Poster Presentation Number 63, Session 1, Thursday 14:00 - 15:30

Testing the limits of AI and machine learning for quantitative taphonomy. Skeletal-part representation as a proxy for the origins of human deposition

Jonas Grabbe[1,2], Ana Pantoja-Pérez [1,3], Ana Serrano-Mamolar[2,4], Antonio Canepa-Oneto[2,4] and Nohemi Sala[1,3]

1 - Centro Nacional de Investigación sobre la Evolución Humana, Burgos, Spain · 2 - Universidad de Burgos, Burgos, Spain · 3 - Centro Mixto UCM-ISCIII de Evolución y Comportamiento Humanos, Madrid, Spain · 4 - Advanced Data MIning Research And (Business intelligence | Bioinformatics | Big Data) LEarning, Escuela Politécnica Superior, Burgos, Spain

The analysis of skeletal-part representation provides a valuable approach within taphonomic research. Differential bone preservation can provide inferences regarding biostratinomic processes, such as episodes of rapid burial linked to funerary behaviors or catastrophic events. However, these patterns can also be significantly altered by fossil-diagenetic processes and excavation circumstances, potentially obscuring the original depositional signals. Furthermore, subtle variations in mortuary or funerary practices—from formal burials to surface deposition—can leave distinct yet easily confounded traces in the assemblage. Without a rigorous integration of geological history, sedimentary context, and detailed taphonomic analyses, interpretations of past behaviors risk conflating cultural signals with post-depositional biases.

In this context, machine learning (ML) approaches offer a complement to traditional taphonomic methods, potentially reducing the subjective biases inherent in traditional visual interpretation. Recently, ML and artificial intelligence (AI) techniques have rapidly expanded in archaeological and paleoanthropological research, with applications extending across diverse taphonomic contexts [1]. Despite their growing adoption, the overall use and effectiveness of ML techniques in taphonomic research remain underexplored. This is primarily due to their recent introduction to the field and the limited availability of high-quality, abundant data.

In this study, we present a comprehensive database of more than 120 sites with Skeletal-Part representation and an analytical pipeline that applies ML techniques for standardized analysis to a subset of this database. This approach enables reproducible classification of taphonomic signatures and supports large-scale comparative studies of depositional behaviors. To build this dataset, we standardized abundance records from the paleoanthropological sites, encompassing skeletal-part representation data from more than 1,000 individuals. Each record incorporates detailed element identification including laterality, geological contexts, chronological periods, and specific depositional contexts (e.g., cannibalism, carnivore activity, geological transport, or funerary practices). Where available, we also distinguished between funerary behaviors, including inhumation, collective burial, and funerary caching.

Raw Minimum Number of Elements (MNE) data were transformed into standardized quantitative variables—such as Minimum Number of Anatomical Units (MNAU), percentage MNAU (%MNAU), and Relative MNAU (ReMNAU). We used various imputation techniques to achieve consistent bone-level resolution across sites. These curated datasets were then analyzed using supervised machine learning classifiers, including Random Forest, Support Vector Machines, and other advanced AI methods, to differentiate the cause of the accumulation and varying bone-level resolution. Model robustness was assessed through stratified cross-validation using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Finally, we analyzed the effect of bone-level granularity on classification performance, assessing how detailed anatomical resolution influenced the accuracy of taphonomic predictions.

Despite ongoing challenges, including post-depositional disturbances and limited sample sizes, our results highlight specific anatomical patterns that most effectively signal different depositional origins. By automating taphonomic classification with transparent, open-access tools, reproducibility across paleoanthropology can be enhanced. Clarifying both the strengths and limitations of ML/AI approaches applied to rigorously curated skeletal datasets, our work enables large-scale comparative analyses of hominin depositional behavior and offers novel quantitative insights into the fossil record.

This research has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant agreement No. 949330) and PID2021-122355NB-C31 and PID2021-122355NB-C33 funded by MCIN/AEI/10.13039/501100011033/FEDER, UE. NS has received funding from the Ministerio de Ciencia e Innovación under the Ramón y Cajal program (RYC2020-029656-I) funded by MCIN/AEI/10.13039/501100011033 and El FSE invierte en tu futuro. The authors wish to thank the DEATHREVOL team.

References: [1] Bellat, M., Figueroa, J.D.O., Reeves, J.S., Taghizadeh-Mehrjardi, R., Tennie, C., Scholten, T., 2025. Machine learning applications in archaeological practices: a review. arXiv preprint arXiv:2501.03840.