

Hierarchical Estimation of Metabolic Index Parameters

Tim!

6/01/2022

Basic Model structure

Metabolic Index

The metabolic index is the ratio of metabolic demands to metabolic supply, both expressed as Arrhenius equations:

$$\phi = pO_2 A_o W^n \exp\left(\frac{E_o}{k_b}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right)$$

where T is temperature (degrees Kelvin), T_{ref} is a specified reference temperature (here chosen to be 15 degrees C), pO_2 is the partial pressure of oxygen (in kPa), W is body size, k_b is Boltzmann's constant, and A_o , n and E_o are estimated parameters. Thus, for any taxa, we define the vector \mathbf{g} as the true mean values of these three parameters for each species.

Parameter Estimation

To estimate these parameters hierarchically, we use the evolutionary random walk model used for other life history traits as described by Thorson (2020). The model tracks the mean values of any taxonomic grouping (the “child”) as a function of the mean values of the higher taxonomic group (the “parent”) in which the taxa is nested. For instance, the mean value for a particular Order will depend on the mean value for that Order's Class. The distribution of trait values within a parent taxonomic grouping is described by a multivariate normal distribution:

$$\mathbf{g} \sim MVN(\mathbf{x}_{p(g)}, \Sigma_{l(g)})$$

where $p(g)$ is the parent taxonomy level for child group g , and $\Sigma_{l(g)}$ is the evolutionary covariance, and $l(g)$ indicates the taxonomic level of group g . For reference $l(g) = 1$ refers to Class.

The estimation procedure works by estimating a vector of elements of the lower triangular matrix, \mathbf{L}_Σ , so the evolutionary covariance matrix equals:

$$\Sigma_{l(g)} = \lambda_l (\mathbf{L}_\Sigma \mathbf{L}_\Sigma^T + \mathbf{D})$$

where \mathbf{D} is a diagonal matrix of the additional and independent evolutionary variance in traits (σ_j^2). The parameter $\lambda_{l(g)}$ is the relative amount of covariance explained by taxonomic level l relative to taxonomic level Class (e.g. $\lambda_{l(g)=1} = 1$).

Currently, the diagonal elements of D are not explicitly being included in the estimation. If we assume that the critical values are those that make $\phi = 1$, we re-arrange the first equation and add random error, then for each experimental observation i :

$$-\log(pO_{2,crit,i}) = \log(A_{o,g[i]}) + n_{g[i]} \log(W_i) + E_{o,g[i]} T'_i + \epsilon_i$$

where

$$T'_i = \left(\frac{1}{k_b} \left(\frac{1}{T_i} - \frac{1}{T_{ref}} \right) \right)$$

and

$$e_i \sim N(0, \sigma)$$

The Data

Penn and Duetsch have compiled several experimental studies relating $pO_{2,crit}$ to temperature and oxygen to which Essington et al. (2022) added information for an additional four species. A brief summary of the number of experimental studies by taxonomic grouping is below

```
## Warning in checkMatrixPackageVersion(): Package version inconsistency detected.
## TMB was built with Matrix version 1.4.1
## Current Matrix version is 1.4.0
## Please re-install 'TMB' from source using install.packages('TMB', type = 'source') or ask CRAN for a
##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##      combine

## # A tibble: 11 x 2
##   Class      NoOrder
##   <chr>      <int>
## 1 Actinopteri      14
## 2 Ascidiacea       1
## 3 Bivalvia         3
## 4 Cephalopoda      3
## 5 Elasmobranchii   1
## 6 Gastropoda       1
## 7 Hexanauplia      2
## 8 Malacostraca     3
## 9 Ophiuroidea      1
## 10 Polychaeta      1
## 11 Thaliacea       1

## # A tibble: 31 x 2
##   Order      NoFamily
##   <chr>      <int>
## 1 Acipenseriformes      1
## 2 Amphipoda             1
## 3 Anguilliformes        1
## 4 Atheriniformes        1
## 5 Blenniiformes         1
## 6 Calanoida             3
## 7 Carcharhiniformes     1
## 8 Cypriniformes         1
## 9 Decapoda             13
## 10 Eupercaria incertae sedis  4
## 11 Euphausiacea          1
## 12 Gadiformes           1
## 13 Gobiiformes           1
## 14 Harpacticoida        1
```

