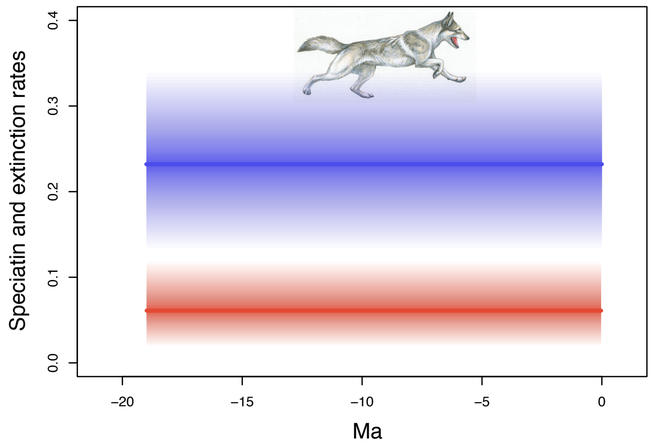
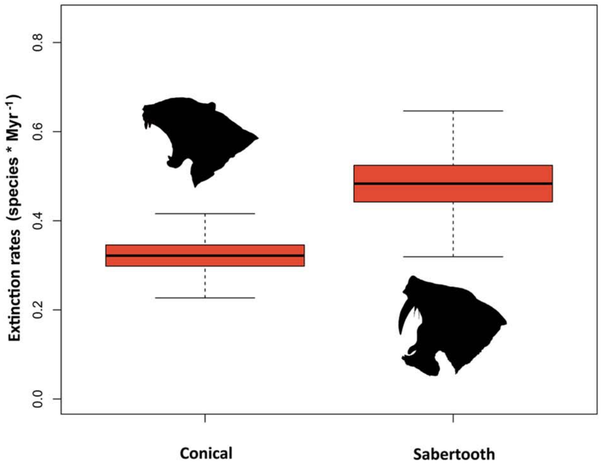
So many flavors of the birth-death model…

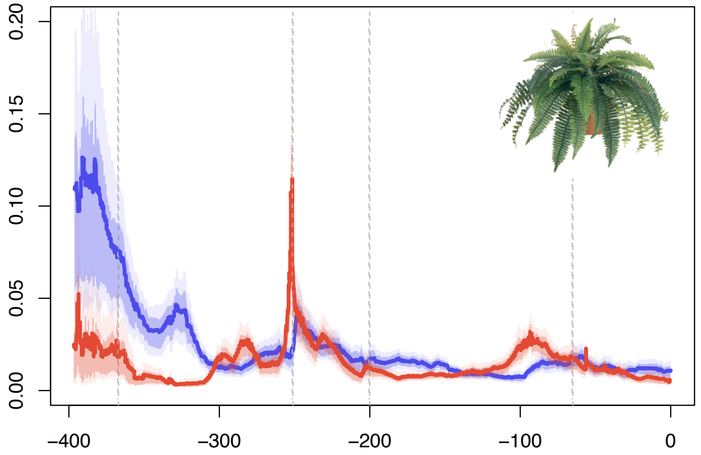
Constant rate



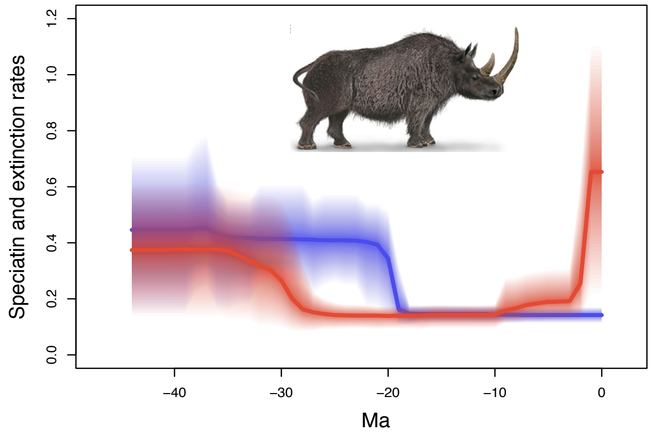
Trait-dependent extinction



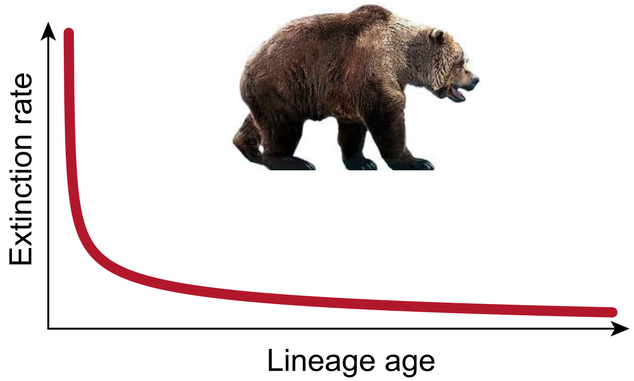
Multivariate BD model



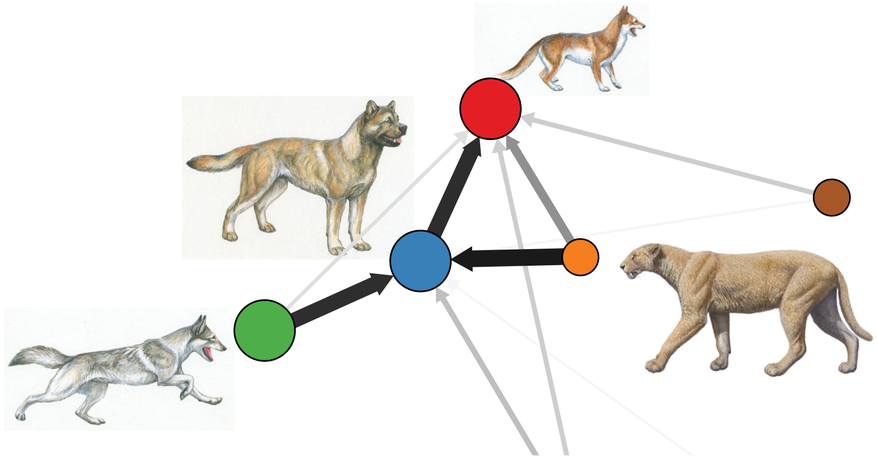
Time-variable rate



Age dependent extinction models

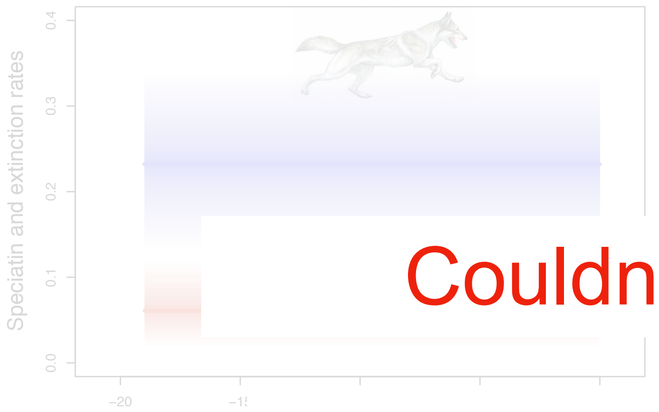


Diversity dependence within and between clades



So many flavors of the birth-death model…

Constant rate



Trait-dependent extinction

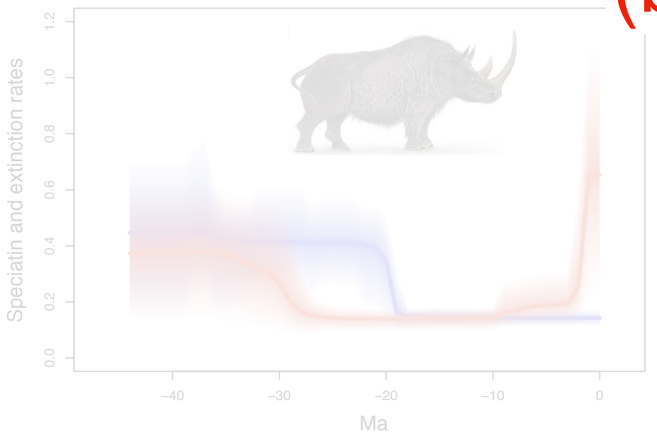
Multivariate BD model

Couldn't we have one model that does it all? 🤔

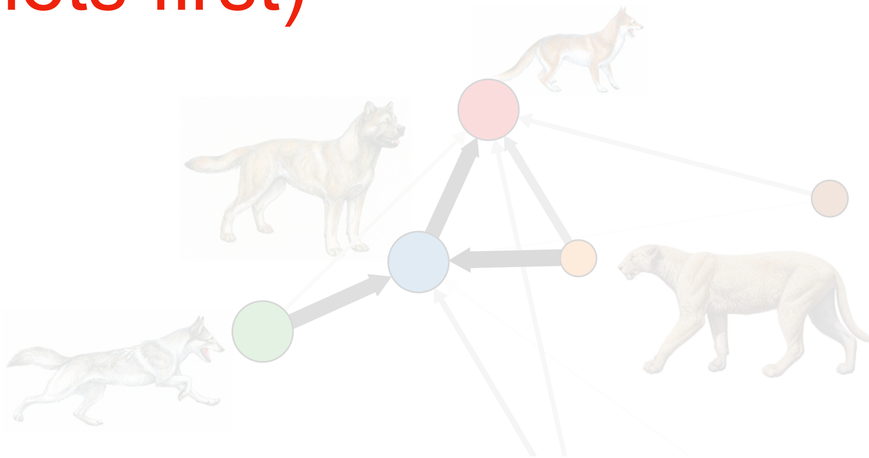
Well, perhaps we can… 🦄

(but let's talk about neural nets first)

Time-variable rate

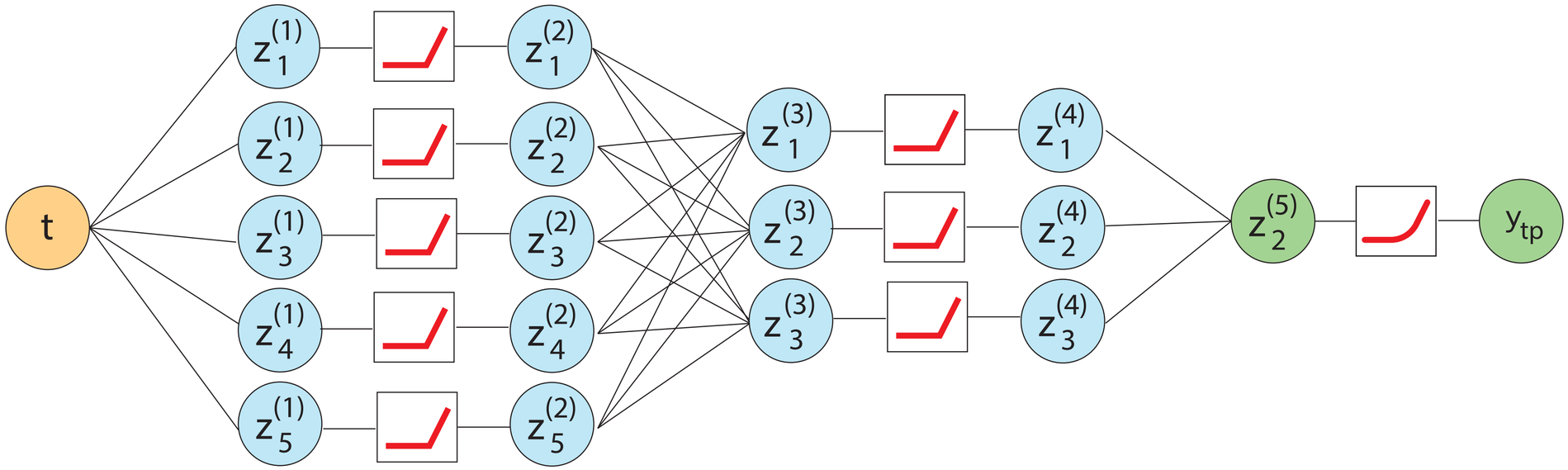


ıce within and between clades



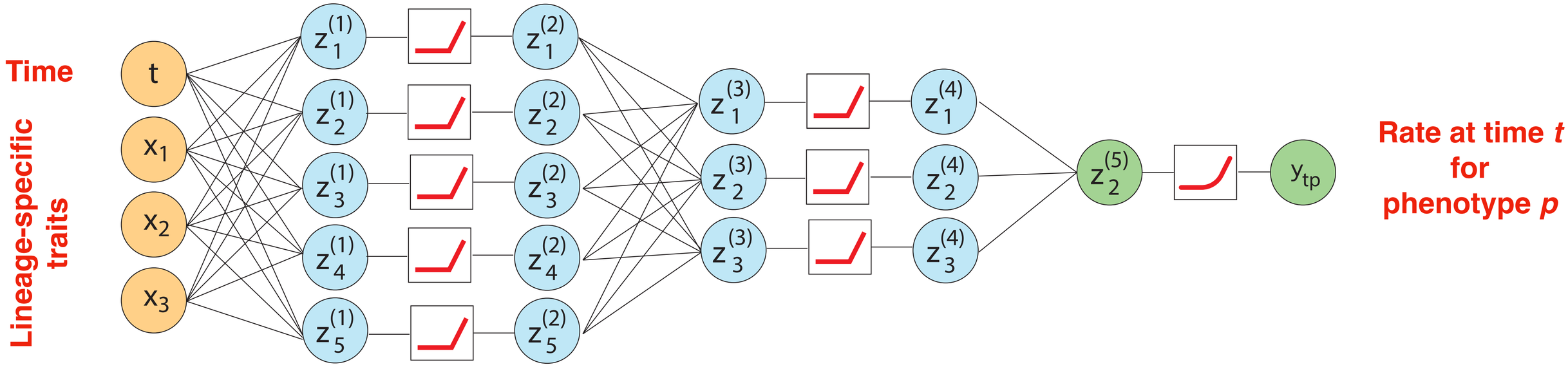
Time-dependent Birth-Death-NN model

Time



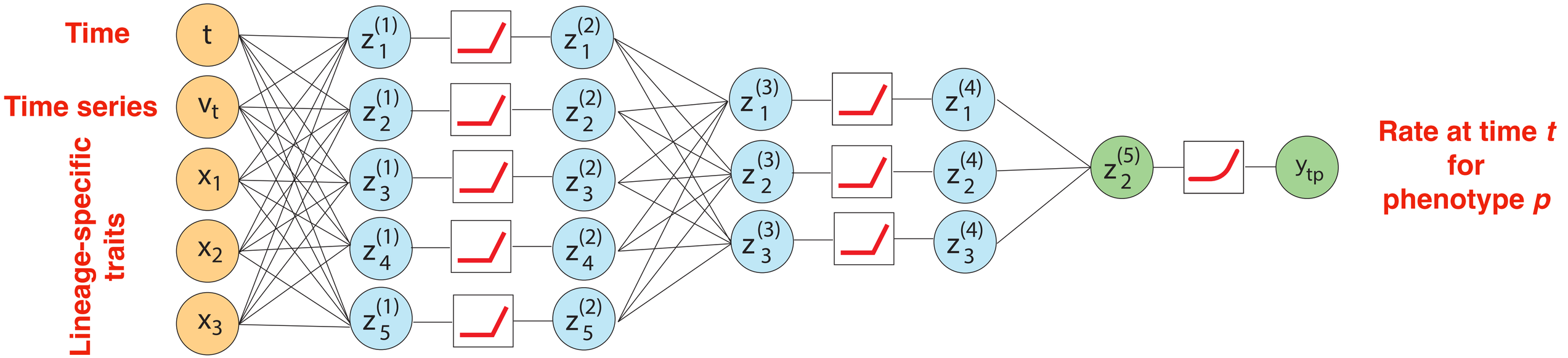
Speciation or extinction rate at time *t*

Time-dependent and trait-dependent BDNN model



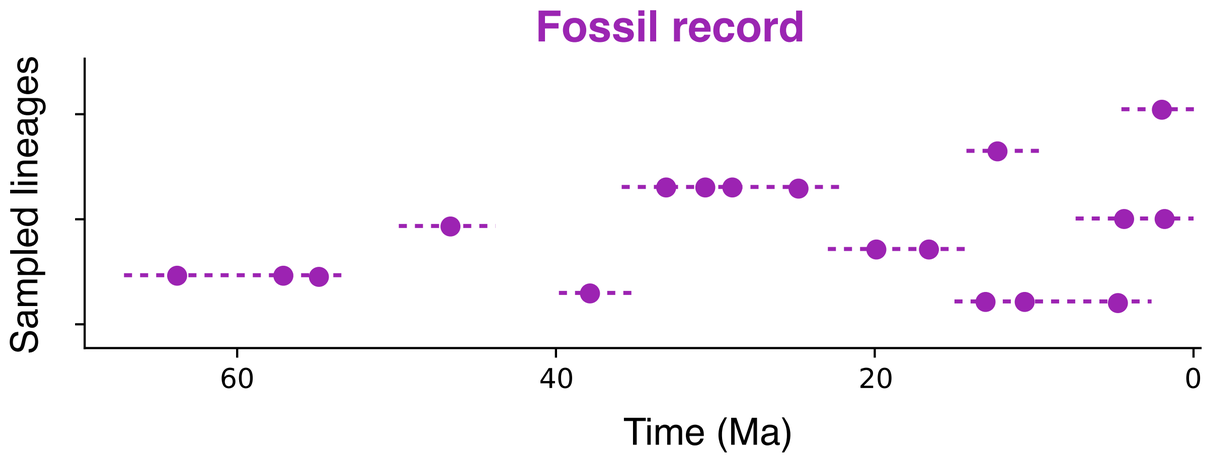
The model can include (continuous and discrete) traits and allows for species-specific rates

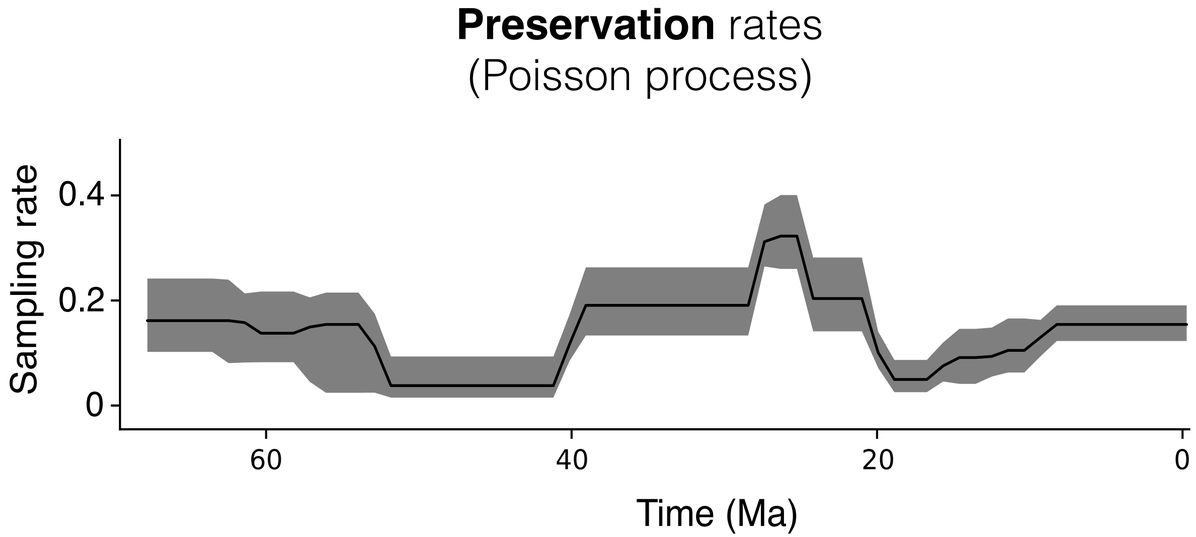
Time-dependent and trait-dependent BDNN model

