O-062

FROM MEASUREMENT ERROR TO HISTORICAL INSIGHTS: A NEW APPROACH FOR CLASSIFYING ZOOLOGICAL SPECIMENS USING PORTABLE NIR SENSORS AND CHEMOMETRICS

Giussani, Barbara[1]; Riu, Jordi[2]; Monti, Manuel[1]; Baruffaldi, Lorenzo[1]; Campeny, Marc[3]; Quesada, Javier[3].

[1]University of Insubria, Science and High Technology Department, Como, Italy; [2]Universitat Rovira i Virgili, Department of Analytical Chemistry and Organic Chemistry, Tarragona, Spain; [3]Natural Sciences Museum of Barcelona, Barcelona, Spain.

Zoological specimens are essential for reconstructing past biodiversity, but their research value relies on metadata such as species, age, and collection context. When key information—especially location and historical period—is lost (due to war, neglect, or poor documentation e.g.), specimen utility is compromised. While biological traits can be recovered, assigning specimens to specific historical periods or collections remains challenging. Traditional research is time-consuming, and some analytical techniques are unsuitable for recent collections, highlighting the need for efficient methods of historical classification. Moreover, given the historical and archaeological value of these samples, it is essential to employ non-destructive analytical techniques that can be applied directly on-site, minimizing the risks associated with handling such valuable and fragile materials.

In this presentation, we will share our initial encouraging results using portable NIR (Near-Infrared) sensors to classify specimens - mammal skeletons of red squirrel (*Sciurus vulgaris*) - from the collections of the Natural Sciences Museum of Barcelona into historical periods. The project aimed to optimize an analytical protocol suitable for practical application in museums, taking into account the various sources of variability that contribute to spectroscopic measurement errors.

The approach focused on studying the multivariate error associated with the measurements and incorporating this information directly into visualization and into classification multivariate models to enhance their performance. The multivariate error was estimated through the analysis of error covariance and correlation matrices of the raw spectroscopic data [1, 2 and references therein].

The first aspect we optimized was the measurement protocol for the samples, which were presented as whole skulls, partially preserved skulls, mandibles, and, in some cases, additional bone elements considered fit for analysis. The study of multivariate error of the raw data proved valuable in selecting the most suitable sampling configuration and determining the optimal number of replicates to perform on each zoological sample.

Subsequently, visualization and classification models were developed using Principal Component Analysis (PCA) and Discriminant Analysis (DA), integrating the information derived from multivariate error estimation [3]. This approach allowed for effective separation of samples into classes representing modern and historical periods. The classification models were validated using external validation.

In this study, we evaluated two of the most widely used portable NIR sensors available on the market. The first, the NeoSpectra Scanner (Si-Ware), operates within the 1250-2500 nm range, while the second, the MicroNIR OnSite-W (VIAVI Solutions), spans from 908 to 1976 nm. Both sensors operate using external reflection. When combined with chemometric data analysis, both sensors were deemed suitable for the intended purpose. In both cases, incorporating multivariate error into the models improved their performance, making the visualization models clearer and more interpretable and the classification models more effective. Differences in the sources of variability between the two sensors were observed, affecting the quality of the spectroscopic signal in different ways and, consequently, the performance of the classification models. Incorporating multivariate error into the models allowed us to identify the optimal solution for each sensor, demonstrating the significant value of studying raw spectroscopic data error when working with handheld NIR devices.

References:

[1] P.D. Wentzell, *Measurement errors in multivariate chemical data*, J. Braz. Chem. Soc. 25 (2014) 183–196

[2] B. Giussani, G. Gorla, J. Ezenarro, J. Riu, R. Boqué, *Navigating the complexity: Managing multivariate error and uncertainties in spectroscopic data modelling*, TrAC 181 (2024) 118051

[3] P.D. Wentzell, D.T. Andrews, D.C. Hamilton, K. Faber, B.R. Kowalski, *Maximum likelihood principal component analysis*, J. Chemom. 11 (1997) 339-366