## 15	Littorinimorpha	1
## 16	Myctophiformes	1
## 17	Myopsida	1
## 18	Mytilida	1
## 19	Octopoda	1
## 20	Oegopsida	1
## 21	Ophiacanthida	2
## 22	Ostreida	1
## 23	Ovalentaria incertae sedis	1
## 24	Pectinida	1
## 25	Perciformes	4
## 26	Phyllodocida	1
## 27	Pleuronectiformes	2
## 28	Salmoniformes	1
## 29	Salpida	1
## 30	Stolidobranchia	1
## 31	Tetraodontiformes	1

A tibble: 53 x 2

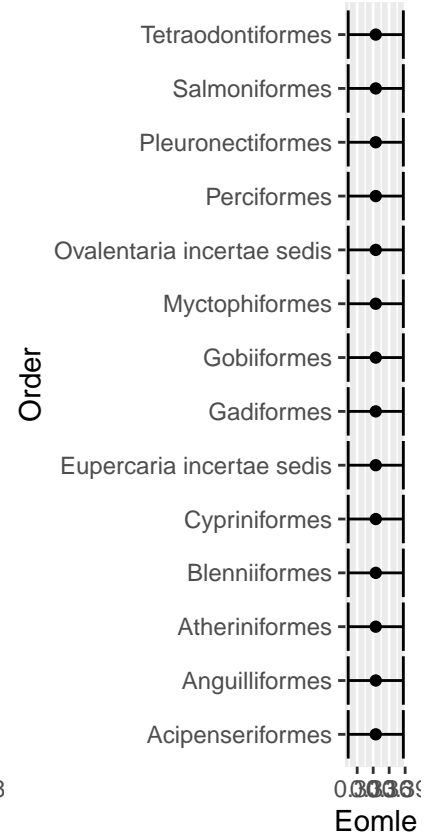
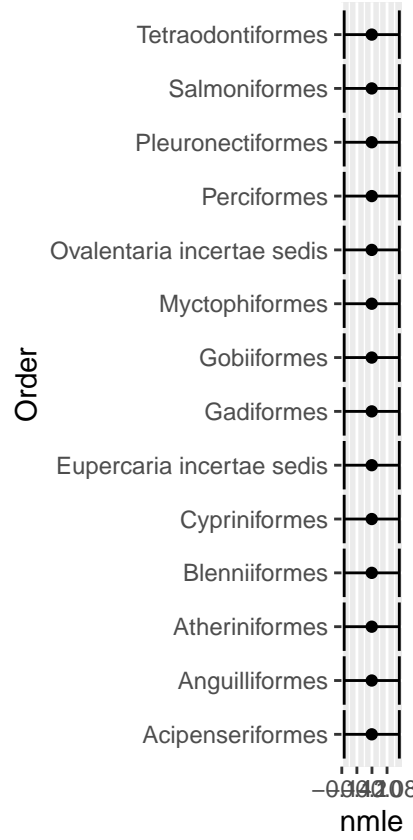
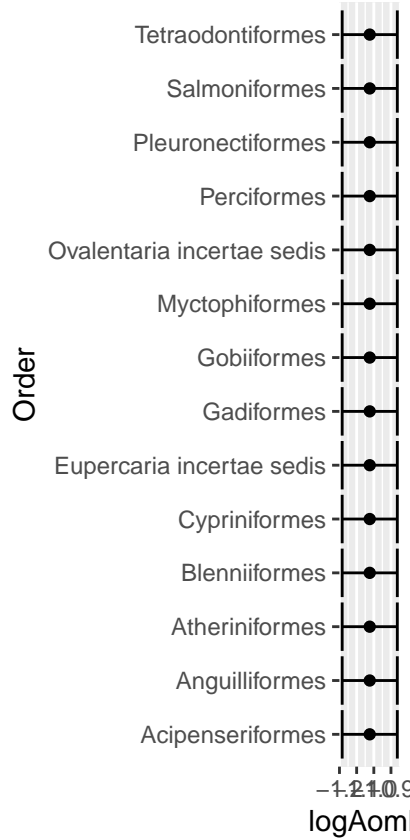
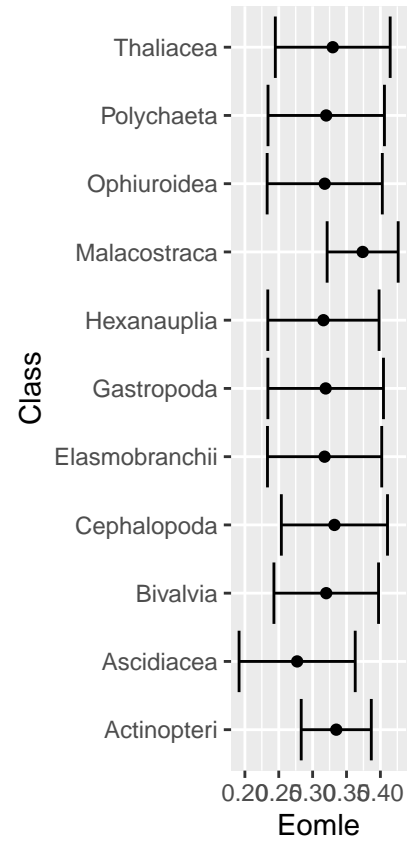
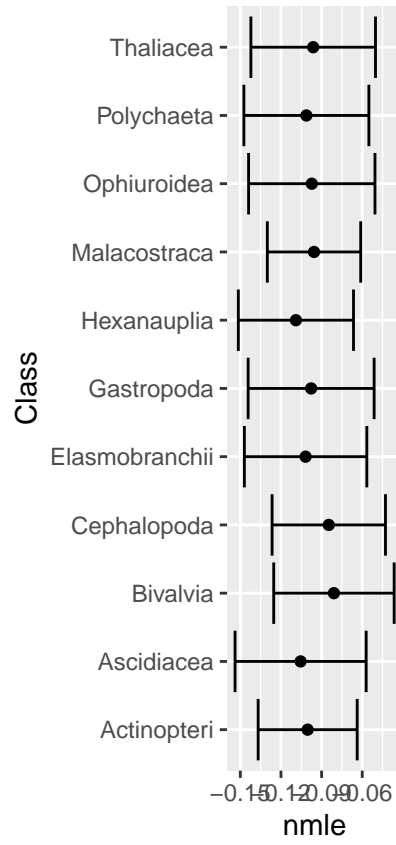
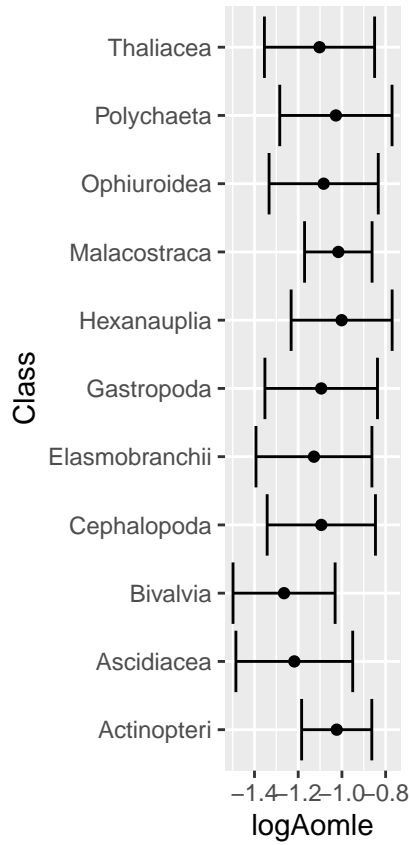
##	Family	NoGenus
##	<chr>	<int>
## 1	Acanthephyridae	2
## 2	Acipenseridae	1
## 3	Anguillidae	1
## 4	Anoplopomatidae	1
## 5	Atherinopsidae	1
## 6	Benthesicymidae	1
## 7	Bythograeidae	1
## 8	Cancridae	1
## 9	Carcinidae	1
## 10	Cyclopteridae	1
## 11	Cyprinidae	2
## 12	Euphausiidae	1
## 13	Gadidae	1
## 14	Gammaridae	1
## 15	Gobiidae	1
## 16	Harpacticidae	1
## 17	Labridae	1
## 18	Loliginidae	1
## 19	Lucicutiidae	1
## 20	Majidae	1
## 21	Megacalanidae	1
## 22	Metridinidae	1
## 23	Moronidae	1
## 24	Munididae	1
## 25	Myctophidae	1
## 26	Mytilidae	1
## 27	Nereididae	1
## 28	Octopodidae	1
## 29	Ommastrephidae	1
## 30	Ophiocomidae	1
## 31	Ophiodermatidae	1
## 32	Oplophoridae	2
## 33	Ostreidae	1