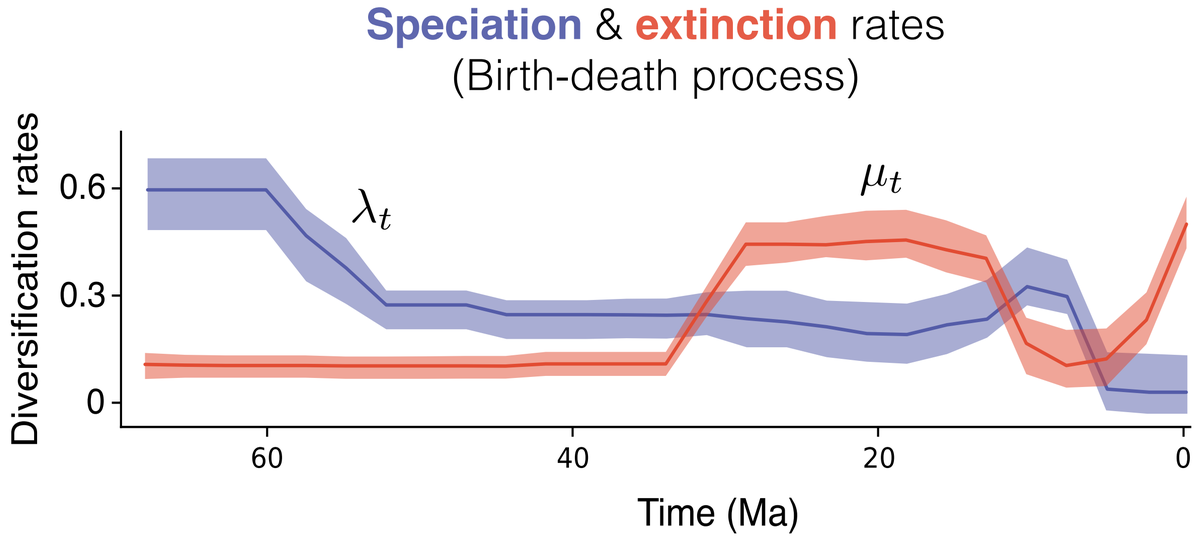


The model can include (continuous and discrete) traits and time series

Bayesian (unsupervised) estimation of speciation and extinction rates from fossils



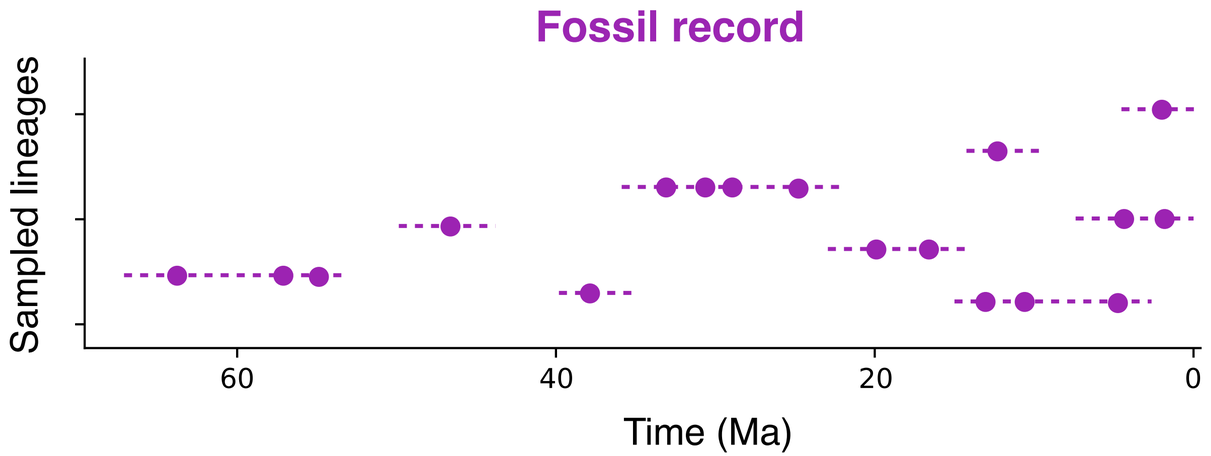




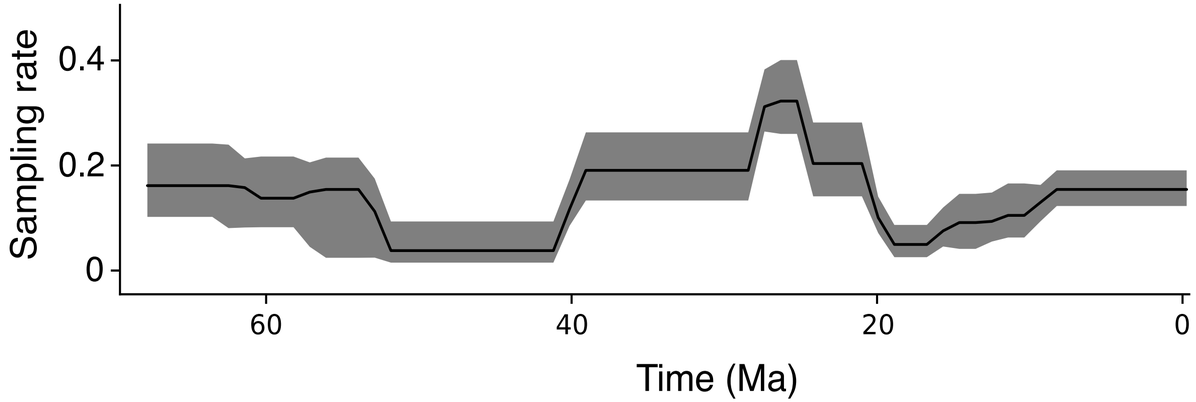
github.com/dsilvestro/PyRate

Silvestro et al. 2019 Paleobiology

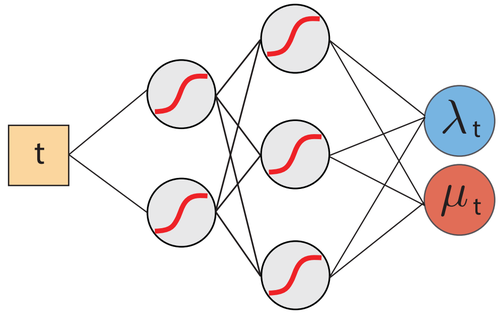
A birth-death neural network model of speciation and extinction



Preservation rates



Time-varying **Speciation** & **extinction** rates modeled by a NN

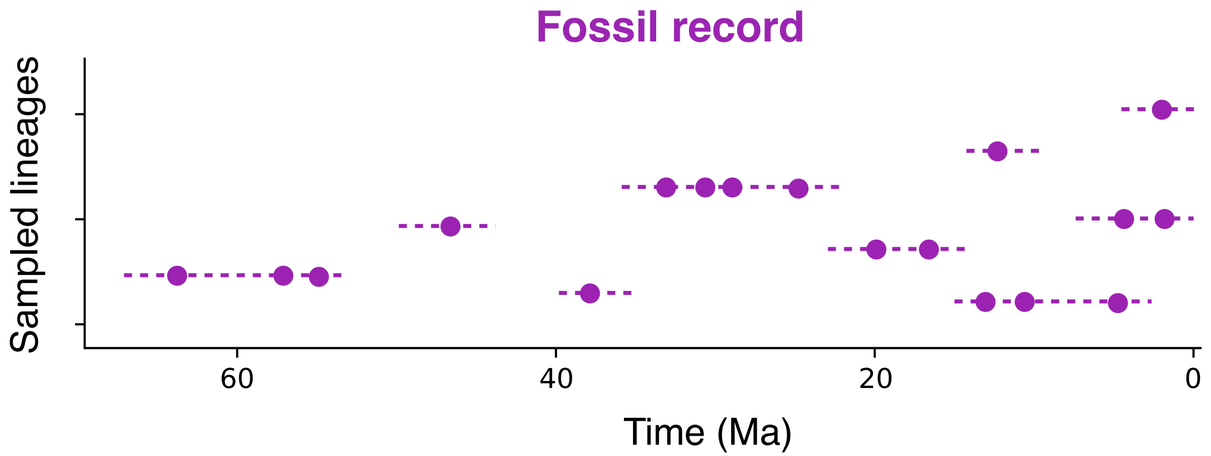


T Hauffe

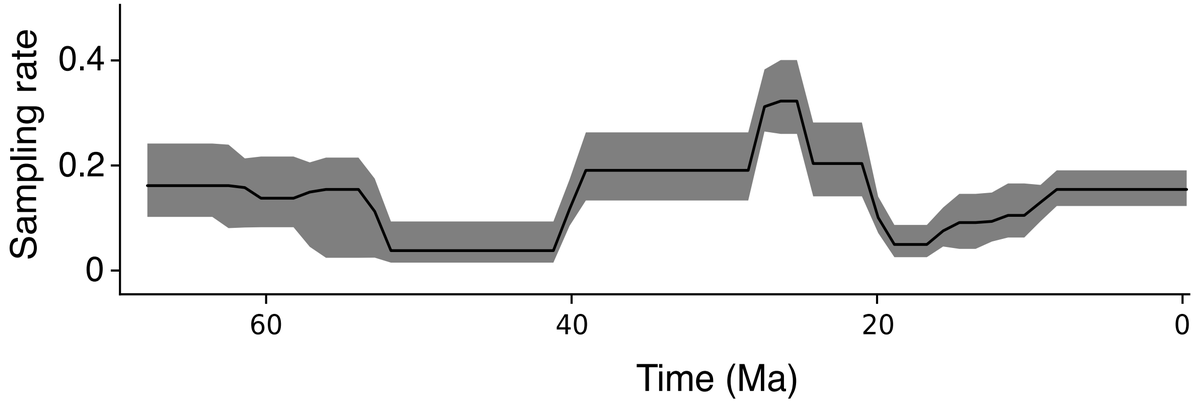
github.com/dsilvestro/PyRate

Hauffe et al. 2024 Science Advances

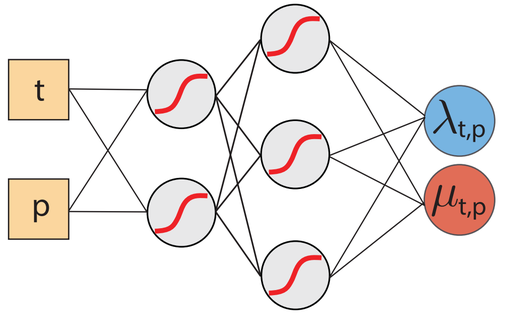
A birth-death neural network model of speciation and extinction



Preservation rates



Species-specific and time-varying
Speciation & **extinction** rates



T Hauffe

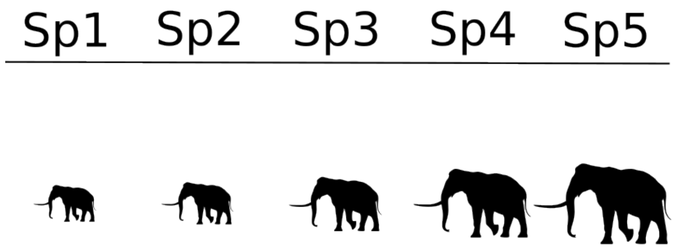
github.com/dsilvestro/PyRate

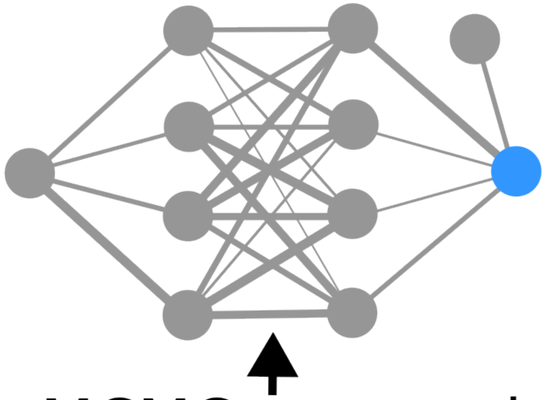
Hauffe et al. 2024 Science Advances

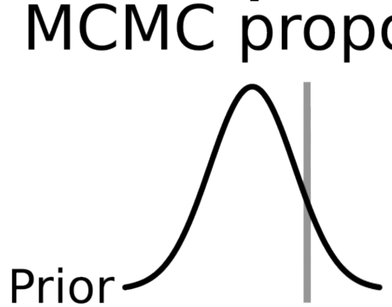
Unsupervised Bayesian neural networks and xAI to infer factors governing diversification

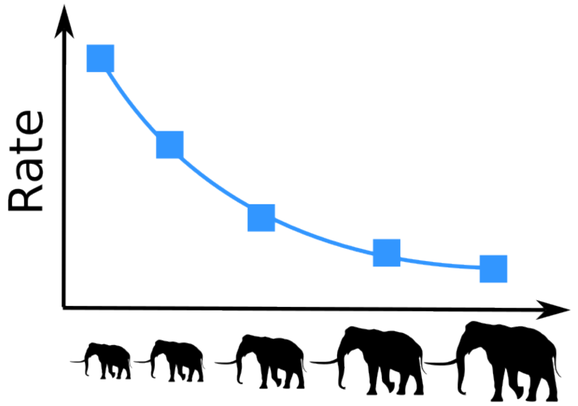
1. Predictors

2. Transformation



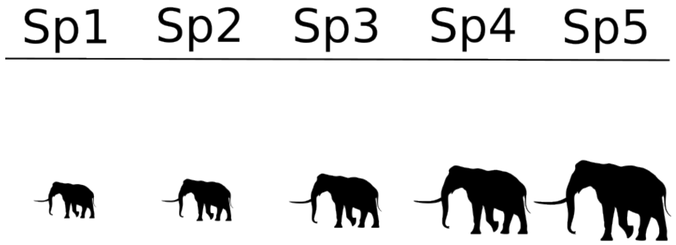




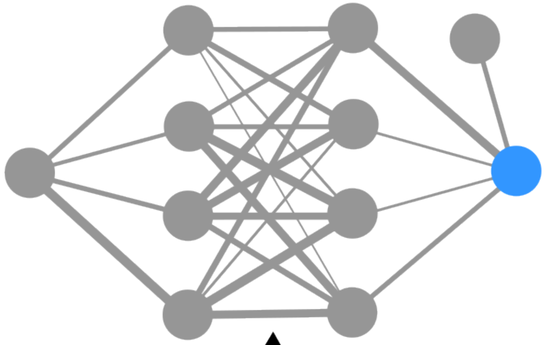


Unsupervised Bayesian neural networks and xAI to infer factors governing diversification

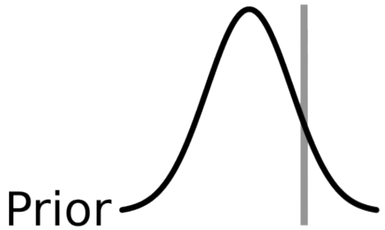
1. Predictors

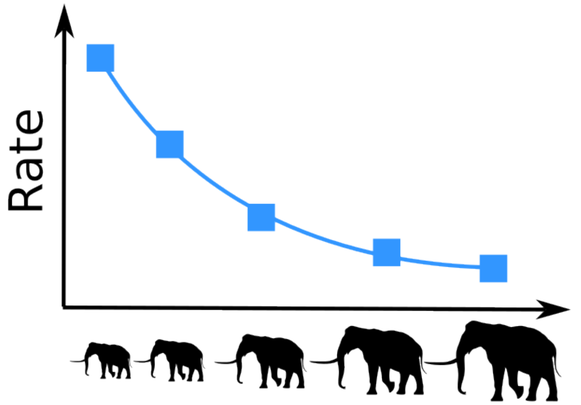


2. Transformation

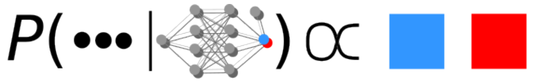


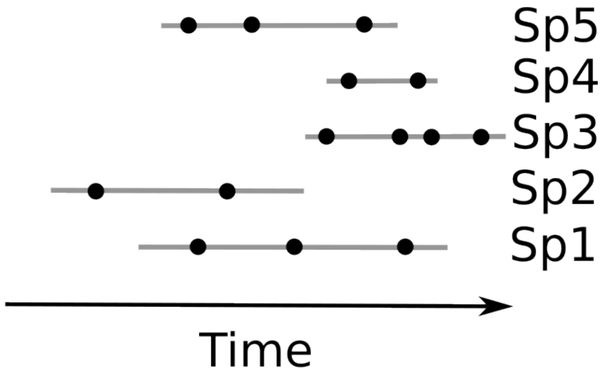
MCMC proposal





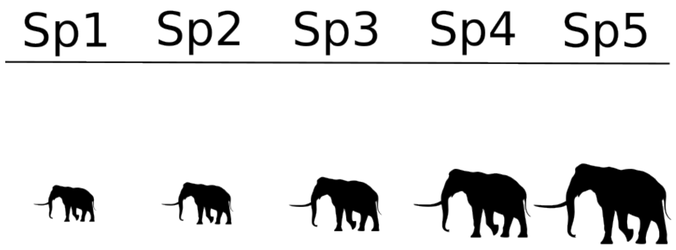
3. Likelihood



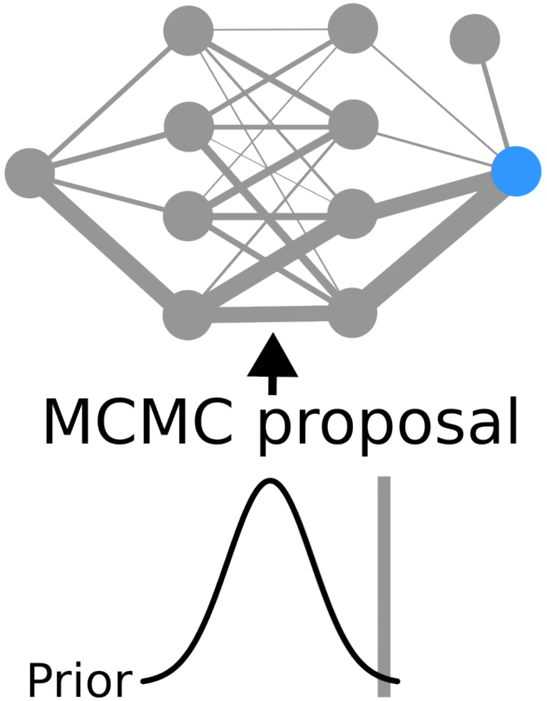


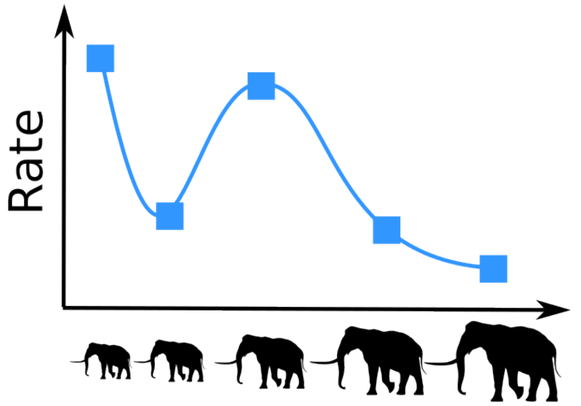
Unsupervised Bayesian neural networks and xAI to infer factors governing diversification

1. Predictors

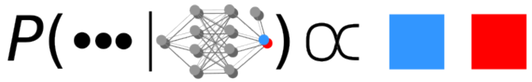


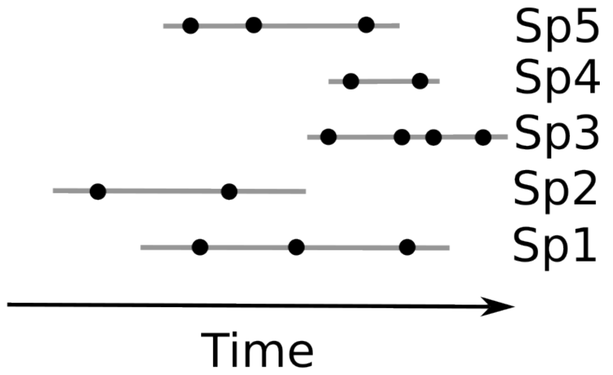
2. Transformation





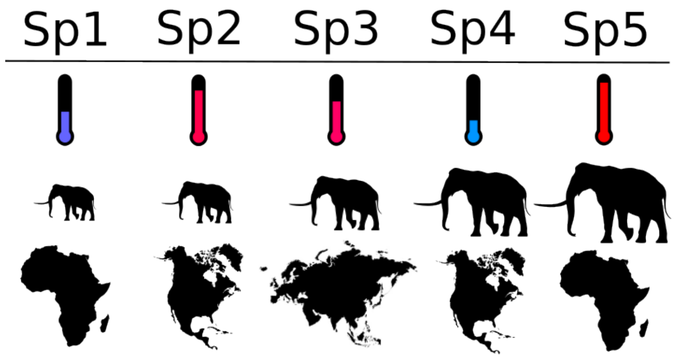
3. Likelihood



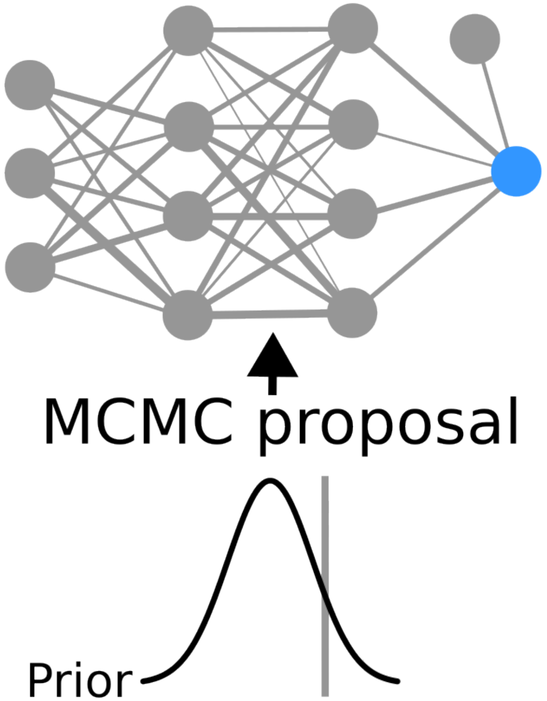


Unsupervised Bayesian neural networks and xAI to infer factors governing diversification

