Primate Allometric Scaling for Predicting Mass

Justin Petluk

April 15, 2020

1 Introduction

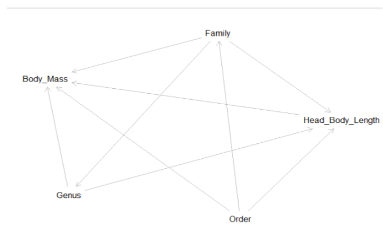
Exploring the relationship of the physical characteristics in animals is important to understanding the animal itself. In the case of this study, the ability to estimate mass using various levels of data can demonstrate that there exists broad relationships which can be used to further narrow down important features.

In this study, we use the taxonomic classifications to group animals of interest to explore the relationships within the physical structures. Taxonomy is leveraged because the inherent structure of the taxonomic classification groups similarities and differences of specific features despite it being absent in most data sets.

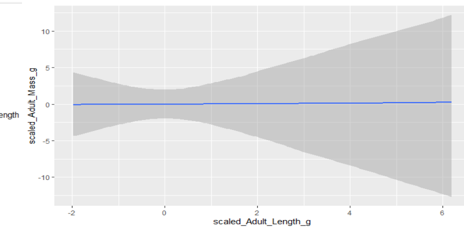
The goal of this study is to evaluate the taxonomic order, primates, and how the head and body length (nose to base of the tail) relates to the mass of the animal. The idea is that the there exists some sort of relationship where the larger the animal, the heavier it becomes. The question is, does this relationship exist? If so, how does this relationship exist across taxonomic classes, and can this be structured in a model to determine upper and lower limits of mass given a measure of size?

2 Methods

Data was collected from 1993-2004 across the world and incorporated previous data sets since 1979 from [4]. Within this data set, the decision was to explore primates. The data was collected up to the species-level of the animals, such that there exists only one value per-species for each category. Due to this approach, the authors decided to use unique median values for each species. This looses individual variation within the species such as sexual dimorphism since they either aggregate or unspecified sex [2]. The data contains freshly-killed and captive animals along with wild. If either body mass or head-body length contained -999 the data was removed from the data set since the data is missing. Terrestriality is either ground dwelling (1) or above ground dwelling (2). As we see in (Fig. 2.a), all samples are either ground dwelling or unspecified (-999).



(a) DAG for Causal Relationships



(b) Coloured by Terrestriality

Figure 1: We see that the prior model used representing the DAG assumes no relationship between mass and head and body length.

By looking at the data presented in (Fig. 2), a directed acyclic graph (DAG) as shown in (Fig.1) is designed to show the causal relationships between variables used. The idea is that there's a relationship between the taxonomic orders and both body mass and head and body length. It is also reasonable to assume that the larger (longer) the primate becomes, the heavier it will weight.

The approach to this data was to use BRMS [1] for a Bayesian multilevel linear regression model to predict the relationship between primate adult body mass and head and body length over their family and genus. We approach this with a linear multilevel regression model that follows:

$\alpha = ScaledAdultMass$
$\beta = ScaledAdultHeadBodyLength$
$\phi = Family$
$\gamma = Genus$

A Gaussian distribution is used with the formula:

$$\alpha \sim \beta + (1 + \beta|\phi/\gamma)$$

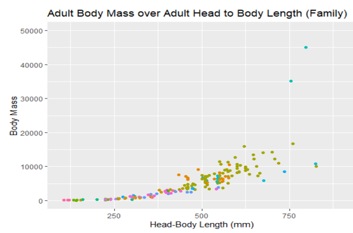
With priors used on scaled variables:

$$beta = Intercept = StandardDeviation = normal(0, 1)$$

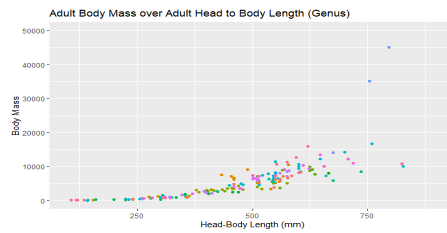
Where normal(0,1) is a normally distributed prior with an intercept of 0 and standard deviation of 1.
We see that this model with the priors given assumes that there is no relationship between the variables used by how the slope is 0 when the model is plotted in (Fig. 1).

3 Results

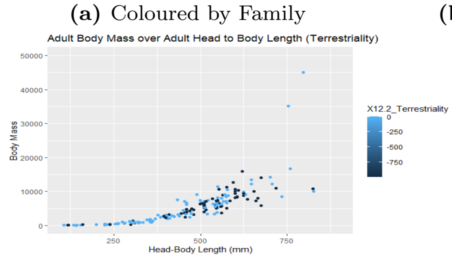
The model produced by the equation has been summarized in (Fig. 3). The model shows that there is a positive effect of the adult body length on



(a) Coloured by Family



(b) Coloured by Genus



(c) Coloured by Terrestriality

Figure 2: The three models here are identical with the exception of the colour coding. As the graphs show, the values in the relation are clustered in both family and genus. We see that this cluster does not exist with terrestriality in this data due to only having one valid class of data available.

estimated mass. Examining the model shows that there's a relatively small amount of error being within 10 percent on all group levels. We see between groups where only family is included and where family and genus are included, it broadens the credible interval for the body length, while narrowing the credible interval for the intercepts.

