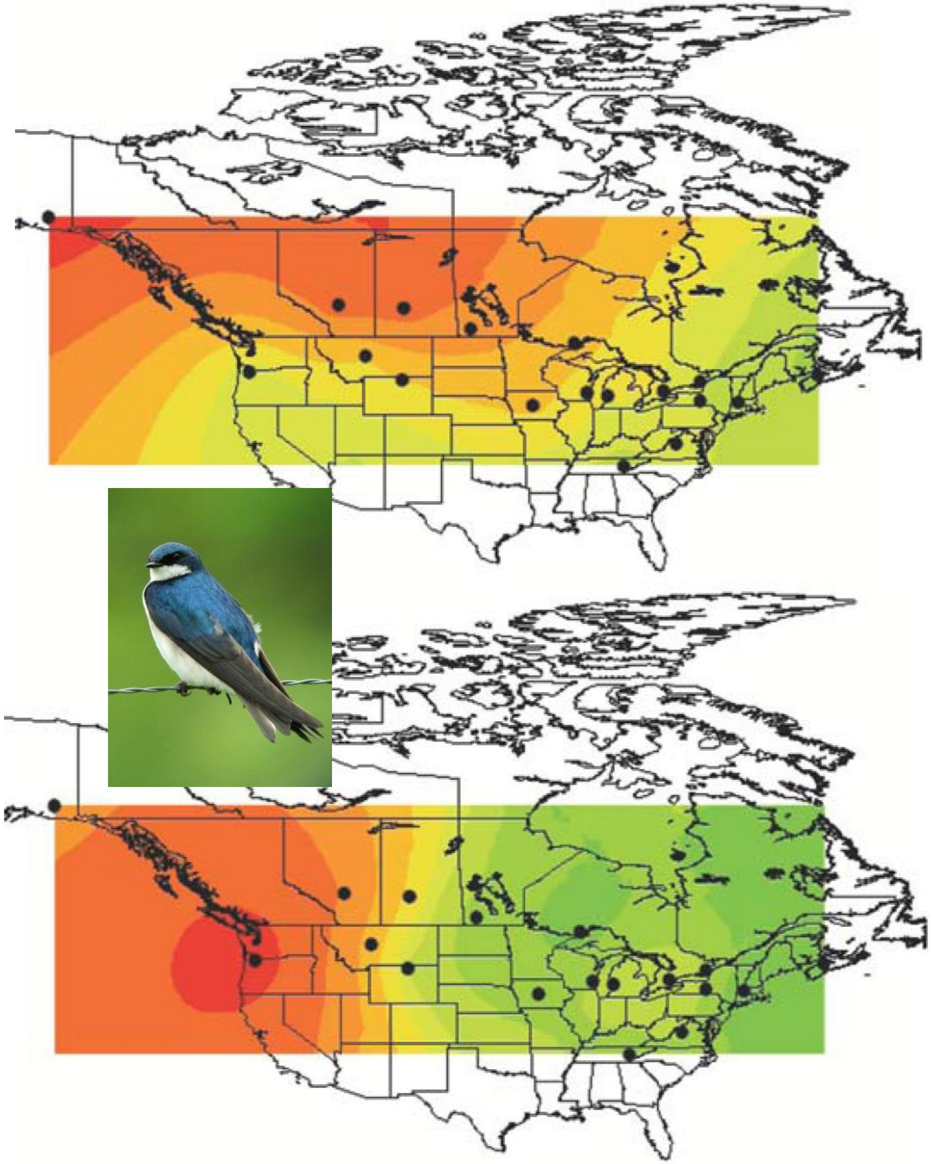
Hobson et al., 2009

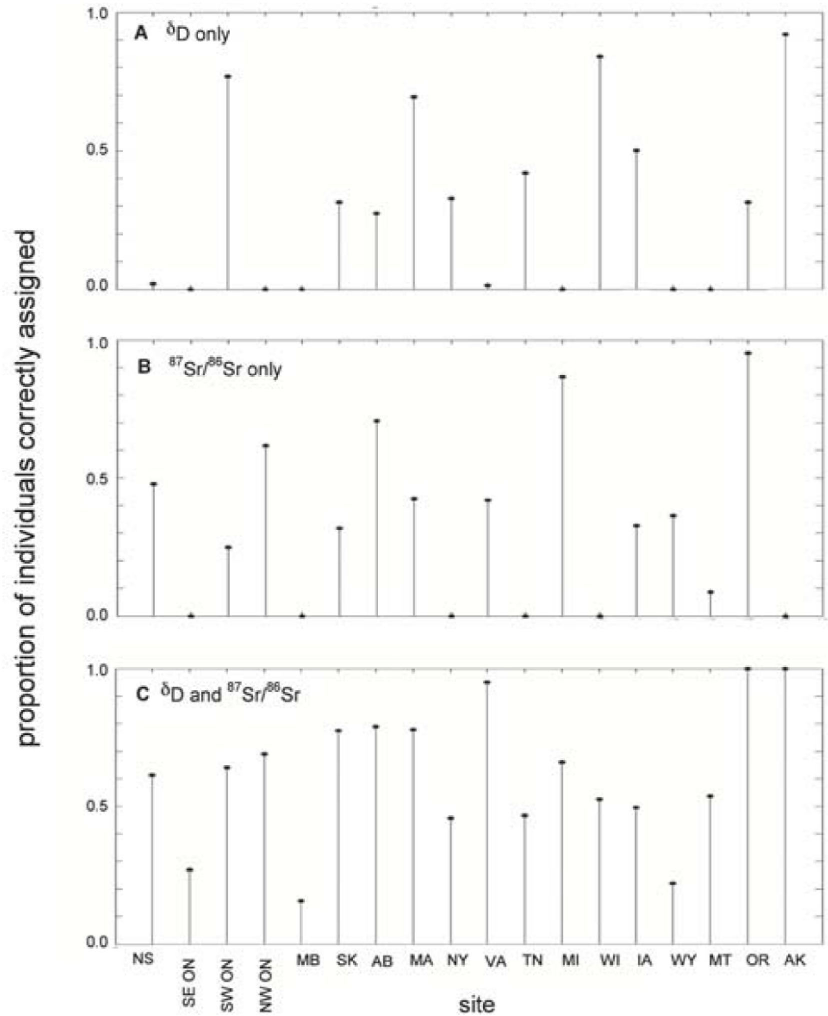
Multiple Markers

- Multiple markers (isotopes, elements, etc.) can increase precision of results
  - Requires more data, more models, more assumptions!
  - Not all marker systems are structured similarly or useful in the same way!