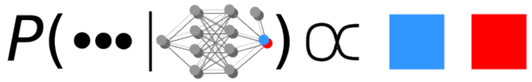
1. Predictors

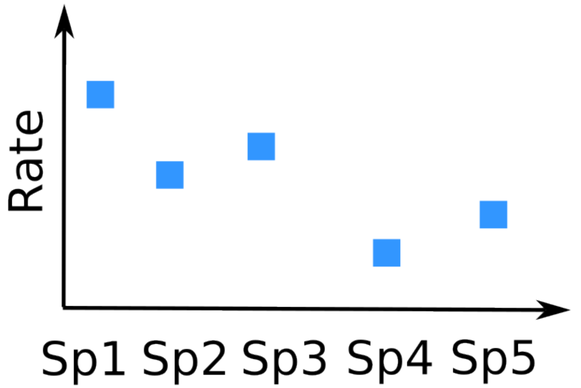


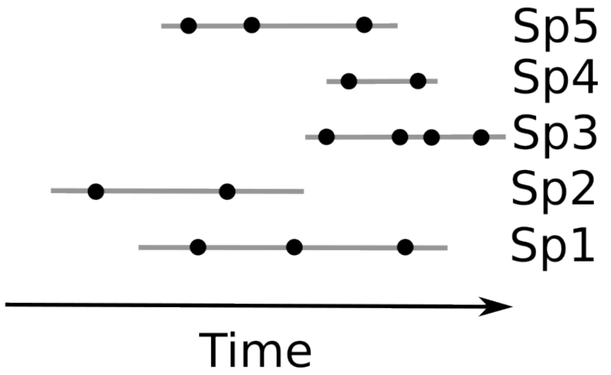
2. Transformation



3. Likelihood







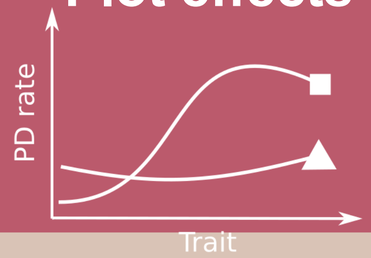
Unsupervised Bayesian neural networks **and xAI** to infer factors governing diversification

Infer BDNN

Coefficient of rate variation

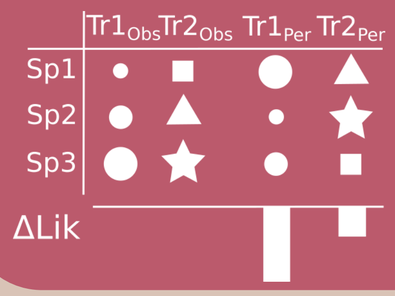
Explainable AI

Plot effects

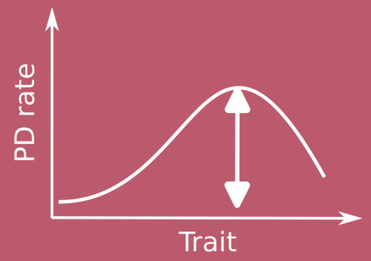


Predictor identification

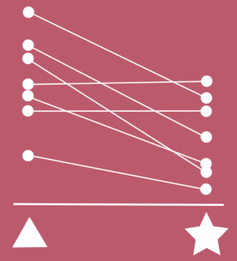
Decrease model fit



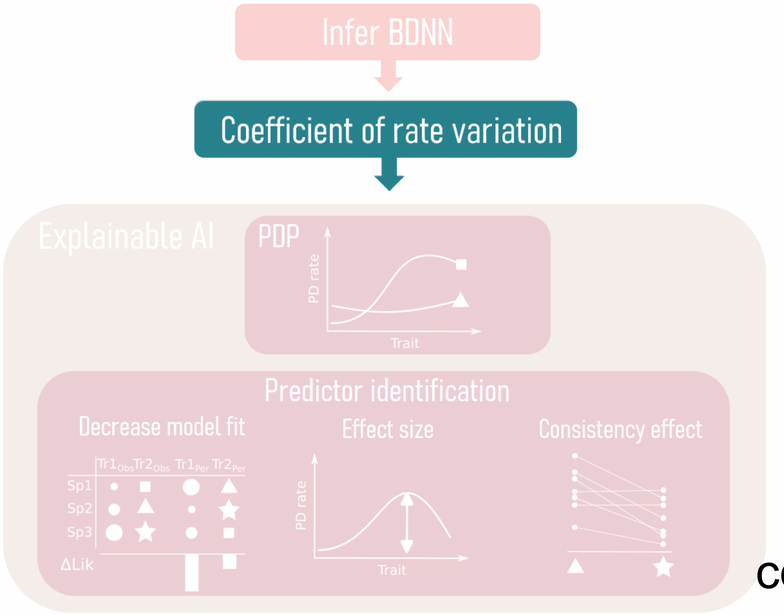
Effect size

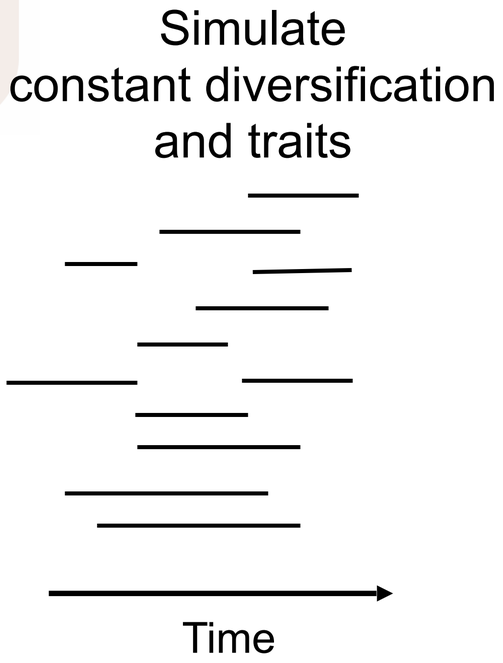


Consistency effect

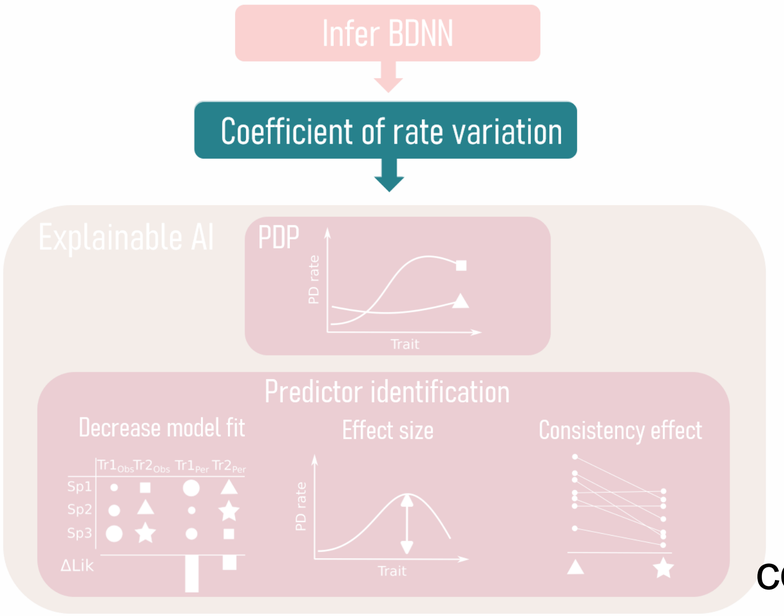


Should we dig deeper into factors? Threshold for variation in species-specific rates





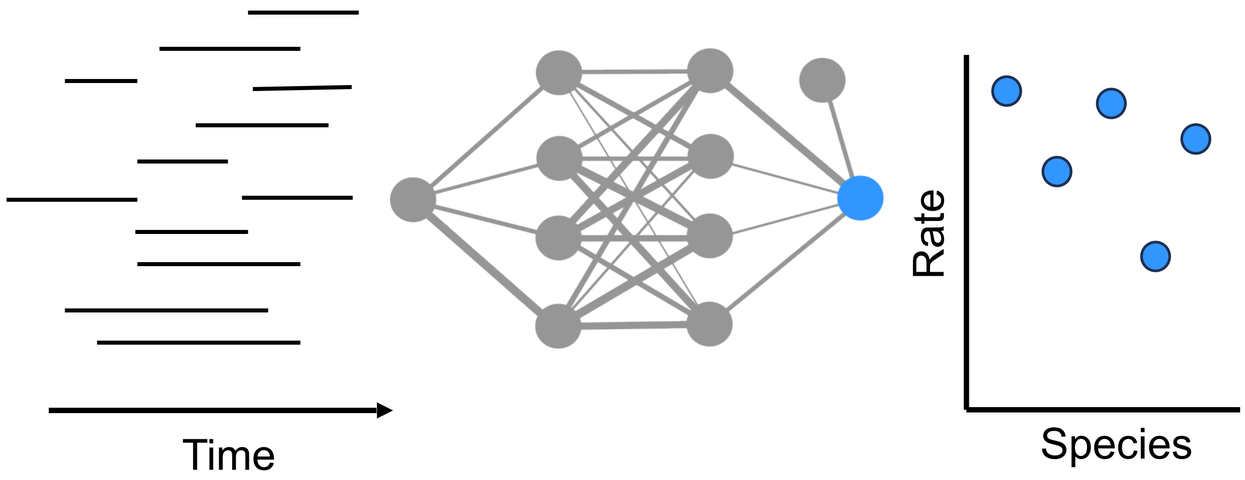
Should we dig deeper into factors? Threshold for variation in species-specific rates



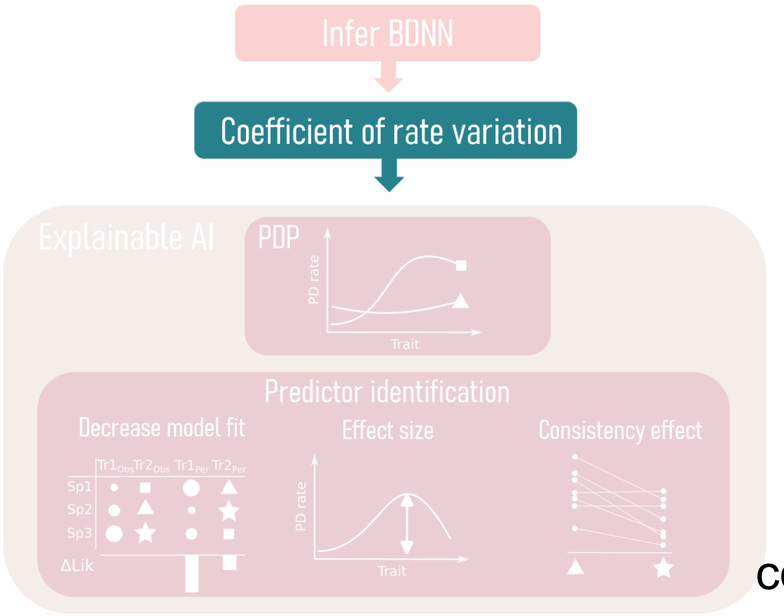
Simulate constant diversification and traits

Infer BDNN

Variation in species-specific rates



Should we dig deeper into factors? Threshold for variation in species-specific rates

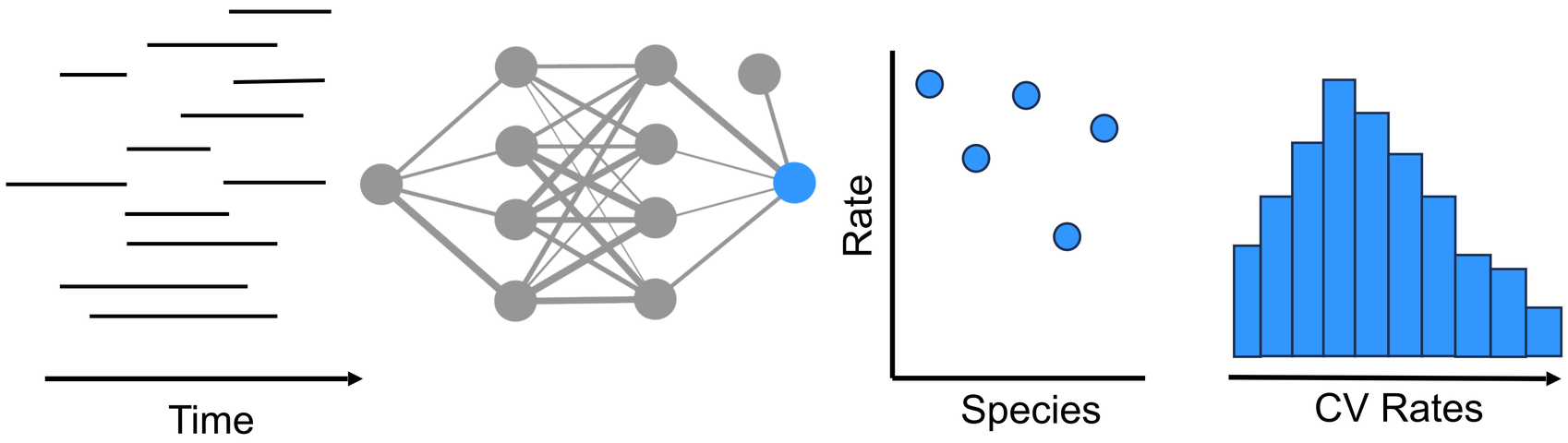


Simulate constant diversification and traits

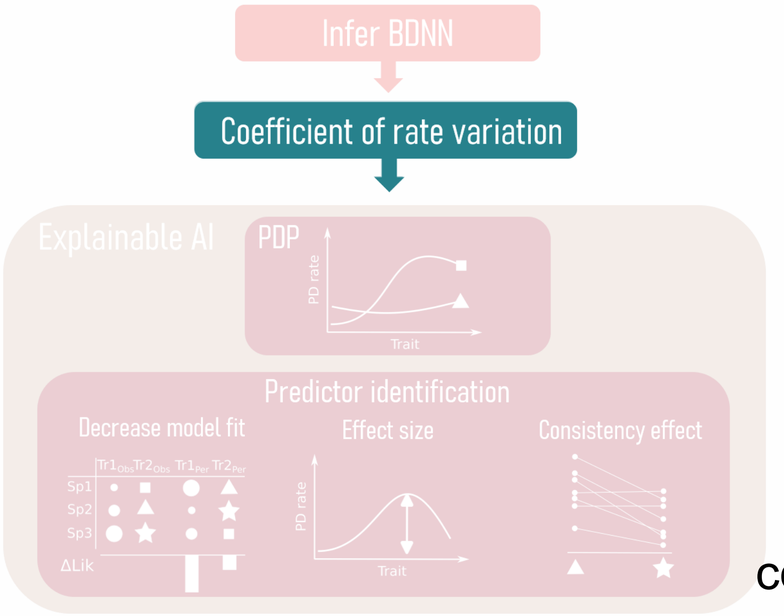
Infer BDNN

Variation in species-specific rates

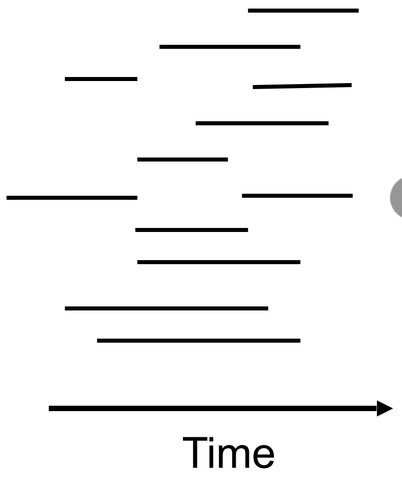
100x simulations



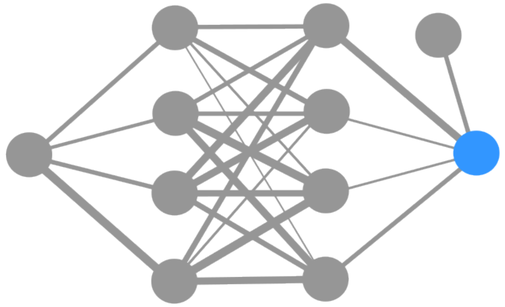
Should we dig deeper into factors? Threshold for variation in species-specific rates



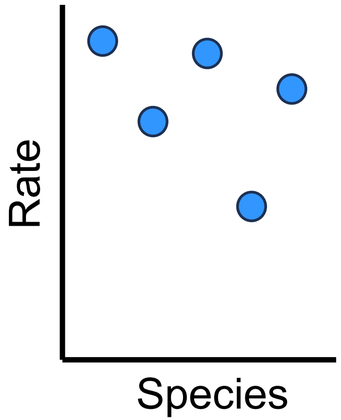
Simulate constant diversification and traits



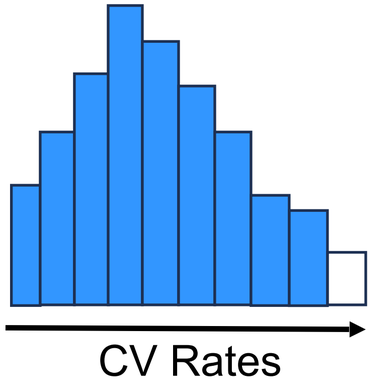
Infer BDNN



Variation in species-specific rates

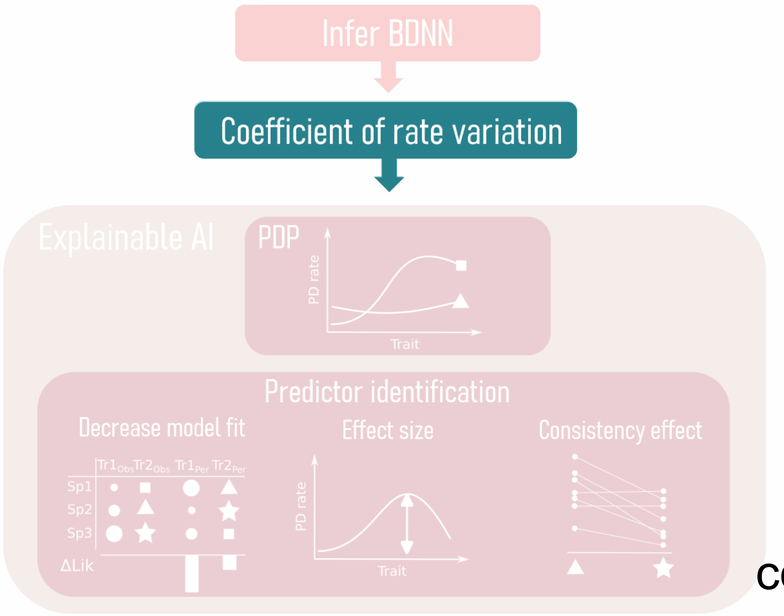


100x simulations to get threshold

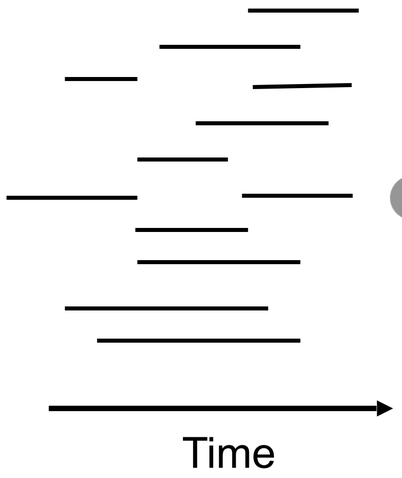


5% false positives

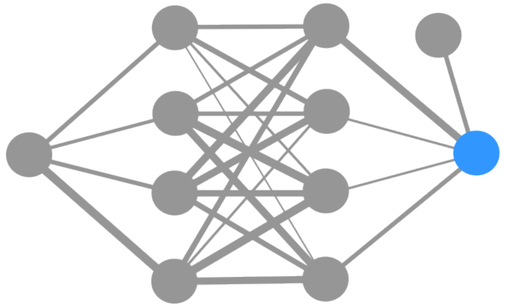
Should we dig deeper into factors? Threshold for variation in species-specific rates