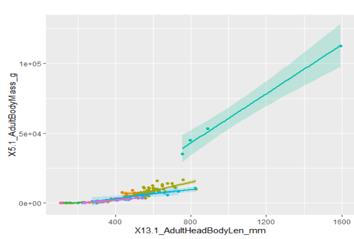
Plotting predictions for both family and genus levels presents a general positive relationship. The model shows that with large samples of data within the families that the trend is positive. The credible interval is for family is larger than that of genus, but there is quite a lot of variability within a taxonomic family alone so this can be explained. The genus level of the models show quite a bit of uncertainty due to the small amount of data presented. We notice that some of the linear estimations produce a negative relationship which is a consequence of the small data set. Since some only have two samples, this can be described as a negative slope, while in reality this might just be the variability in the mass-size relationship.

4 Discussion

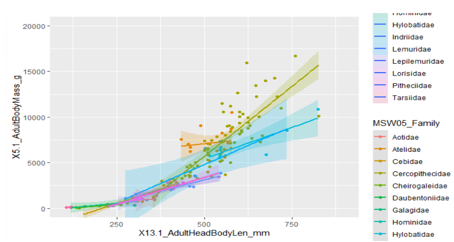
This study produced broad relationships between mass and head plus body lengths. It was able to show that generally, the two are positively correlated even at a family level. The model suggests that more information on the more specific levels would result in a more accurate estimation at a more specific

Group-Level	Effects:						
~Family	**Estimate**	**Est.Error**	**l-95%**	**CI**	**u-95%**	**CI**	**Rhat**
sd(Intercept)	0.36	0.1	0.2	0.58	1	1062	1667
sd(Adult Body Length)	0.38	0.08	0.25	0.58	1	1842	2368
cor(Intercept, Adult Body Length)	0.98	0.05	0.84	1	1	1117	1028
~Family:Genus	**Estimate**	**Est.Error**	**l-95%**	**CI**	**u-95%**	**CI**	**Rhat**
sd(Intercept)	0.1	0.02	0.06	0.14	1	1435	2736
sd(Adult Body Length)	0.1	0.02	0.06	0.14	1	1540	2087
cor(Intercept, Adult Body Length)	0.92	0.09	0.67	1	1	1399	1659
Population-Level	**Estimate**	**Est.Error**	**l-95%**	**CI**	**u-95%**	**CI**	**Rhat**
Intercept	-0.18	0.1	-0.4	0.03	1.01	764	1402
Adult Body Length	0.31	0.11	0.1	0.53	1.01	796	1504
Family							
	Estimate	**Est.Error**	**l-95%**	**CI**	**u-95%**	**CI**	**Rhat**
sigma	0.11	0.01	0.1	0.13	1	3249	3318

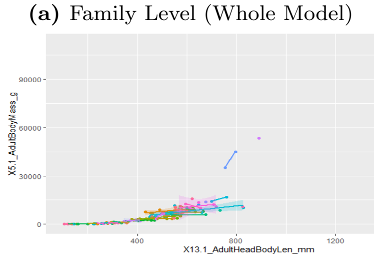
Figure 3: Summary table of the model.



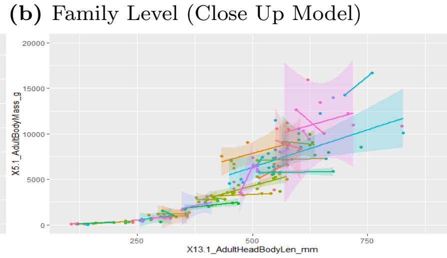
(a) Family Level (Whole Model)



(b) Family Level (Close Up Model)



(c) Genus Level (Whole Model)



(d) Genus Level (Close Up Model)

Figure 4: These figures evaluate the multilevel predictions per level.

level, which may be preferred.

Some limitations to the model is the data set not containing important information that may influence the data. This includes sex, an accurate age, and values not placed into a mean value. All of these things influence the variability in different ways and different weights. These may also influence the results in a non-linear form given a multilevel model. We also notice that this data set lacks the definition of a potentially important variable; terrestriality.

This study looks at the relationship between physical properties of primates and the relationship with their mass. Mass can describe the relationship with an animal's environment such as food availability, and this can also impact growth rate [3]. By evaluating the relationship between physical size, growth rates and mass, this can provide a more in-depth understanding of the animals being studied.

This study showed that there is a relationship between body mass and head-body length across different families of primates. This study shows that the relationship can better be described along side other variables that contribute to the weight of a primate that does not exist within the dataset used. The lack of important variables is a limitation to the model and inhibits the full capabilities of exploring the physical relationships. Future studies should approach this with more physical characteristics in mind. The data should also not aggregate individual information to display individual variability. As for the model used, a multilevel approach seems to be appropriate, though the genus level needs more evidence.

References

[1] Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. Journal of statistical software, 80(1), 1-28.

[2] Clutton-Brock, T. H., Harvey, P. H., Rudder, B. (1977). Sexual dimorphism, socionomic sex ratio and body weight in primates. Nature, 269(5631), 797-800.

[3] Gavan, J. A., Swindler, D. R. (1966). Growth rates and phylogeny in primates. American journal of physical anthropology, 24(2), 181-190.

[4] Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., ... Connolly, C. (2009). PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals: Ecological Archives E090-184. Ecology, 90(9), 2648-2648.