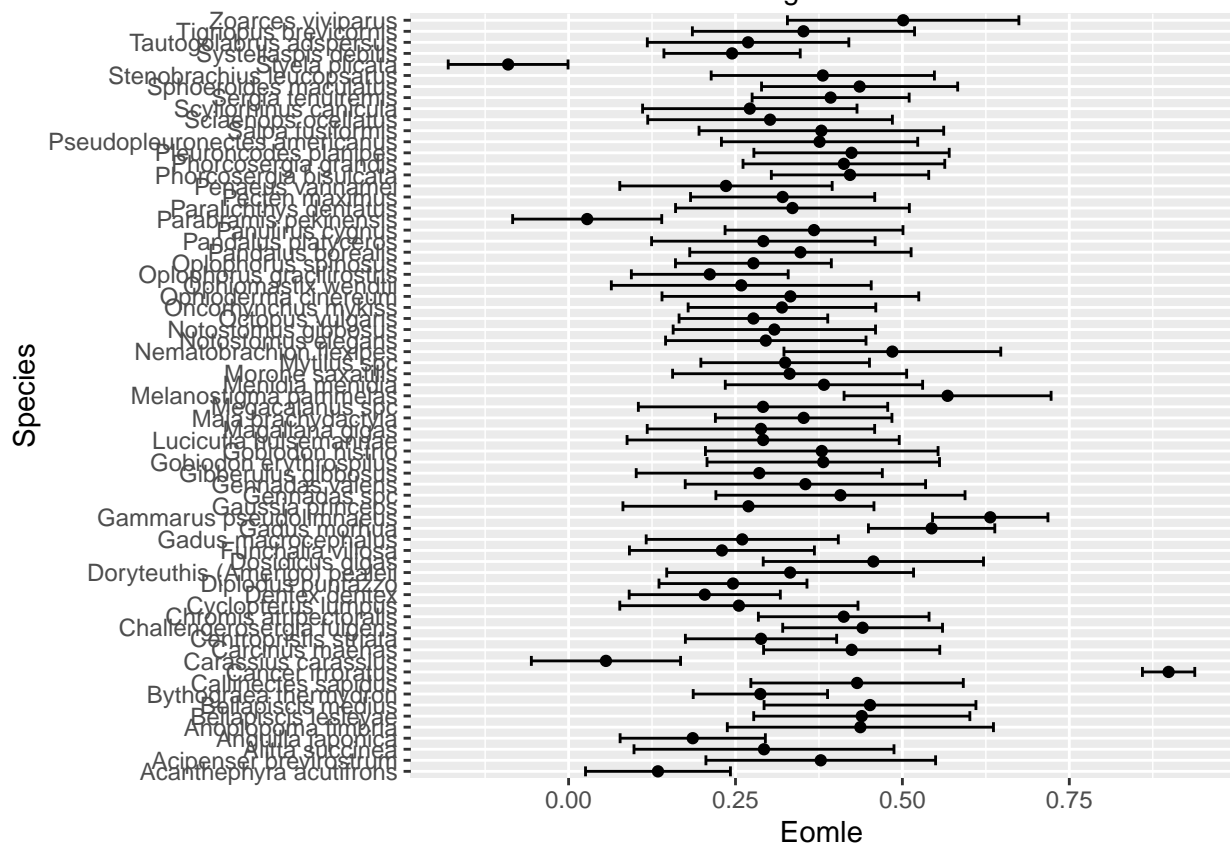
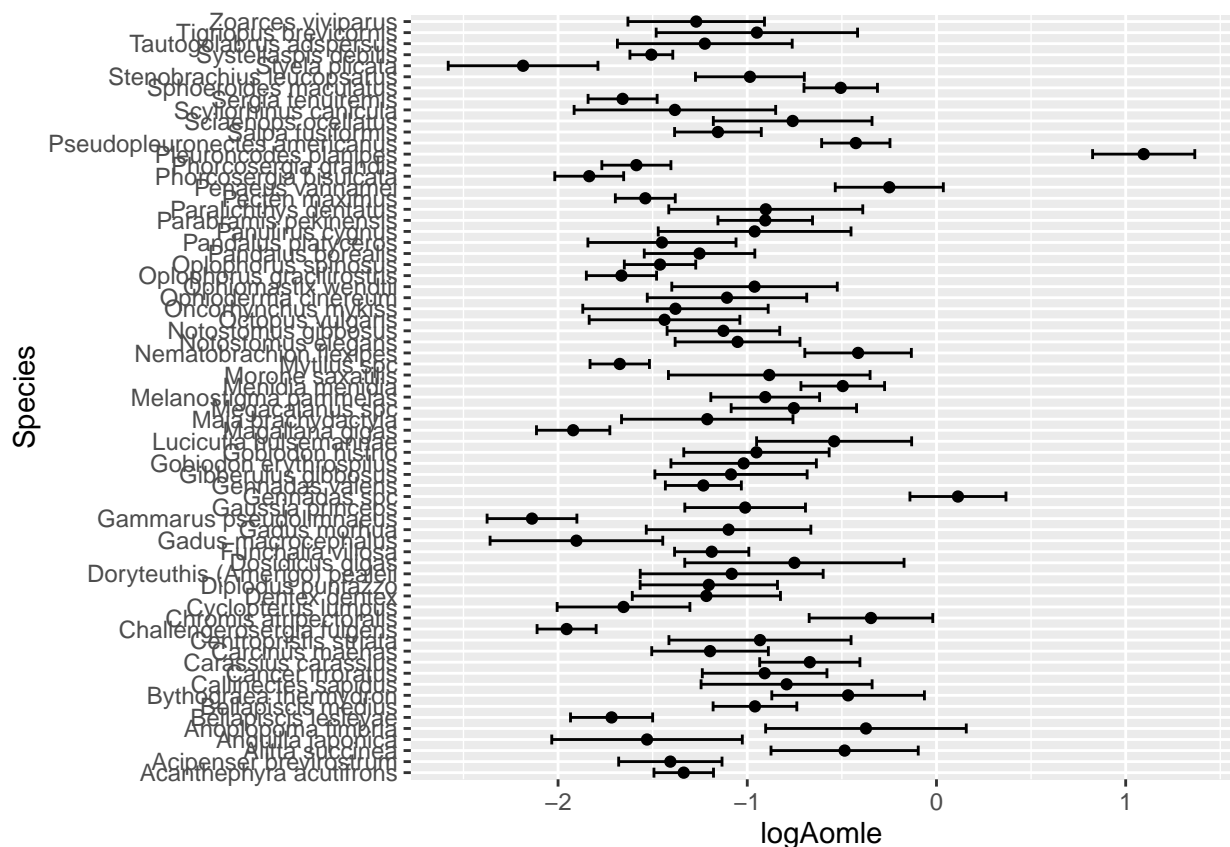
```
## 34 Palinuridae      1
## 35 Pandalidae      1
## 36 Paralicthyidae  1
## 37 Pectinidae      1
## 38 Penaeidae       2
## 39 Pleuronectidae  1
## 40 Pomacentridae   1
## 41 Portunidae      1
## 42 Salmonidae      1
## 43 Salpidae        1
## 44 Sciaenidae      1
## 45 Scyliorhinidae  1
## 46 Sergestidae     3
## 47 Serranidae      1
## 48 Sparidae        2
## 49 Strombidae      1
## 50 Styelidae       1
## 51 Tetraodontidae  1
## 52 Tripterygiidae  1
## 53 Zoarcidae       2
```