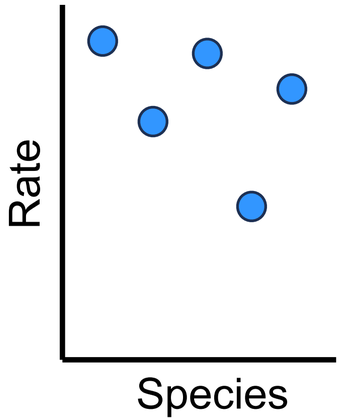
Simulate constant diversification and traits



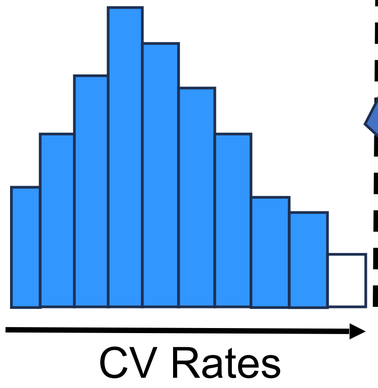
Infer BDNN

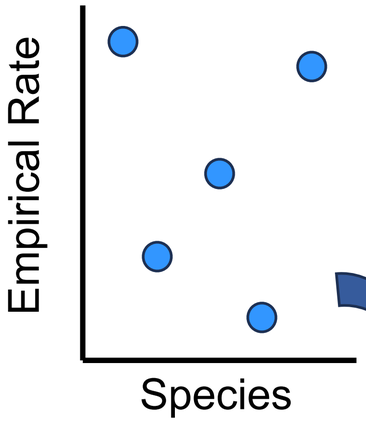


Variation in species-specific rates



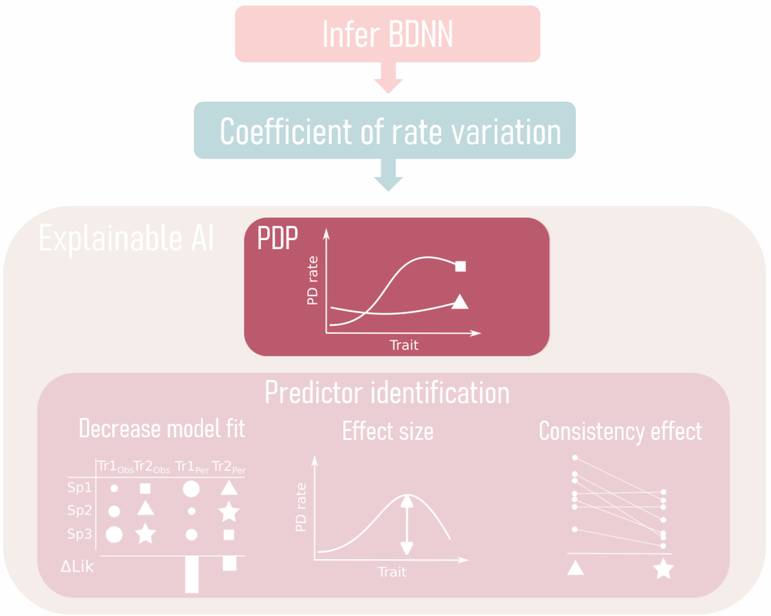
100x simulations to get threshold ?