- Two approaches:
  - Assume independence, calculate joint conditional probability $P(B_1, B_2|A_i) = P(B_1|A_i) \times P(B_2|A_i)$
  - Characterize the covariance structure of the joint spatial distribution of the markers

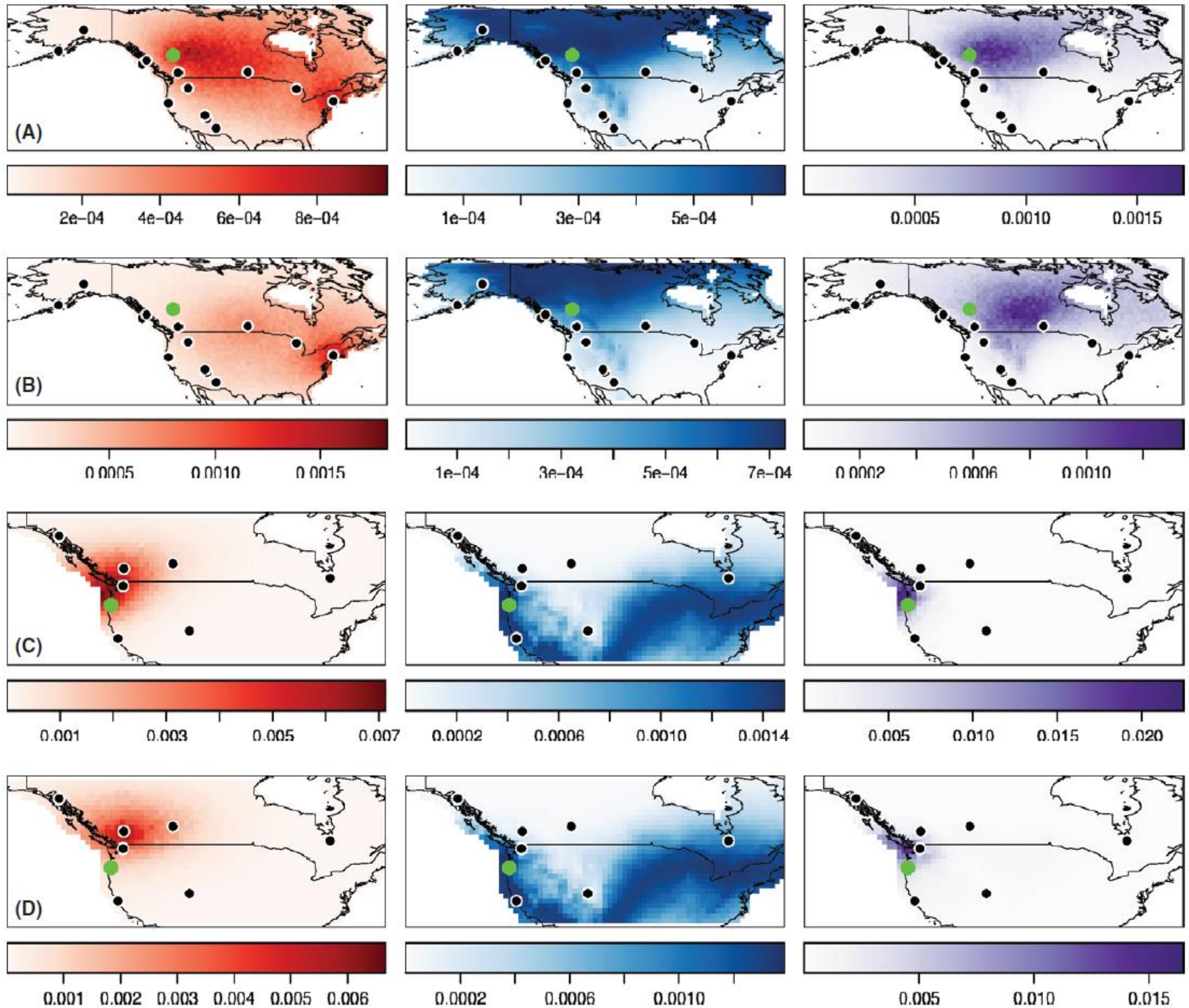
Multiple Markers





Sellick et al, 2009

Multiple Markers



Rundel et al., 2013