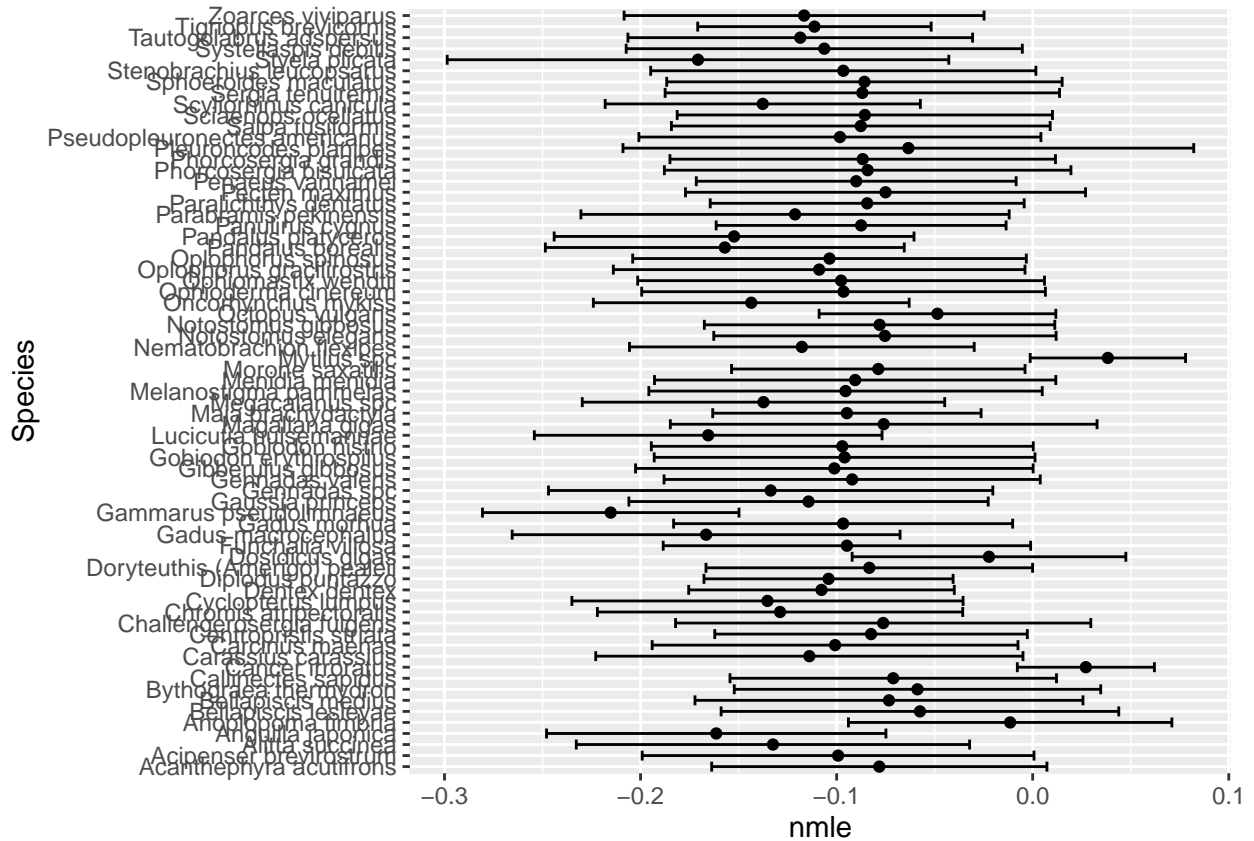
Model Estimates

```
## [1] 0
## Warning in sqrt(diag(object$cov.fixed)): NaNs produced

##           Estimate Std. Error
## alpha_j    -9.728885e-02 0.03318474
## alpha_j    -1.095815e+00 0.15046686
## alpha_j     3.235070e-01 0.05293057
## L_z        -8.226241e-03 0.01068749
## L_z         9.764702e-02 0.28085772
## L_z         1.997444e-01 0.14230360
## L_z        -4.829897e-02 0.05883830
## L_z         3.708694e-02 0.06927391
## L_z        -7.641764e-06 0.15932988
## log_lambda -1.741090e+01      NaN
## log_lambda  1.503793e+00 0.61682740
## log_lambda  1.233099e+00 0.62670717
## logsigma   -1.320911e+00 0.04445586
```







The data do not allow us to estimate Order-level mean trait values, as the maximum likelihood estimate of $\lambda_{l=2}$ is converging to 0 i.e., the estimated variance of traits among orders within a class is zero. This is not an inherent problem with the structure of the data: I have simulated the data using the same structure and find that I can reliably estimate order-level mean trait values. Rather, it suggests that there is not sufficient variation across Orders given the variation in species within Orders and Families.

Nonetheless, the estimated parameters largely make sense, and we minimally can generate priors for an out-of-sample taxa based on Class (though the differences there are small) or Family (where we have limited taxonomic coverage).

It is notable that the estimated body size coefficients are so close to 0. In my simulation testing it was very hard to accurately estimate n unless there were really huge differences among taxa (and even then it was tough). My guess it is because a lot of experiments / taxa don't have enough sizes to inform this, and the cross-species comparisons are heavily confounded with taxonomy.

What does it mean

This is a decent starting point for integrating into species distribution modeling. Comparing spatially explicit models with statistical presence-only data will be a useful next step.