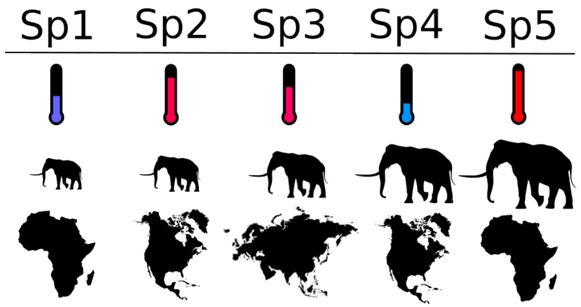


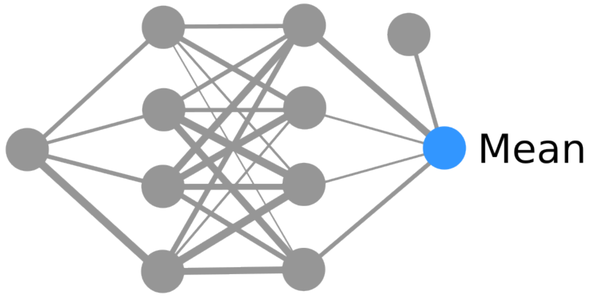


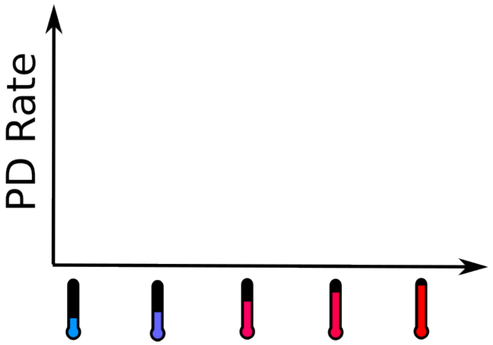
5% false positives

Visualize effect on rates

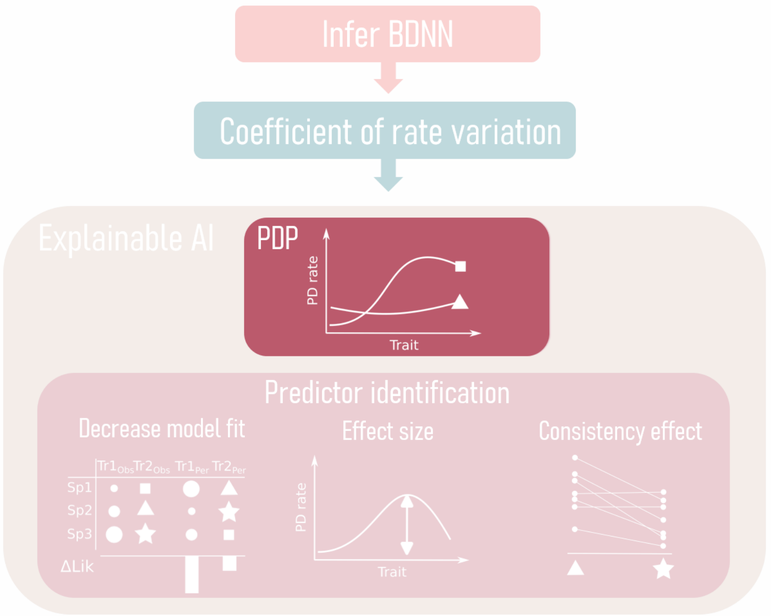


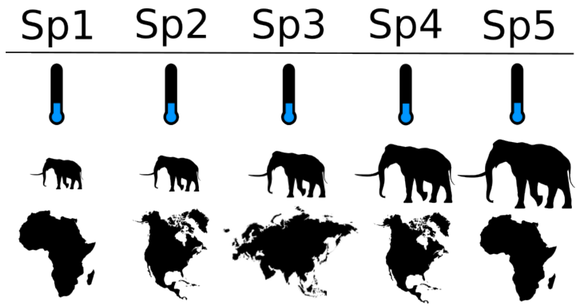


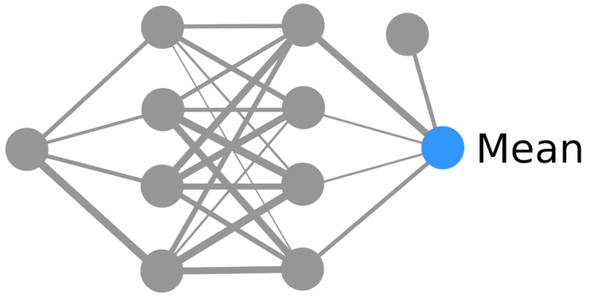


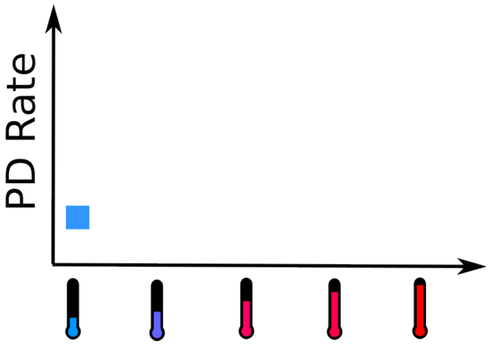


Visualize effect on rates

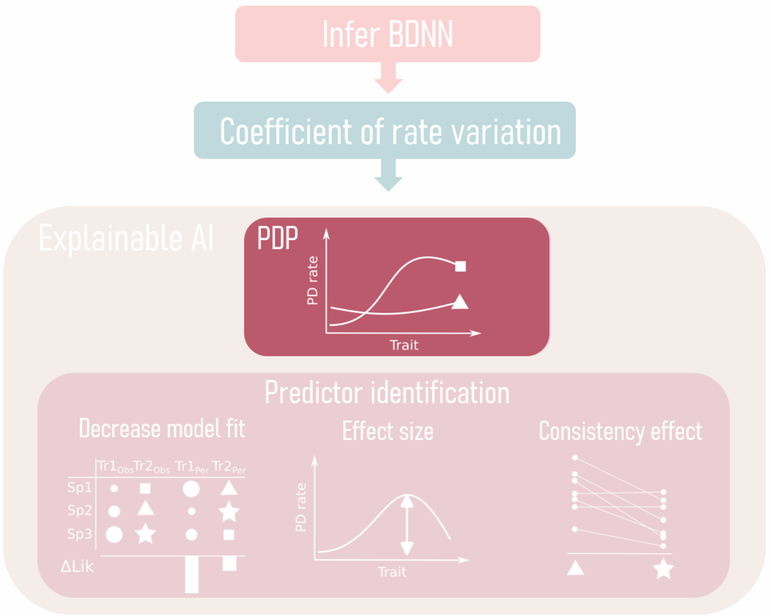


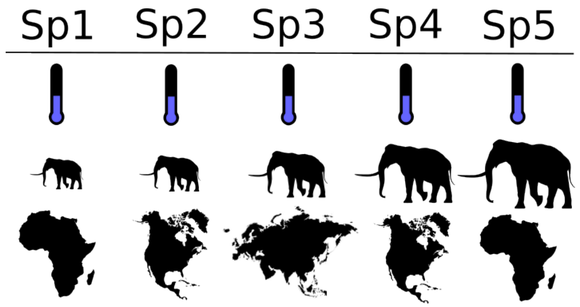


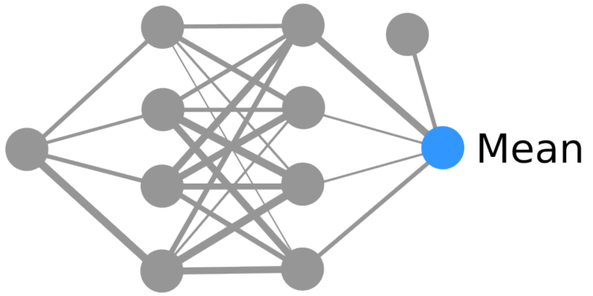


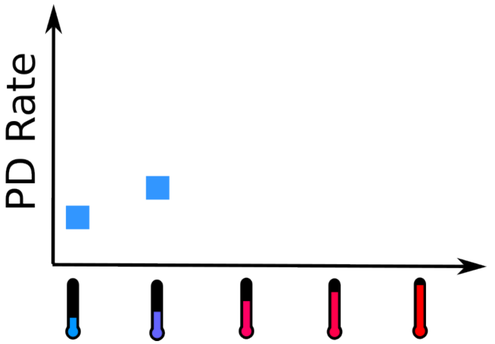


Visualize effect on rates

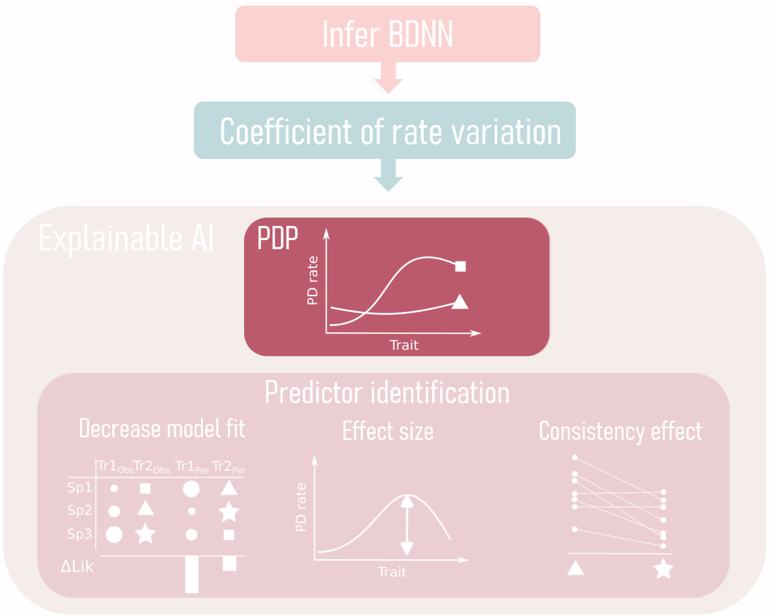


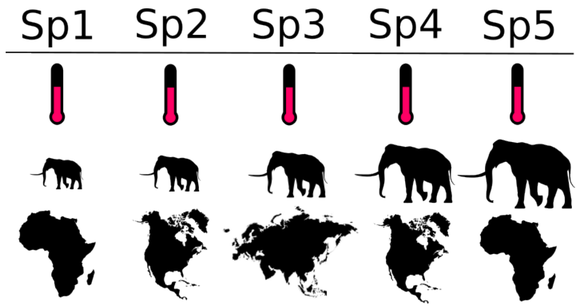


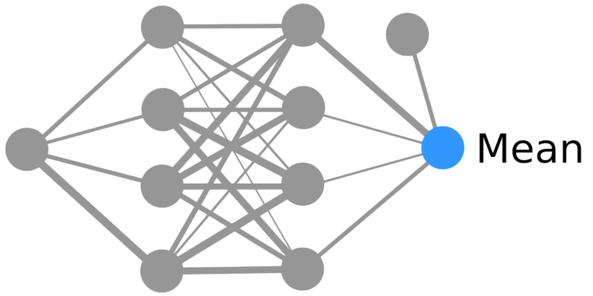


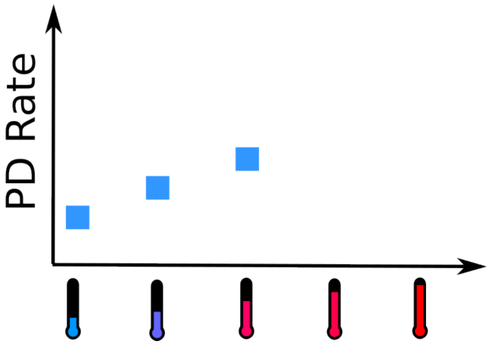


Visualize effect on rates

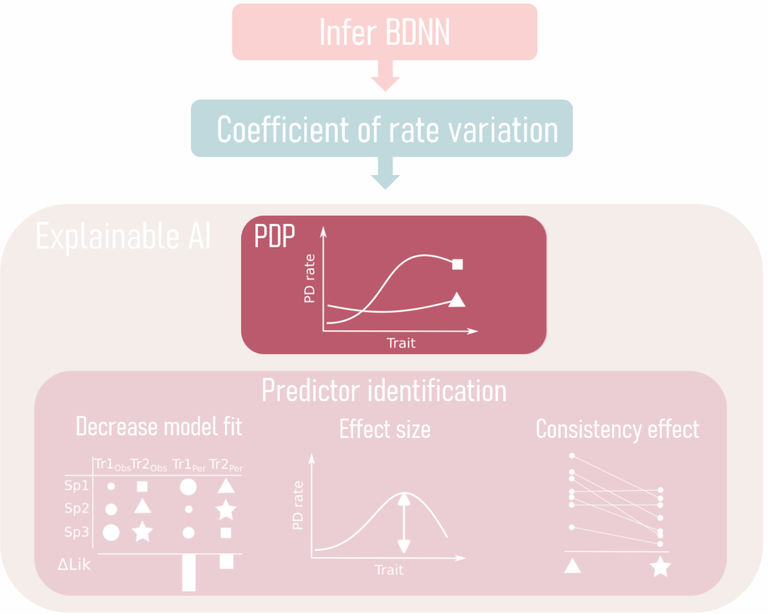


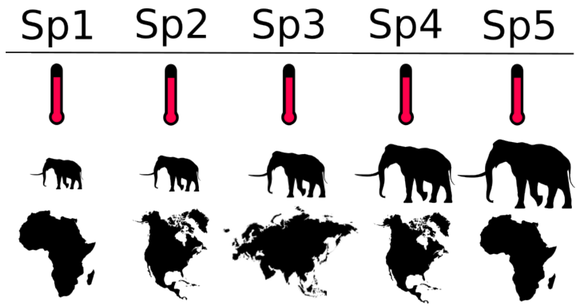


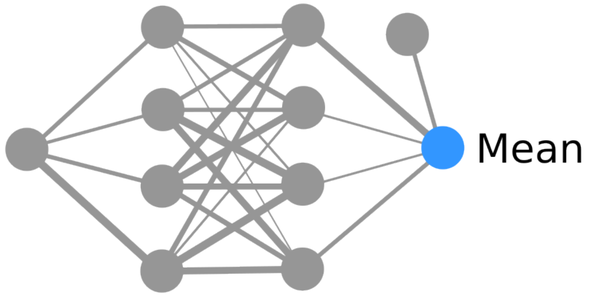


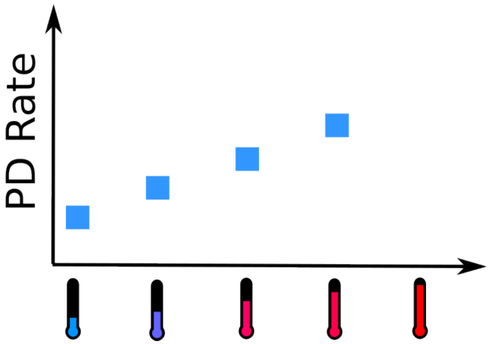


Visualize effect on rates

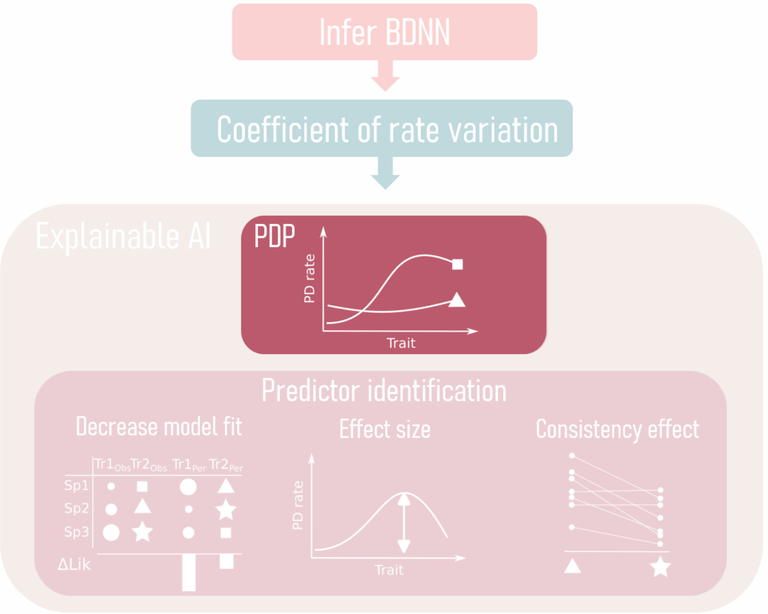


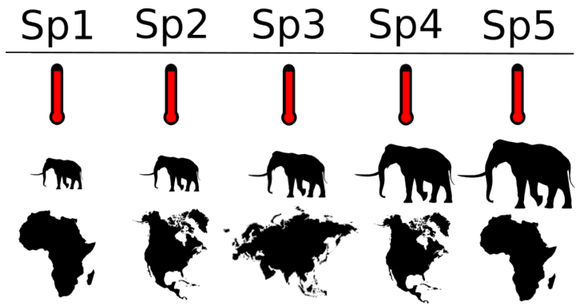


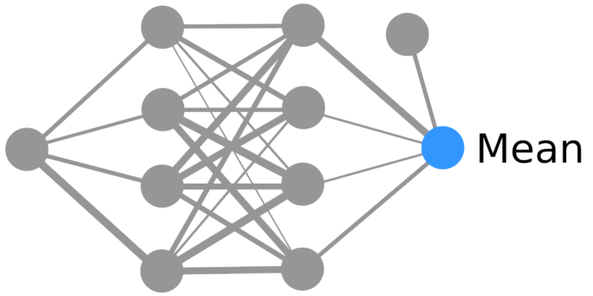


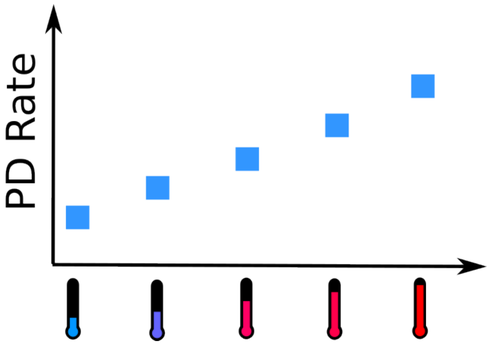


Visualize effect on rates

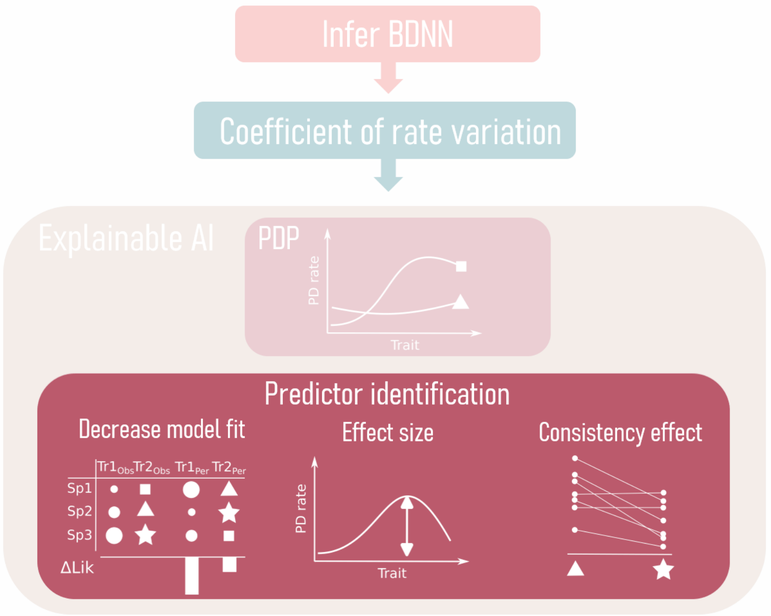


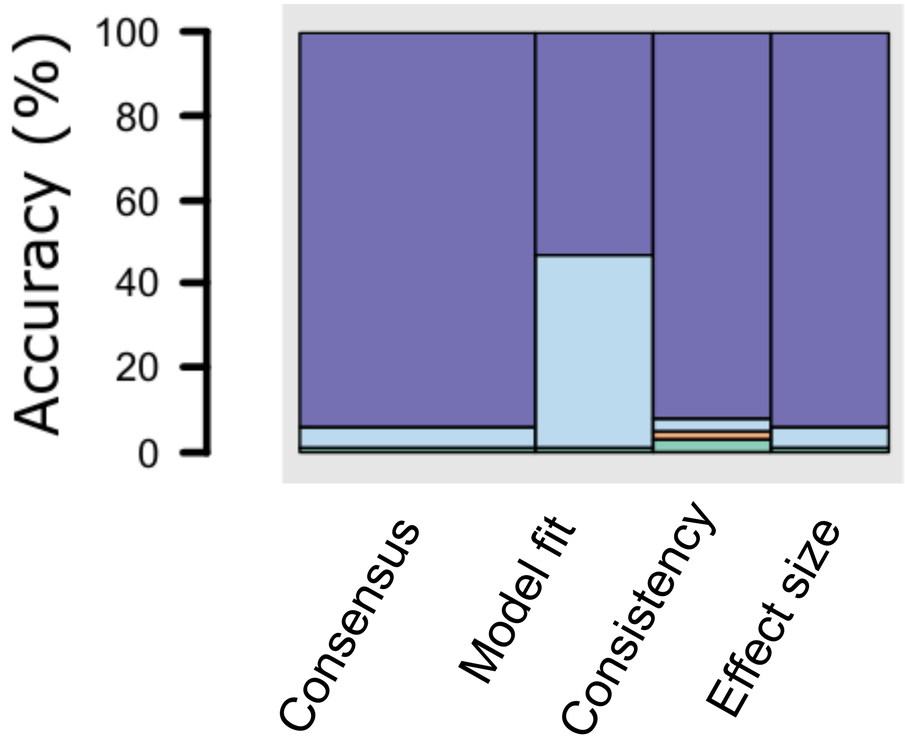




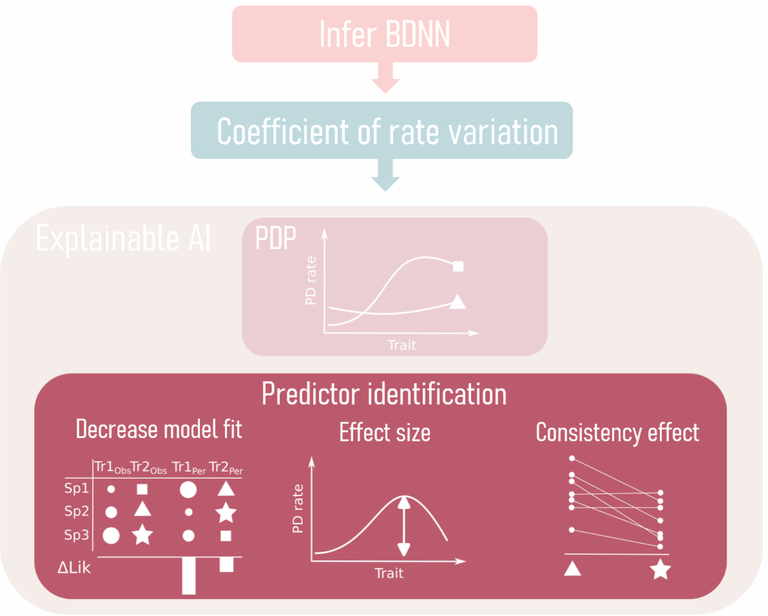


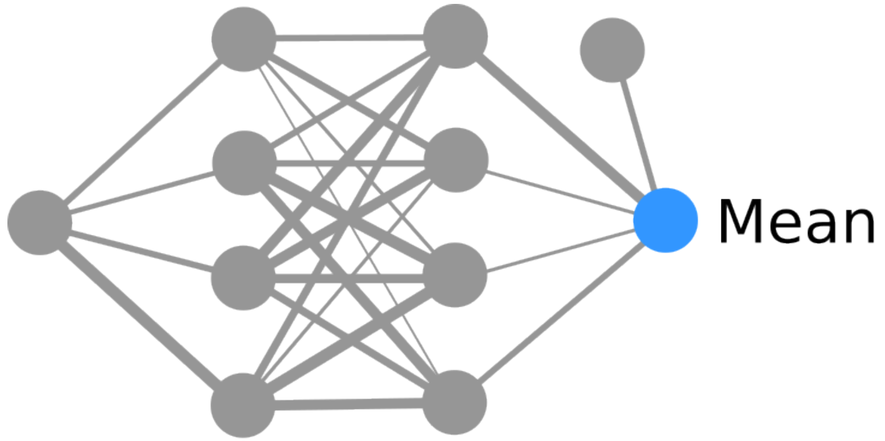
Identify rate predictors with consensus among xAI approaches

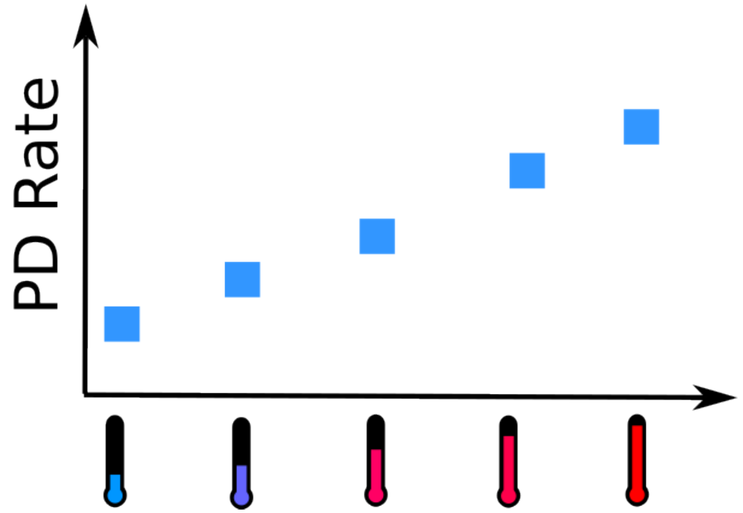




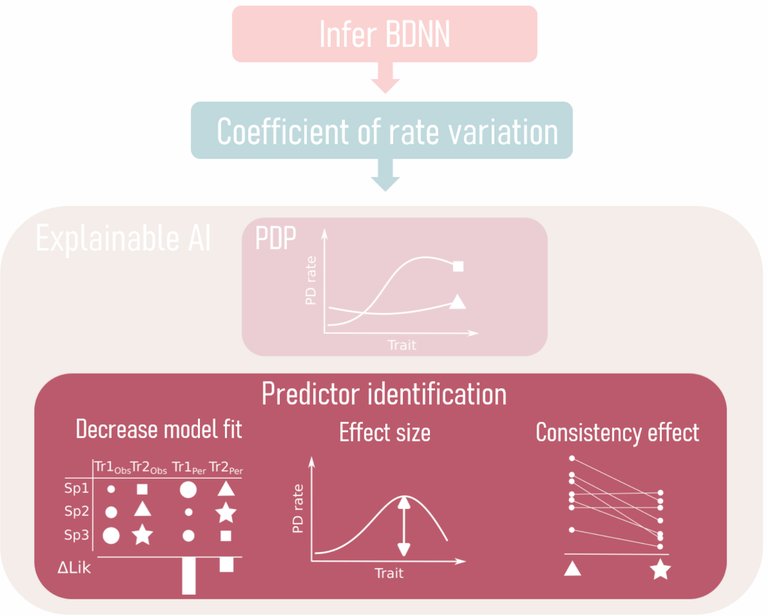
xAI: Consistency through posterior probability

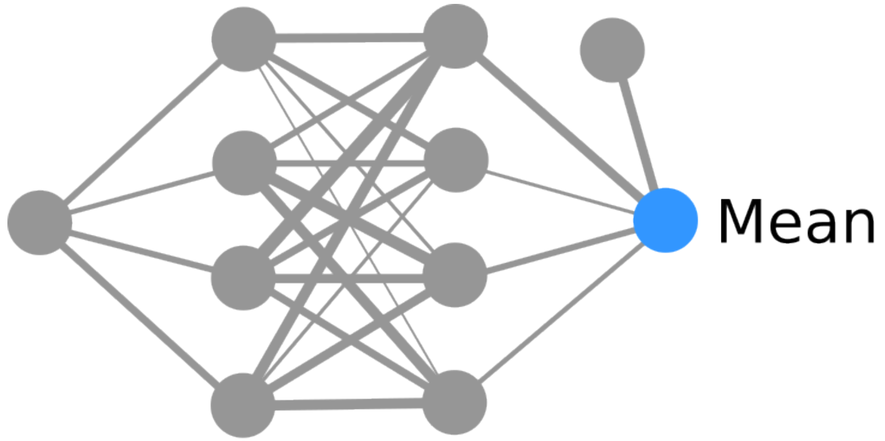


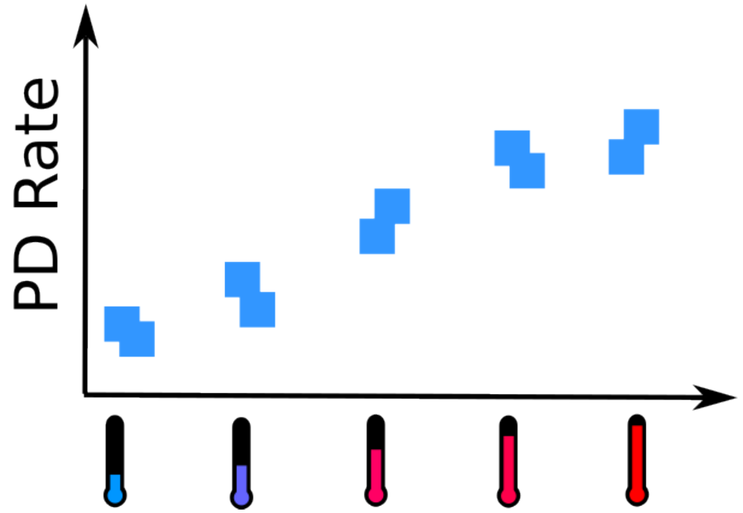




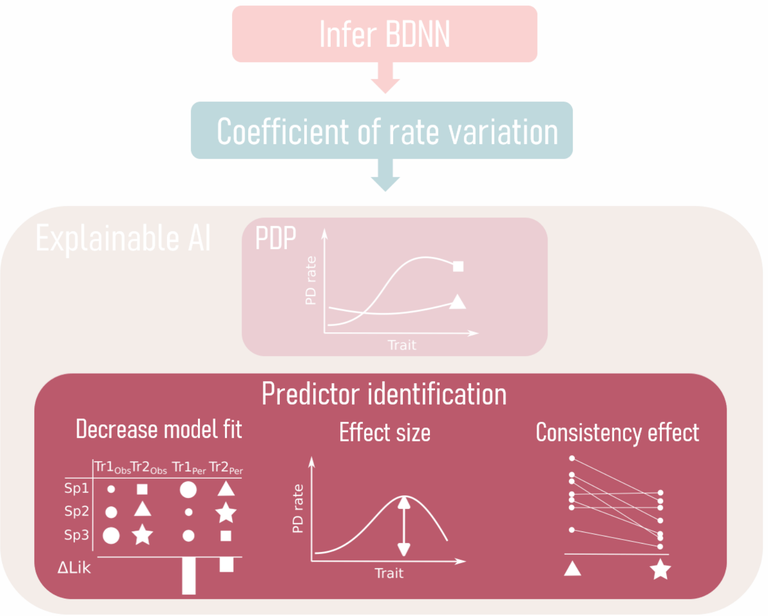
xAI: Consistency through posterior probability

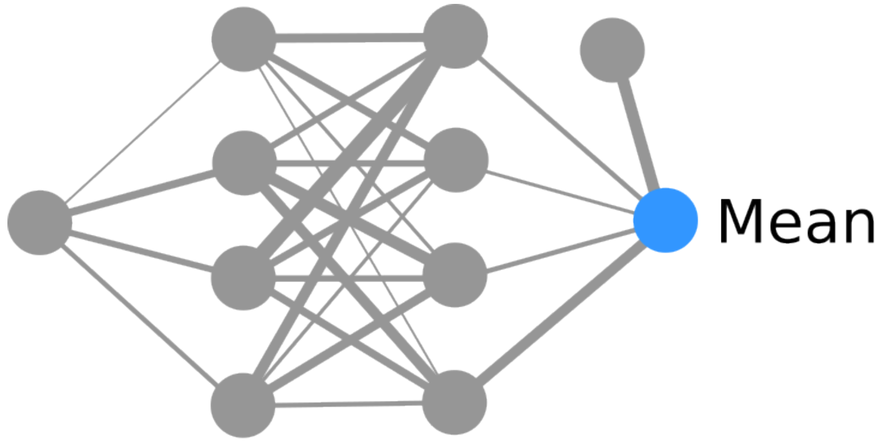


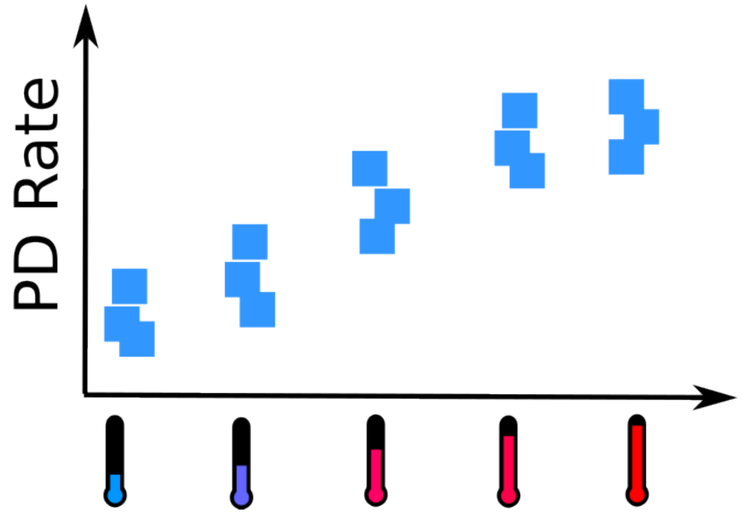




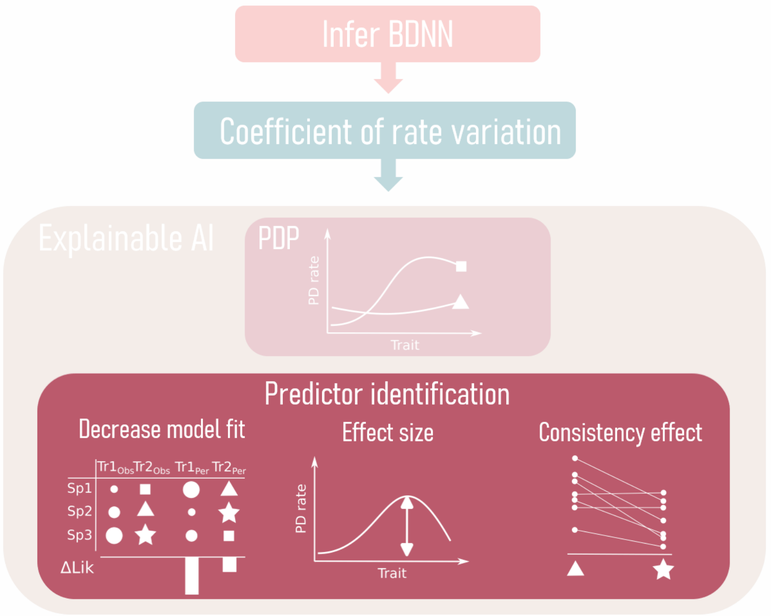
xAI: Consistency through posterior probability

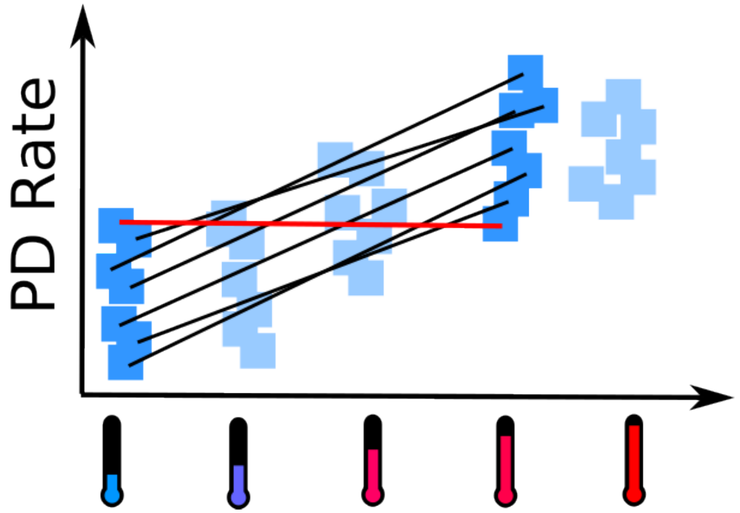






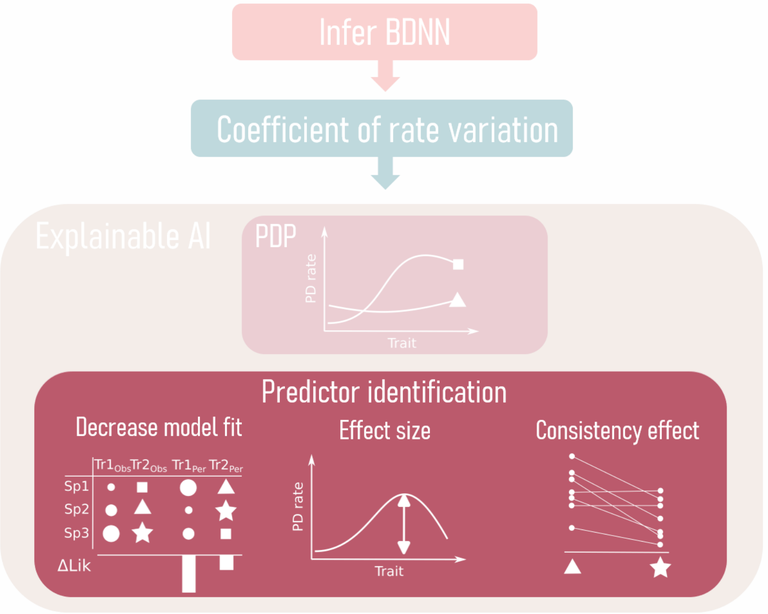
xAI: Consistency through posterior probability

