Why Does Measurement Matter?

Measurement generally requires defining scales for data, conditioned on a data-generating process



Measurement is a process of transferring observations to a symbolic space subject to isomorphic transformations





Human-devised scales are often inadequate for measurement



Evolutionary processes are high-dimensional and have many potential confounders

(1) Houle, D., Pélabon, C., Wagner, G. P., & Hansen, T. F. (2011). Measurement and Meaning in Biology. https://doi.org/10.1086/658408. (2) Raup, D. M. (1966). Geometric Analysis of Shell Coiling: General Problems. http:// www.jstor.org/stable/1301992. (3) Mridul Khurana et al (2024). Hierarchical Conditioning of Diffusion Models Using Tree-of-Life for Studying Species Evolution. https://arxiv.org/abs/2408.00160. (4) Mohannad Elhamod et al (2023). Discovering Novel Biological Traits From Images Using Phylogeny-Guided Neural Networks. https://arxiv.org/abs/2306.03228. (5) Gomulkiewicz, R., Kingsolver, J., Carter, P., & Heckman, N. (2018). Variation and Evolution of Function-Valued Traits. https://doi.org/10.1146/annurev-ecolsys-110316-022830. (6) Dean C. Adams, A Generalized K Statistic for Estimating Phylogenetic Signal from Shape and Other High-Dimensional Multivariate Data. https://doi.org/10.1093/sysbio/syu030.

E¹ E² X E⁴

Latent Space Phenotyping for Measuring **Complex Evolutionary Traits**

Caleb Charpentier, Josef Uyeda Virginia Tech Department of Biological Sciences

Traits can take numerical, functional, or other complex values

KGML Performance \propto Accuracy + Simplicity + Consistency Rapid progress on any biological problem rests on the hope that there is at least one viewpoint to each problem that makes causation relatively simple. -- David Houle For every complex problem there is a solution which is clear, simple and wrong. b) RGB space -- Often attributed to H.L. Mencken d) 27 bin **Phylogenetic Tree Training Stage** (discretized at 4 levels) 12 5 11 14 Hierarchical Embedding Species A f) 125 bins Inference Stage **Perturbation at Level 3**

· L3

Phylogenetic grounding can make trait discovery more interpretable

Traits on Complex Scales

Biological patterns can be represented in diverse mathematical ways

Age (days)

b Aware researcher's plot

Tests of Performance

A Generalized K Statistic for Estimating Phylogenetic Signal from Shape and Other High-Dimensional Multivariate Data *6

Department of Ecology, Evolution, and Organismal Biology, and Department of Statistics, Iowa State University, Ames IA, 50011, USA; *Correspondence to be sent to: Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011, USA;



(b) Phylogenetic ground truth

Level 4	\mathbf{E}^{1}	\mathbf{E}^2	E^3	E^4
			_	
Level 3	E^1	E^2	E^3	X
Level 2	E^1	E^2	Х	Х
Level 1	E^1	Х	Х	Х
(a) Trait Masking				

Various cross validation and summary statistic-based tests can be used to evaluate model performance



Diffusion

Simulated images from the TraitBlender pipeline





DEAN C. ADAMS*



