Proposal Math 655 – fall 2023

Author: Viviana Romero Alarcon

Problem

Body mass is one of the significant variables in the biology of an organism. Body mass (BM) could provide insights into their physiology, ecology, or behavior. Indeed, BM could influence speciation and extinction rates, driving patterns such as giantism in some vertebrates like dinosaurs, whales, and ratites. However, while obtaining BM data from extant species is relatively straightforward, it becomes challenging when attempting to obtain this variable for extinct organisms. Therefore, prediction becomes the only viable method to obtain this feature.

Various methods exist to reconstruct body size, shape, and weight from preserved skeletons (Fossils). The most popular approaches are volumetric density (VD) and extant scaling (ES). Although VD was the initial proposal for calculating BM, ES is currently the most widely used tool. It is based on linear models, is easy to replicate and use, and is more cost-effective compared to the scans used by VD. However, current models have faced criticism due to their potential to over- or under-estimate BM. Another challenge arises from the distinct distribution of mass in quadrupeds compared to bipeds, leading to more complexities in calculating the BM of bipeds as predictors need to be restricted only to forelimbs. Furthermore, the wide range of BM distribution in modern birds, ranging from 50g to 130,000g, presents an additional challenge, especially when considering the presence of giantism in the Palaeognatha group, which is represented as an outlier in the distribution.

To address the biped issue, some researchers have proposed models that exclusively utilize modern birds (excluding Palaeognathes), incorporate humans, or eliminate the humerus as a covariant. However, no attempts have been made to fit a model using different datasets, including ratites as an essential part of the sample. Re-evaluating the proposed models is thus warranted, considering that this group could serve as a good proxy for the BM of dinosaurs.

Similarly, in the case of sexual dimorphism, it is another unexplored feature that cannot be deduced from the fossil record. There are limited instances where complete skeletons could offer some insights into the sex of the specimens. Researchers have attempted to address this using morphometric geometric approaches or classic statistics (MANOVA, ANOVA, T-test). Still, it has not been evaluated as a classification problem using machine learning. Therefore, testing BM or femur length as a proxy for sexual dimorphism in biped dinosaurs seems imperative.

Methods:

I will collect data from four different papers[1,2,3,4]. These datasets provide information on body size in current bipeds (Birds and humans), femur length, and femur circumference. Additionally, they contain information on sex in modern birds, and BM calculated using the VD approach. In the case of Palaeognathae, there are only 2 observations with complete information and only 8 with BM and femur length. Hence, my initial step will be to fit a model enabling me to predict femur circumference from femur length. This is because, according to the literature, femur circumference is the best predictor of BM in bipeds.

Moreover, I will fit linear models using Femur circumference (FC), Femur Length (FL), and an index of FC and FL to predict BM. Also, I will test a quadratic model representing the allometric behavior of the BM variable. I will select the best model using k-fold cross-validation, estimate the Mean Squared Error (MSE) of my best model, and compare it with the previous models in the literature. Finally, I will evaluate how my best model and previous models predict BM in modern birds and biped dinosaurs, comparing it with BM estimated by VD.

In the case of sexual dimorphism, I will explore Femur circumference (FC), Femur Length (FL), and estimated BM as predictors of sex. For this, I will use classification approaches such as logistic regression, LDA, and QDA and measure the accuracy of the prediction.

Bibliography

1 Campione, N. E., & Evans, D. C. (2020). The accuracy and precision of body mass estimation in non-avian dinosaurs. Biological Reviews, 95(6), 1759-1797.

2 Chan, N. R. (2017). Phylogenetic variation in hind-limb bone scaling of flightless theropods. Paleobiology, 3 43(1), 129-143.

3 Campione, N. E., Evans, D. C., Brown, C. M., & Carrano, M. T. (2014). Body mass estimation in non-avian bipeds using a theoretical conversion to quadruped stylopodial proportions. Methods in Ecology and Evolution, 5(9), 913-923.

4 Field, D. J., Lynner, C., Brown, C., & Darroch, S. A. (2013). Skeletal correlates for body mass estimation in modern and fossil flying birds. PLOS one, 8(11), e82000.