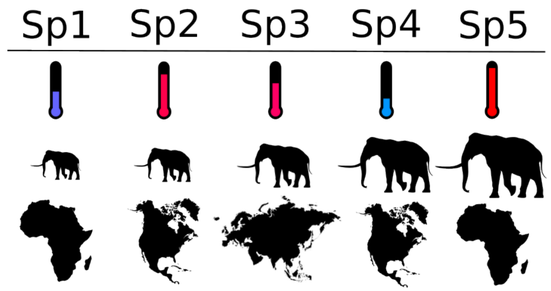


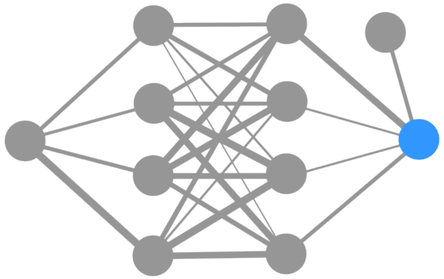


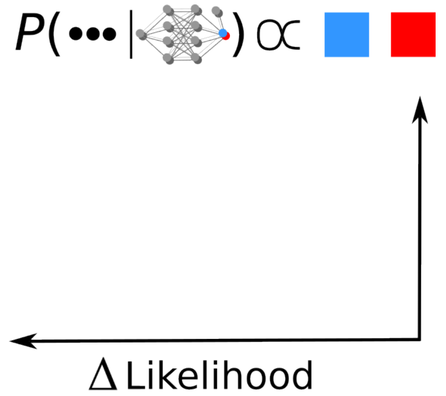
85 % of PD rates are greater than comparative temperature

xAI: Decrease in model likelihood when permuting features

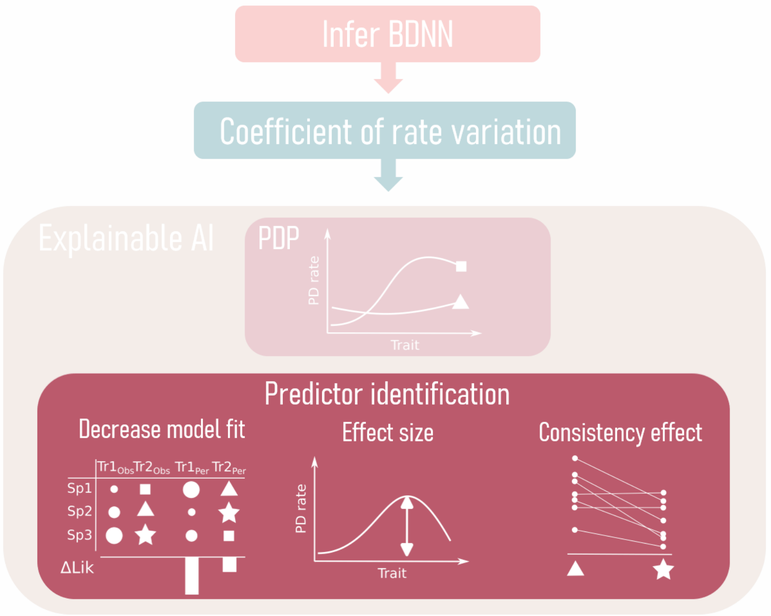


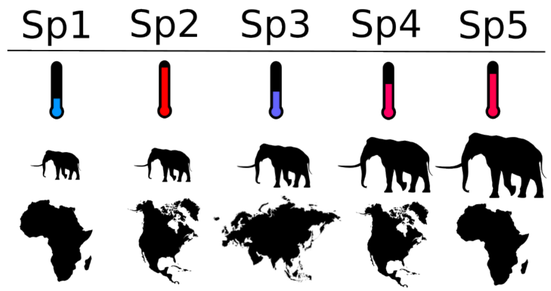


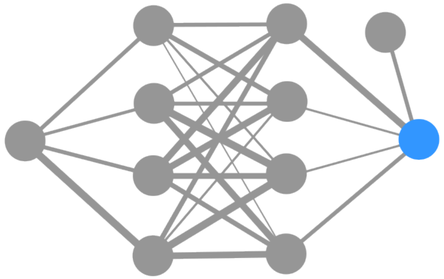


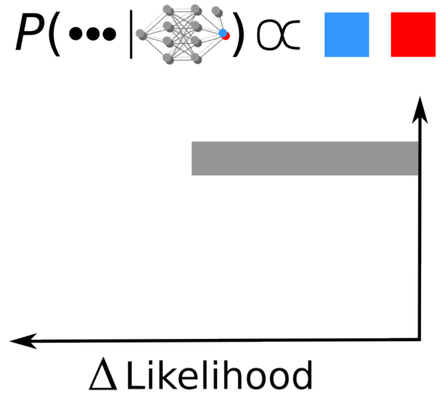


xAI: Decrease in model likelihood when permuting features

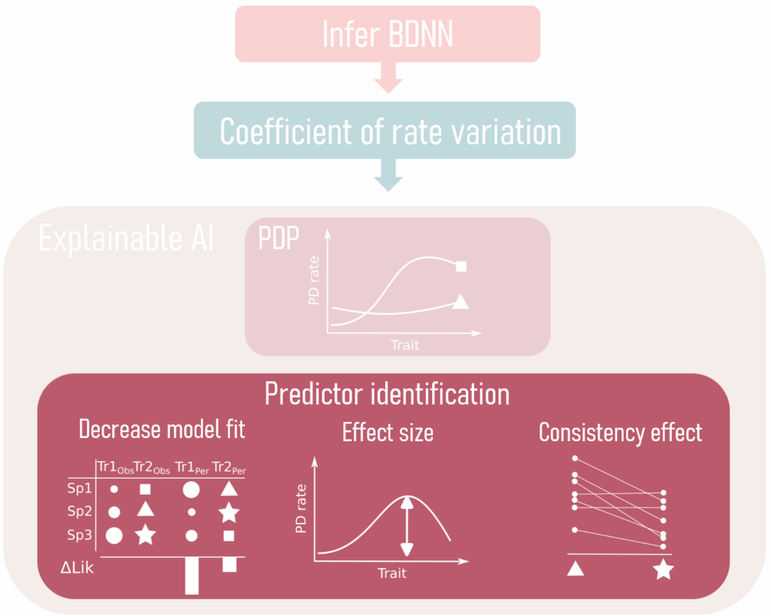


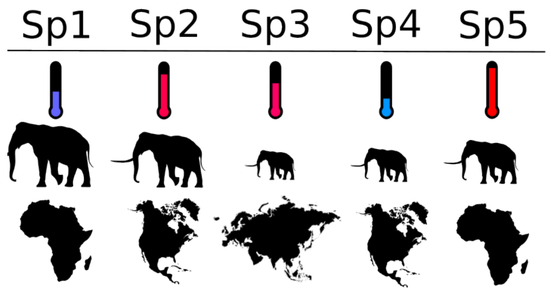


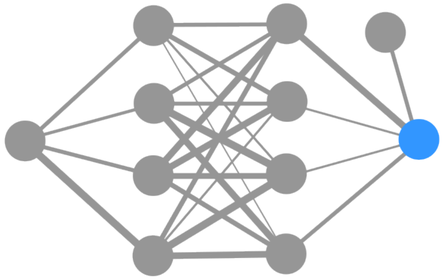


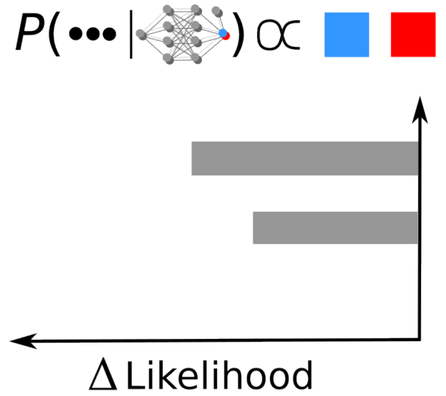


xAI: Decrease in model likelihood when permuting features

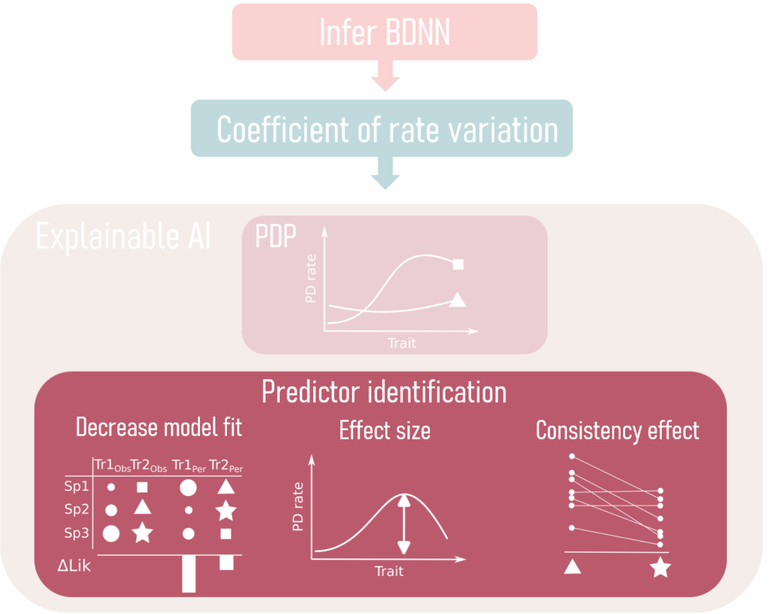


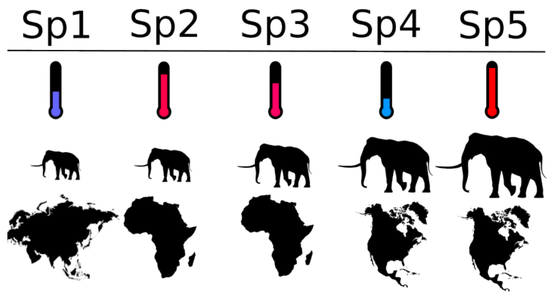


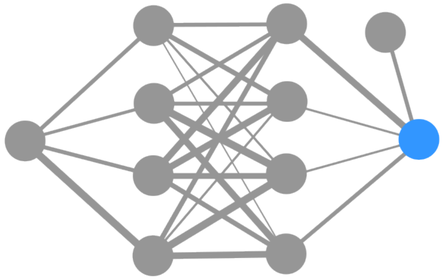


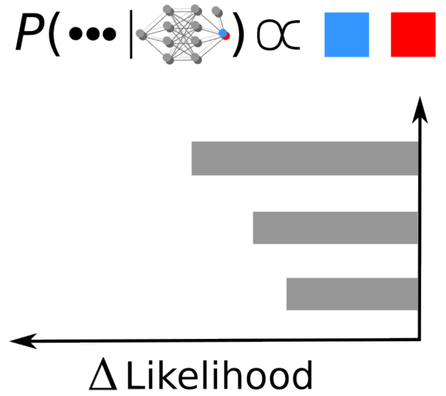


xAI: Decrease in model likelihood when permuting features

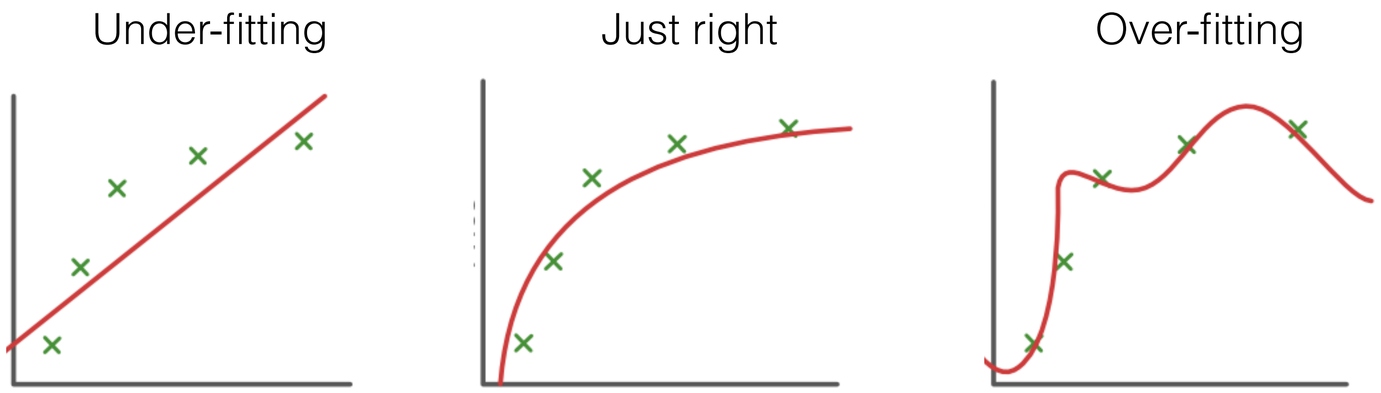




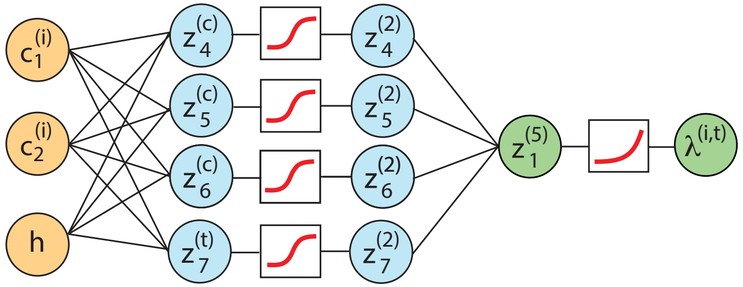


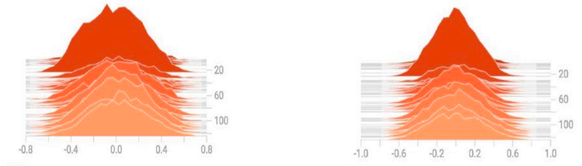


What about overfitting? 🤔



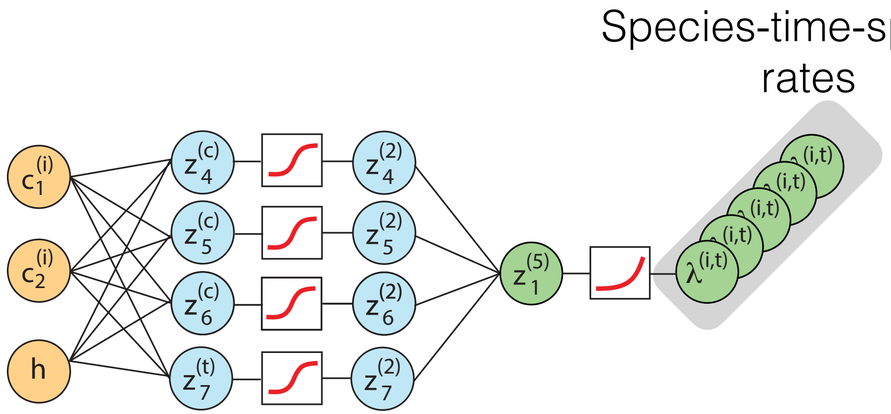
Model regularization #1 - priors on NN weights

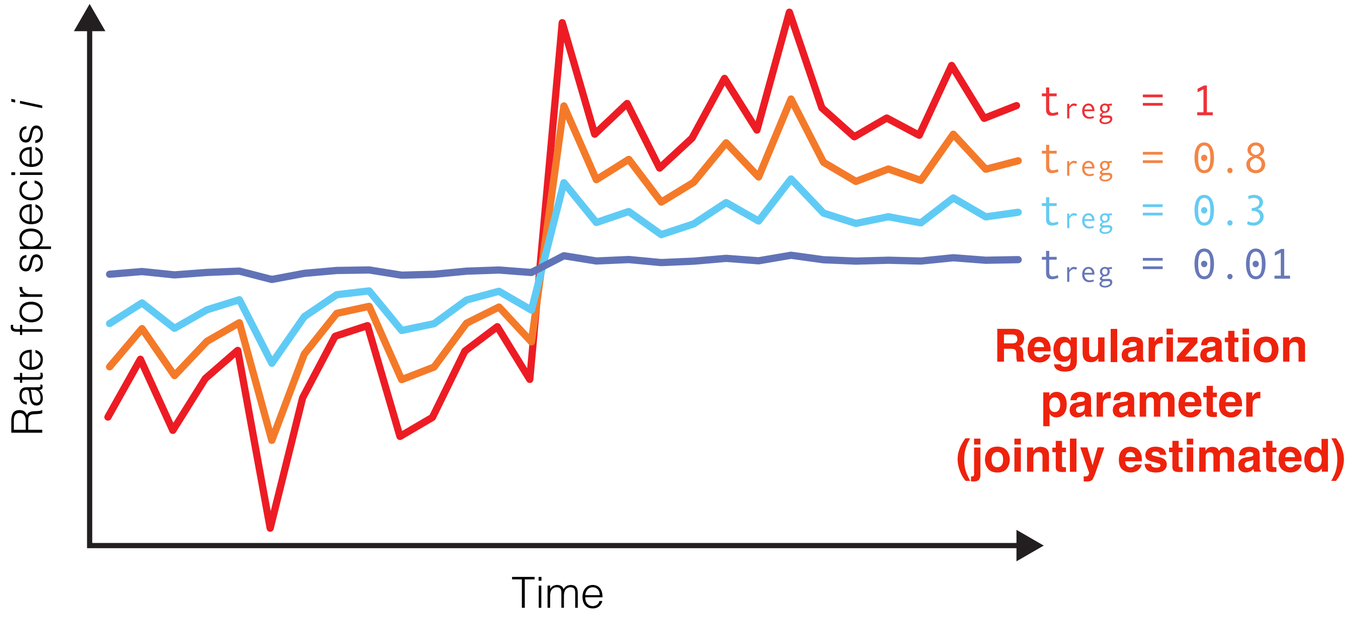




Silvestro & Andermann 2020 arXiv

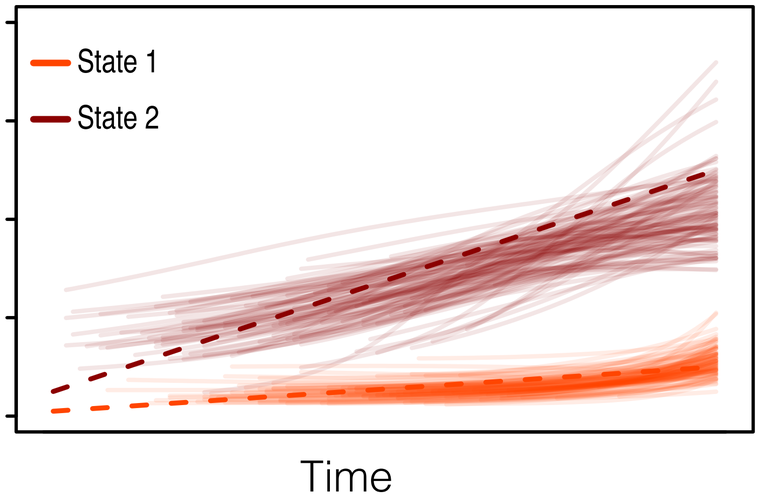
Model regularization #2 – Bayesian shrinkage



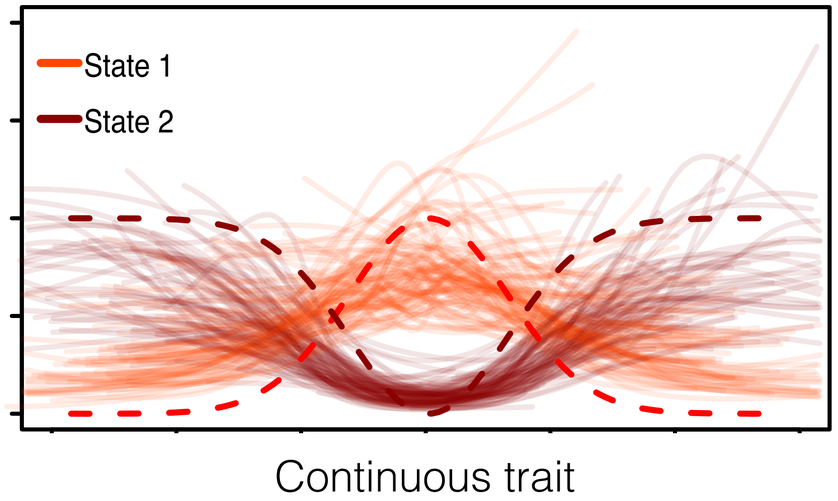


Time-dependent and trait-dependent BDNN model

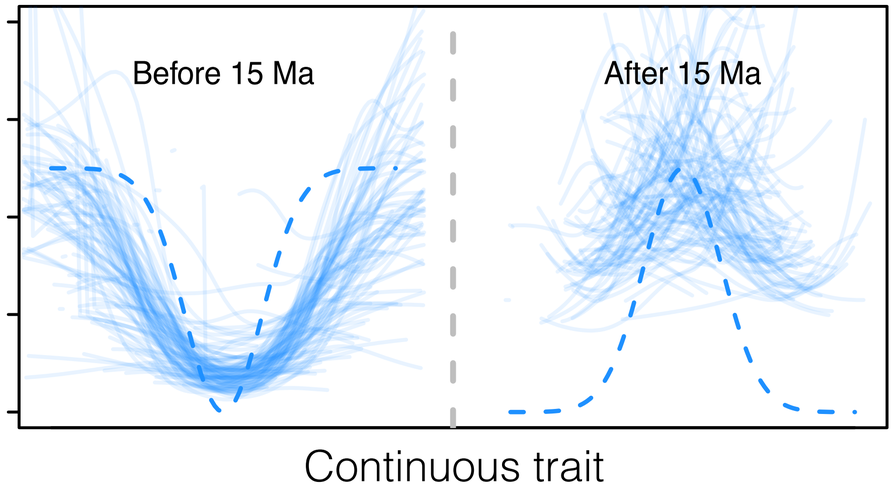
Time and trait dependent rates



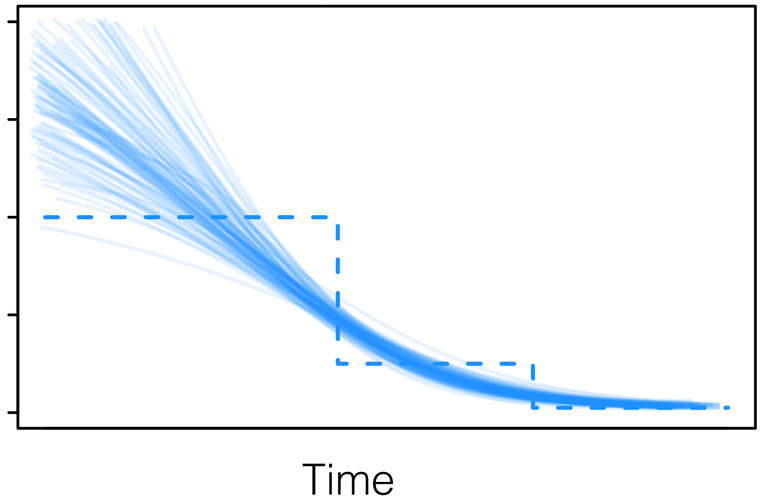
Categorical and continuous trait dependent rates



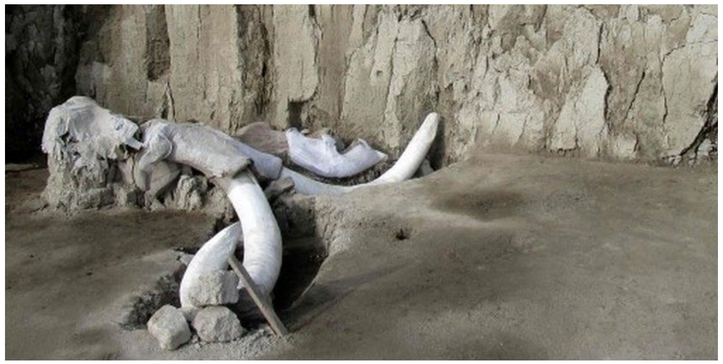
Time and trait dependent rates



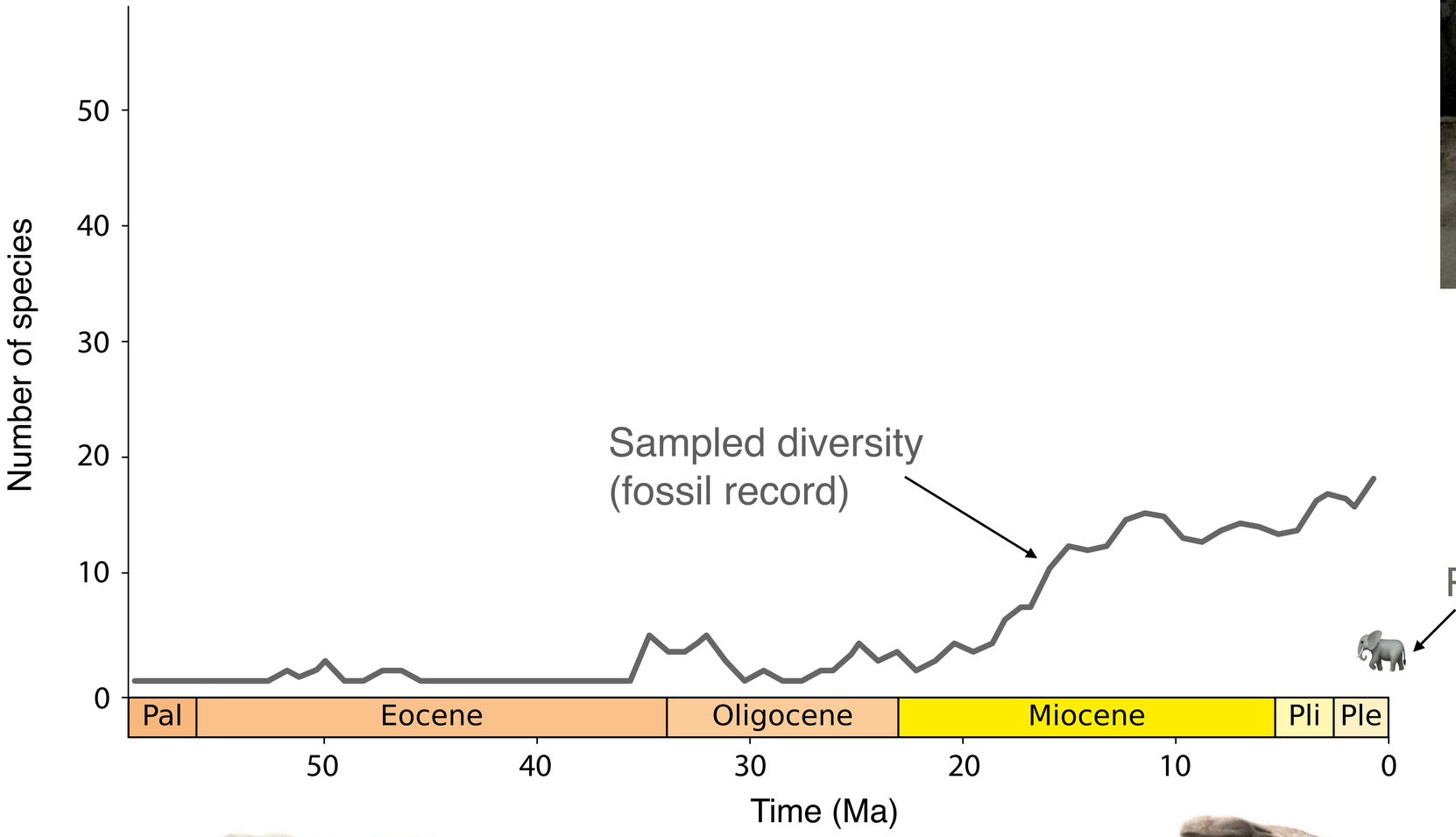
Time dependent rates



The rise and fall of elephants



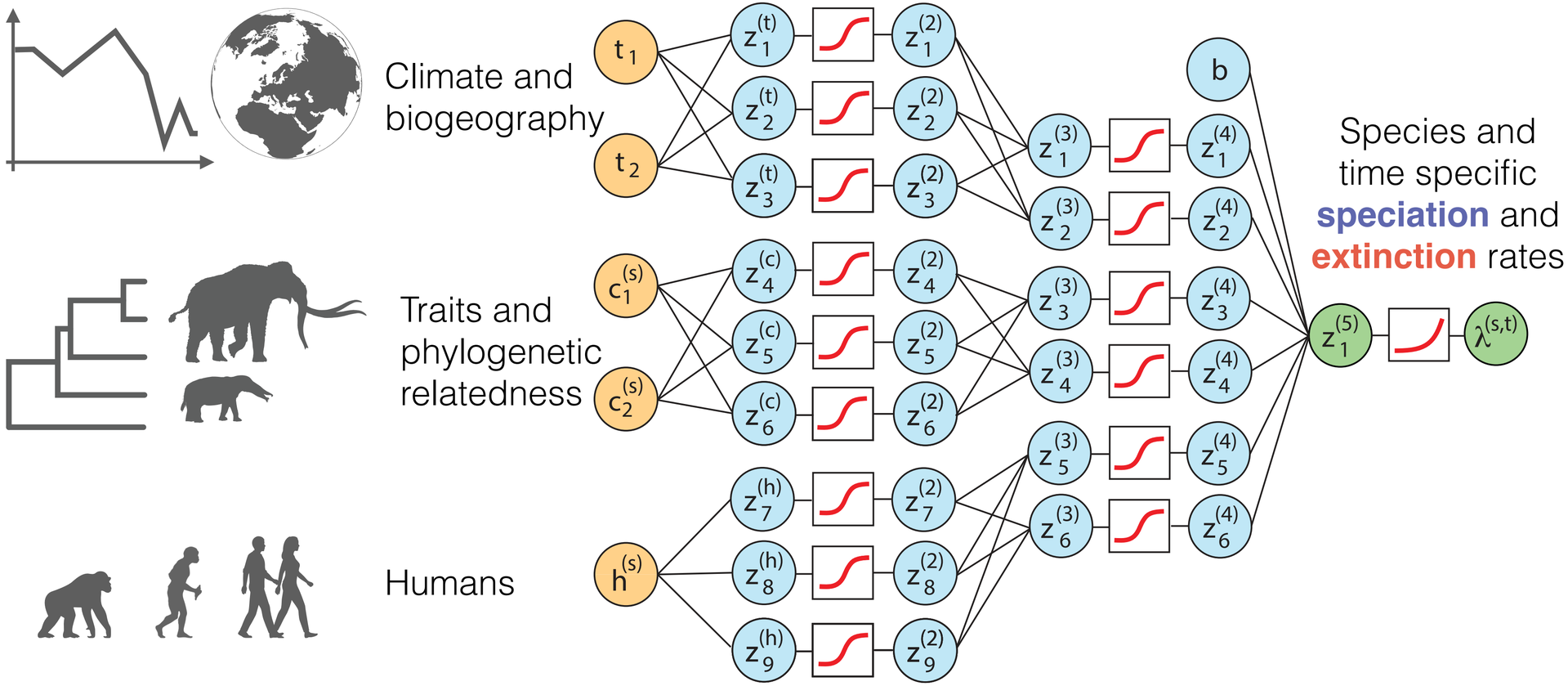
- 2,100 fossil occurrences
- 1,600 localities in 5 continents
- 180 species



Unsupervised Bayesian neural network model of speciation and extinction

Predictors

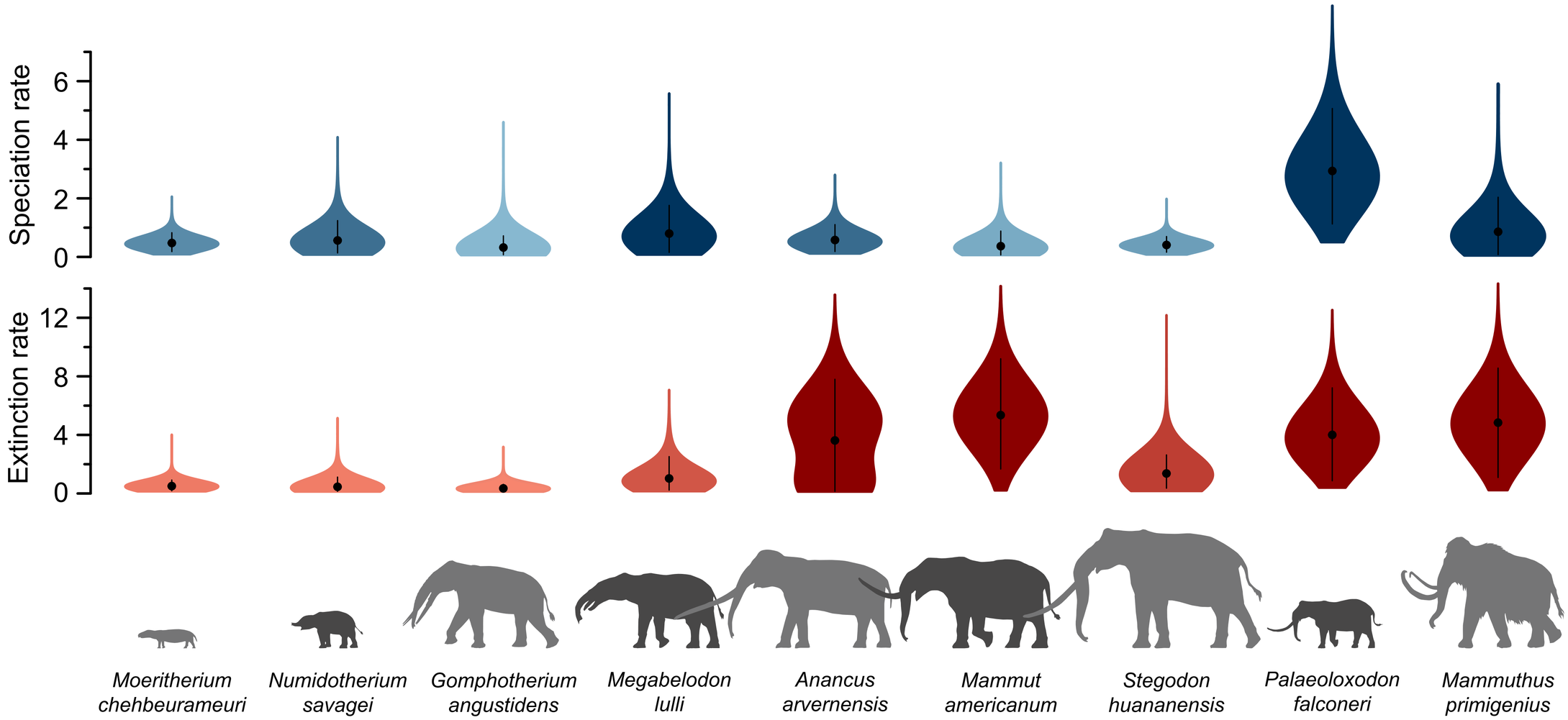
Resulting rates



github.com/dsilvestro/PyRate

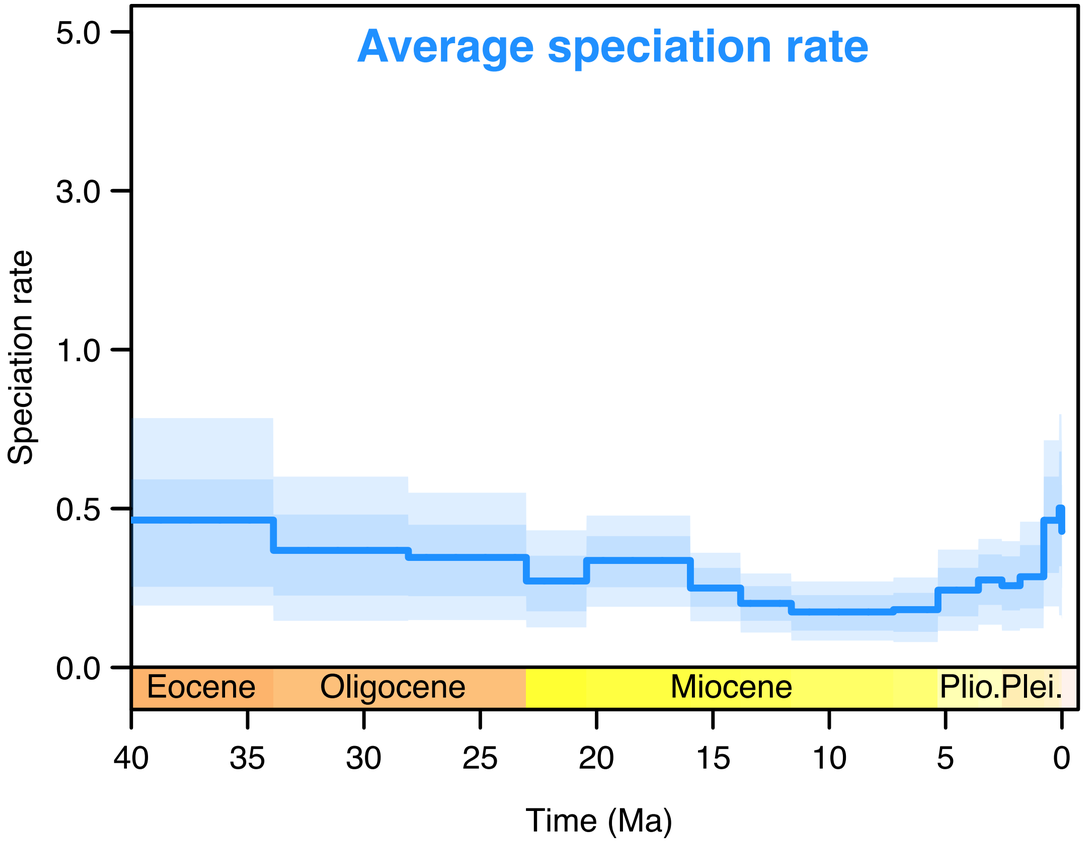
Hauffe, Cantalapiedra, Silvestro 2024 Science Adv

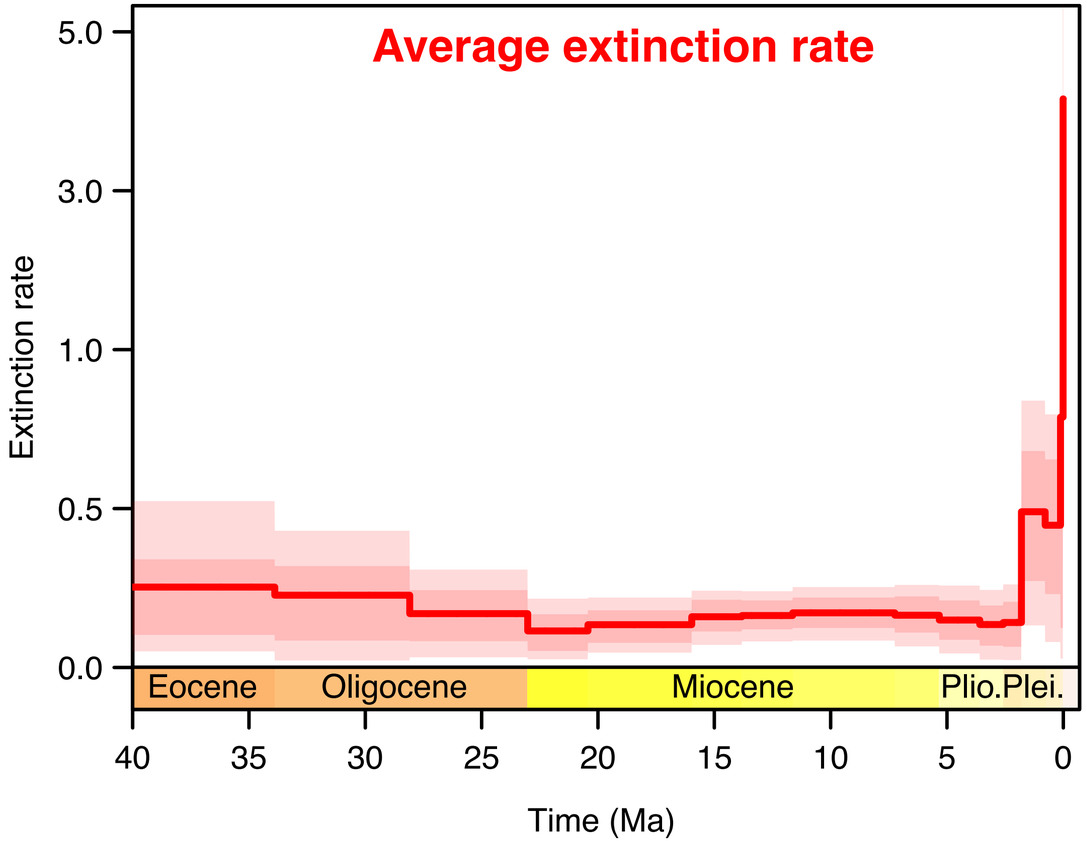
Marginal per-species speciation and extinction rates



Hauffe, Cantalapiedra, Silvestro 2024 Science Adv

Average speciation and extinction rates through time

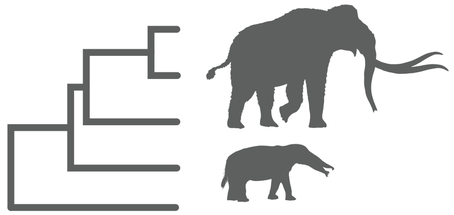




Unsupervised Bayesian neural network model of speciation and extinction

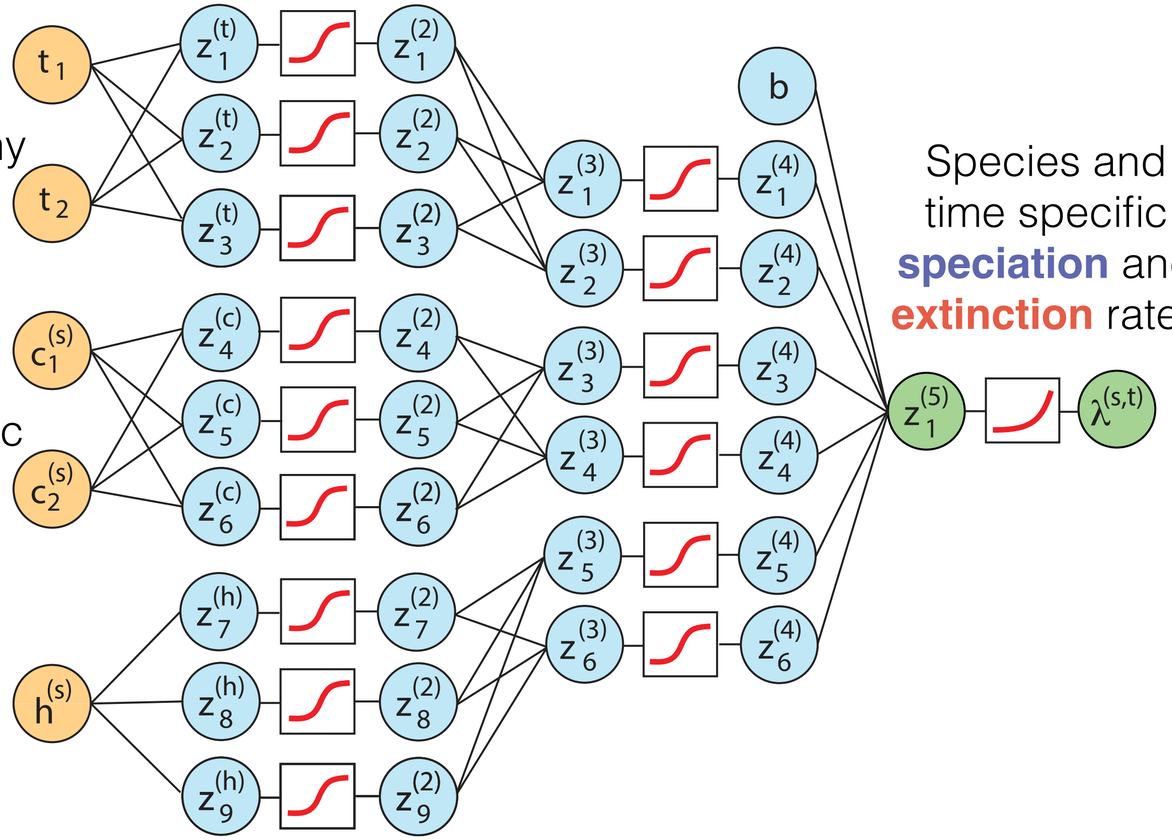
Predictors

Climate and biogeography



Traits and phylogenetic relatedness

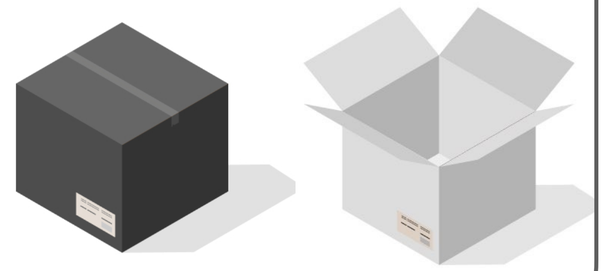
Humans



Resulting rates

Species and time specific **speciation** and **extinction** rates

eXplainable AI tools to infer the predictors of rate changes



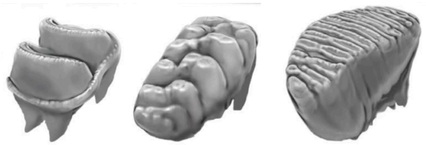
T Hauffe

github.com/dsilvestro/PyRate

Using tools from Explainable AI to determine the drivers of speciation and extinction

1. Permutations (Partial dependence plots)
2. Marginal probabilities
3. SHAP values (mean absolute deviation from the mean rate across all species caused by a predictor)

Ranking the predictors of speciation



Dietary specialization

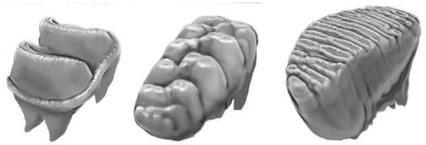
Biogeography

Time

Ranking the predictors of extinction

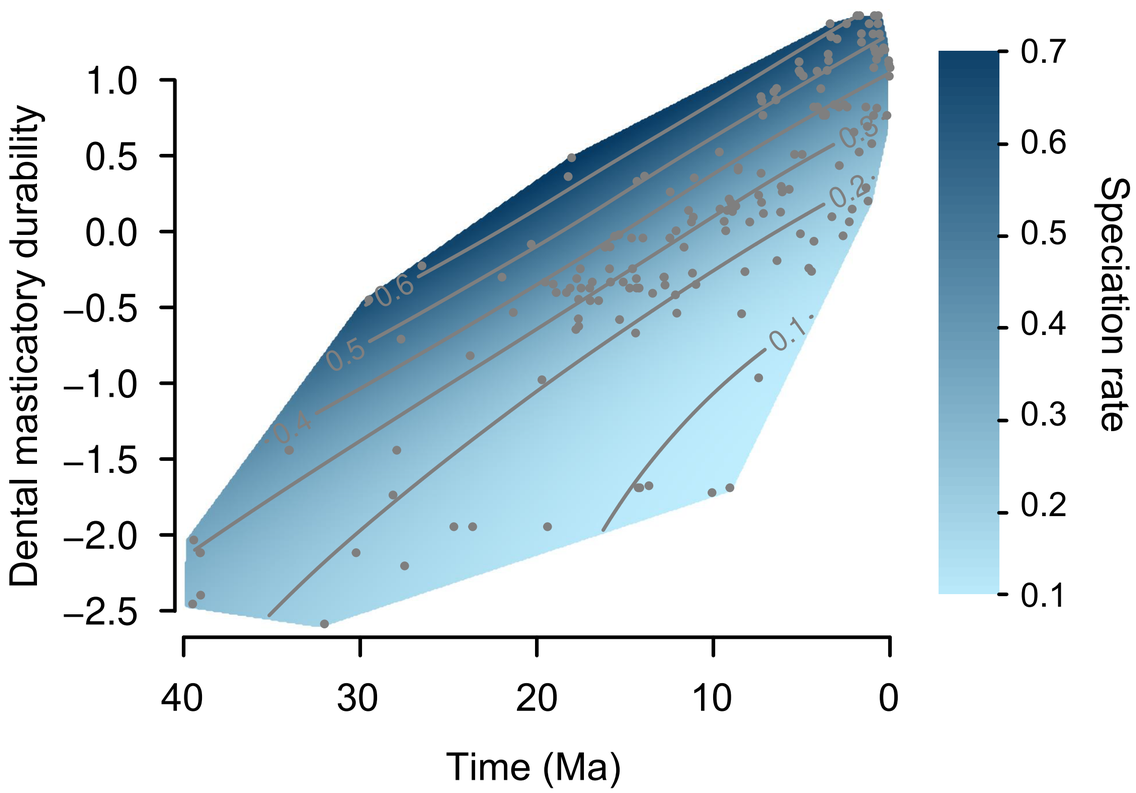
Human impact

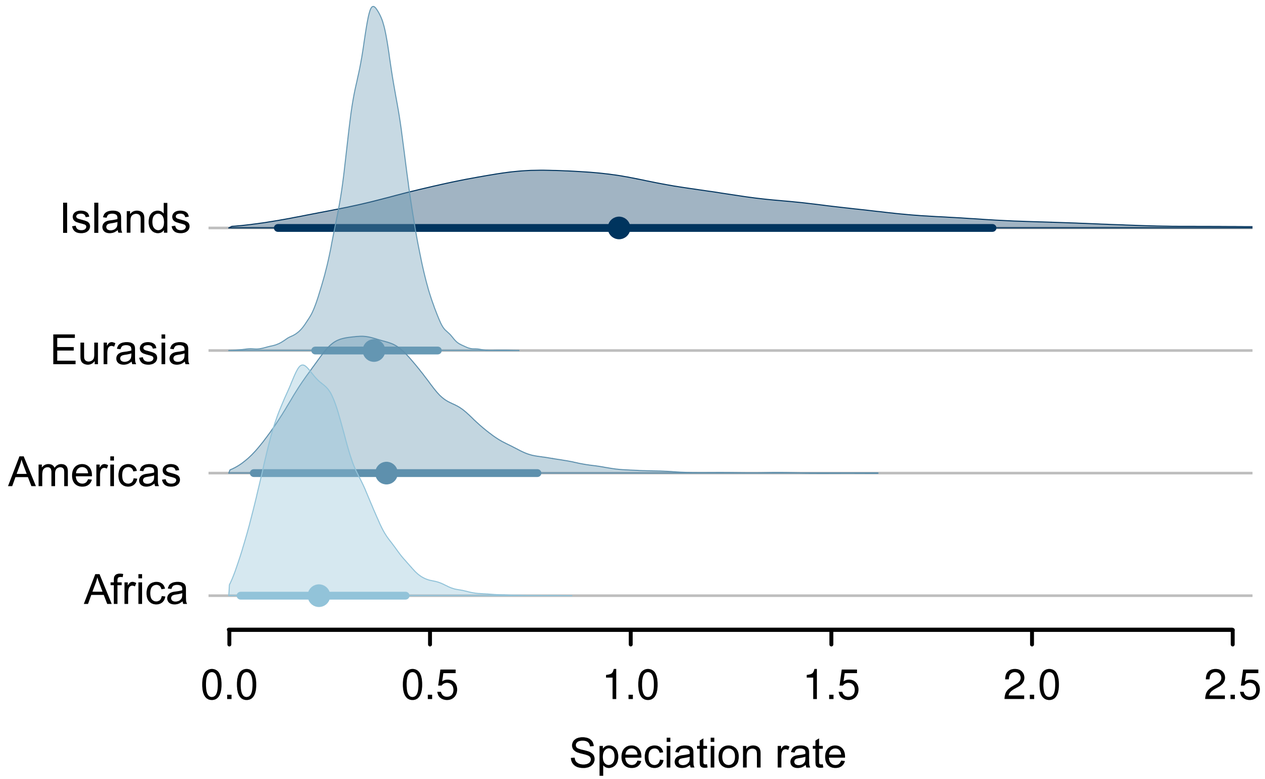
Biogeography



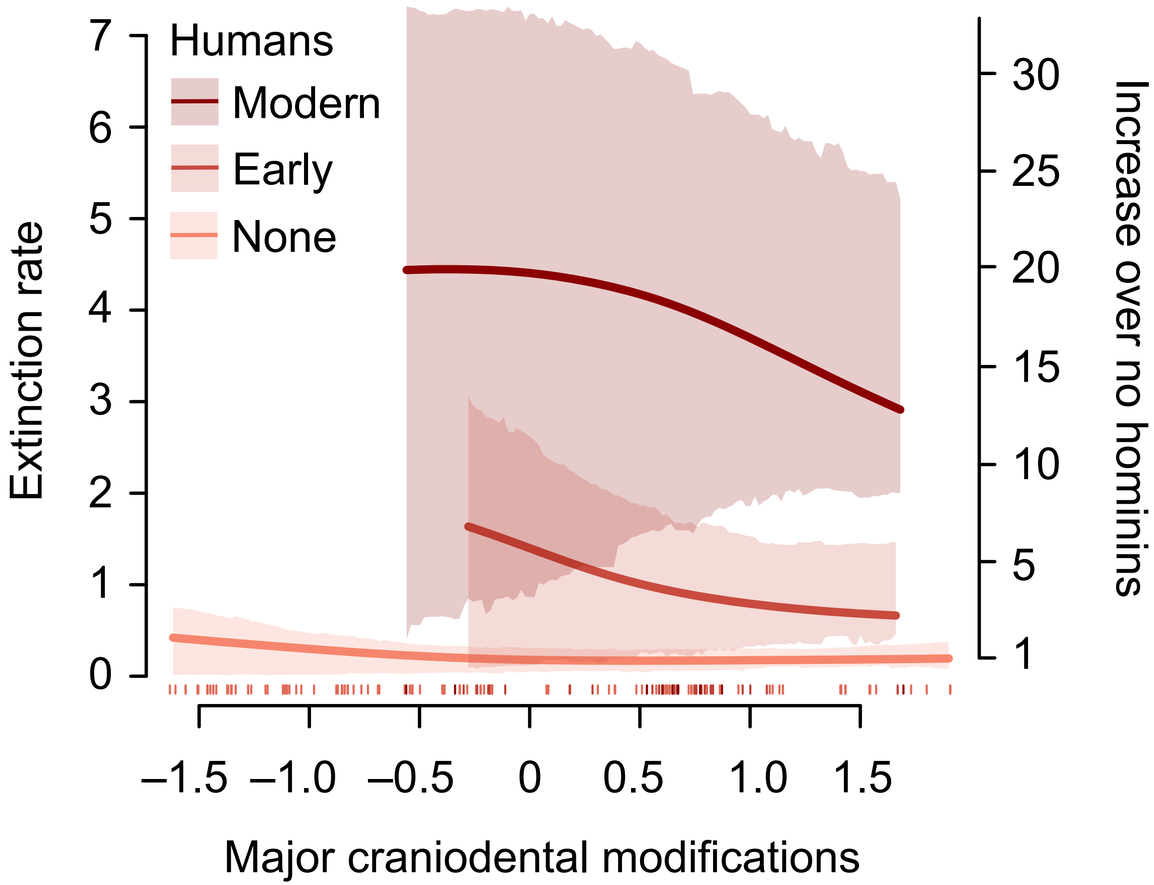
Ecomorphological trait

Speciation rates linked with ecomorphology and biogeography

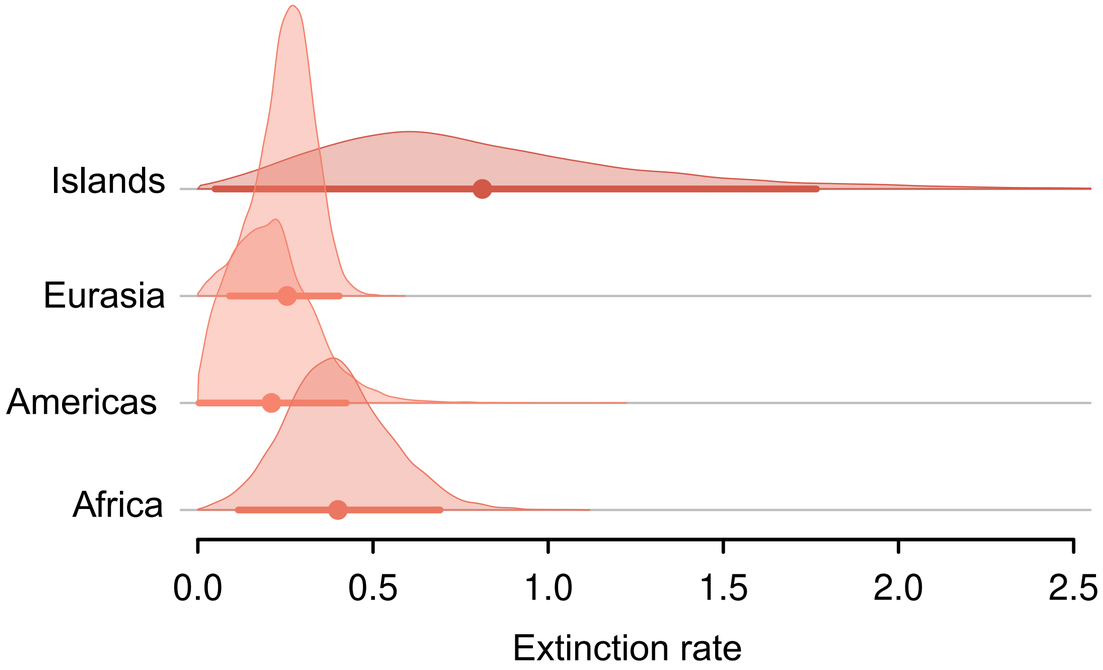


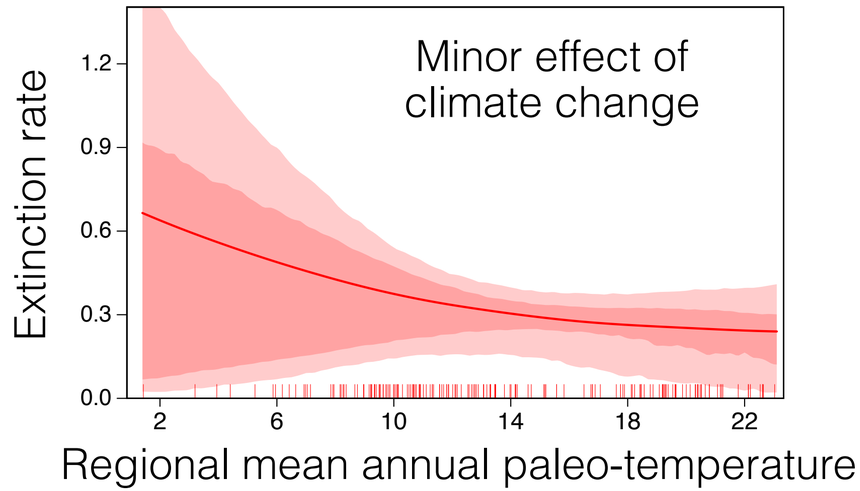


Extinction rates linked with overlap with humans and biogeography



15-20 fold increase in extinction rates linked with modern humans





The BDNN model is available on GitHub (with tutorials)

T Hauffe

The BDNN model

The birth-death neural-network model modulates speciation and extinction rates per lineage and through time as a function of

1. **time**
2. one or multiple **categorical and/or continuous traits** (e.g. diet, body mass, geographic range)
3. one or more **time-dependent variables** (e.g. paleotemperature)
4. **phylogenetic relatedness** (e.g. classification into higher taxa or phylogenetic eigenvectors)

As the function is based on a fully connected feed-forward neural network, it is not based on *a priori* assumptions about its shape. For instance, it can account for non-linear and non-monotonic responses of the rates to variation in the predictors.
It can also account for any interactions among the predictors.

The parameters of the BDNN model are estimated jointly with the origination and extinction times of all taxa and the preservation rates. The output can be used to estimate rate variation through time, across species, and to identify the most important predictors of such variation and their individual or combined effects.

Setting up a BDNN dataset

The BDNN model requires occurrence data in the standard PyRate format. It additionally can use species and time specific data. A table with species-specific trait data can be loaded in the analysis using the `-trait_table` command, while a table with time-series predictors can be loaded using the `-BDNNtimevar` command.

We provide an example dataset based on Hauffe et al 2022 MEE. This includes genus level occurrence data of northern hemisphere Cenozoic carnivores, a table with a paleotemperature time series, and a table with lineage-specific traits: log-transformed body mass, taxonomic information (family-level classification), and continent of

github.com/dsilvestro/PyRate

Is the BDNN a solution to all fossil diversification questions?

The BDNN is data-hungry! (Probably not suitable for poorly sampled fossil clades)

The BDNN is CPU intensive

Formal model testing provides a more direct (arguably) more powerful way to test hypotheses