

## I. INTRODUCTION

Scrap recycling serves as a reliable, sustainable, and cost-efficient source of raw materials, promoting economic and political independence while reducing potential conflicts with mining activities. Real-time quantitative analysis of alloy compositions enables precise monitoring of heterogeneous mixtures, facilitating alloy-to-alloy recycling and optimal material utilization. However, accurate and rapid analysis of these complex metal flows is a challenging task.

Prompt Gamma Neutron Activation Analysis (PGNAA) captures neutron-induced gamma emissions, providing detailed elemental fingerprints non-destructively. It is established in diverse sectors, from security screening and petrochemical exploration to real-time process control in cement and coal industries [1]. This technique is particularly suited for real-time monitoring of metal alloy compositions, enabling quantitative assessment of alloy compositions.

Building on our previous work on categorical spectral classification of copper and aluminium alloys, we achieved promising results using Maximum Log-Likelihood methods and neural networks, with synthetic datasets generated via probability-distribution-based sampling of long-time spectra [2]. Real, independent measurements confirmed the robustness of this approach.

However, prior studies focused on categorical classification, assigning each sample to a single class. Consequently, alloy mixtures lying between two classes can only be assigned to one of them, which prevents an unambiguous analysis and poses a limitation for accurately predicting material compositions in future applications.

In this work, we present for the first time a regression-based approach for predicting PGNAA spectra of metal alloys, enabling the estimation of arbitrary alloy mixtures. This represents an important step toward advanced recycling applications, particularly in metal recycling.

To evaluate this regression framework, synthetic datasets are generated from long-time spectra using a probability-distribution-based sampling approach. These datasets represent different alloy ratios, modeled as convex linear combinations of two reference spectra. Each mixture's composition is uniquely determined by a single mixing parameter, which is estimated via Maximum Likelihood analysis of the available materials, forming the basis of our regression approach.

## II. DATA & PGNAA

PGNAA enables continuous material analysis by irradiating samples with thermal neutrons, which excites the atomic nuclei and induces nuclear reactions. As these nuclei relax, they emit characteristic prompt gamma radiation unique to the elemental composition of the sample. This radiation is captured using a gamma-ray spectrometer, as illustrated in Fig. 1. All spectra presented in this study were acquired with a high-purity germanium (HPGe) detector. However, as shown by the comparison in Fig. 1, the classification of metal alloys becomes significantly more challenging with shorter measurement times due to increased noise and a reduced number of registered gamma events. As a basis for the following analysis, we utilize

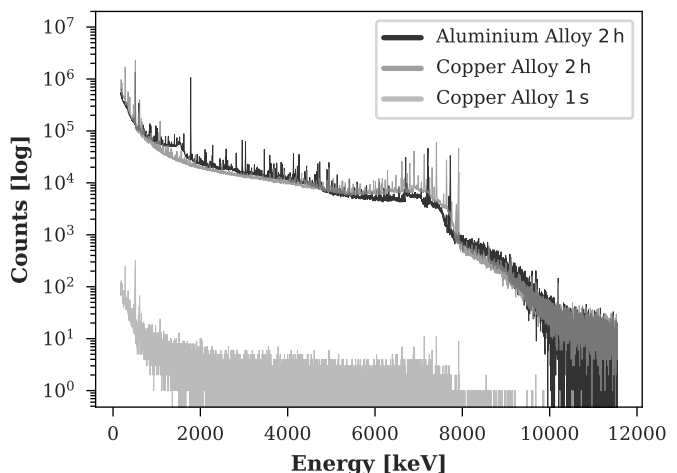


Fig. 1: Spectra of Al\_2 (aluminium) and Cu\_5 (copper) alloys were measured for 2 h, plus one Cu\_5 spectrum for 1 s.

long-time measurements<sup>1</sup> of Aluminium and copper alloys, denoted as Al\_1, . . . , Al\_5 and Cu\_1, . . . , Cu\_5, respectively, each with an acquisition time of 2 h.

## III. DATA CREATION

The sampling method is based on the long-time measurements described previously, which serve as an empirical foundation for data generation. By normalizing these spectra—dividing the counts of each energy channel by the total count sum—the data is transformed into a discrete probability distribution representing a single gamma event. Short-time measurements are simulated by drawing 50 000 counts from the long-time spectrum according to the relative probabilities of each energy channel, corresponding to multinomial sampling [2]. During two-hour measurements, the large number of recorded gamma events ensures that the distributions are representative of the true spectra, according to the law of large numbers, providing a robust template for generating realistic short-time measurements.

## IV. RESULTS

In this section, the Maximum Log-Likelihood method is employed to perform regression-based predictions for measurements of mixed materials. We consider two distinct alloys, an Aluminium alloy and a Copper alloy, with discrete probability distributions  $p = (p_1, \dots, p_n)$  and  $q = (q_1, \dots, q_n)$ , and corresponding photon counts  $X = [x_1, \dots, x_n]$  and  $Y = [y_1, \dots, y_n]$ , respectively (Chapter III). A mixture with mixing parameter  $\lambda \in [0, 1]$  is then given by the linear combination

$$\text{Mix}(\lambda) = \lambda X + (1 - \lambda) Y,$$

where each channel in the mixture is a weighted sum of the corresponding channels in the reference spectra. As  $p$  and  $q$  are normalized probability distributions, the distribution of the mixed material  $\text{Mix}(\lambda)$  is given by the linear combination  $m(\lambda) = \lambda p + (1 - \lambda)q$ . Thus, for any given  $\lambda$ ,  $m(\lambda)$  defines the theoretical probability distribution for the corresponding mixing ratio between the Aluminium and copper

<sup>1</sup>The spectra used in this study, as well as additional materials, are available online: <https://www.kaggle.com/datasets/smartfactoryowl/metalclass>

alloys. Furthermore, let  $c = (c_1, \dots, c_n)$  denote a short-time measurement of photon counts originating from a mixture  $\text{Mix}(\lambda)$  for a fixed but unknown parameter  $\lambda$ . The goal is to estimate the mixing parameter  $\lambda \in [0, 1]$ , which represents the proportion of aluminium alloy within the mixture. The parameter  $\lambda$  from the short-time measurement  $c$  is estimated by maximizing the log-likelihood function under a multinomial counting model:

$$\ell(\lambda) = \log p(c \mid \text{Mix}(\lambda)) = \sum_{i=1}^n c_i \log m_i(\lambda).$$

The likelihood function evaluates how well the observed counts  $c$  align with the theoretical mixture  $m(\lambda)$ . By maximizing this function, the measurement is assigned to the specific distribution that provides the best statistical description of the observed data.

A closed-form analytical solution for the equation  $\ell'(\lambda) = 0$  generally seems not to exist. Therefore, the optimal mixing parameter  $\lambda$  is determined numerically by minimizing the negative log-likelihood function within the interval  $[0, 1]$ . We employ the bounded scalar minimization algorithm, which requires on average only 0.056 s per estimate. In the fol-

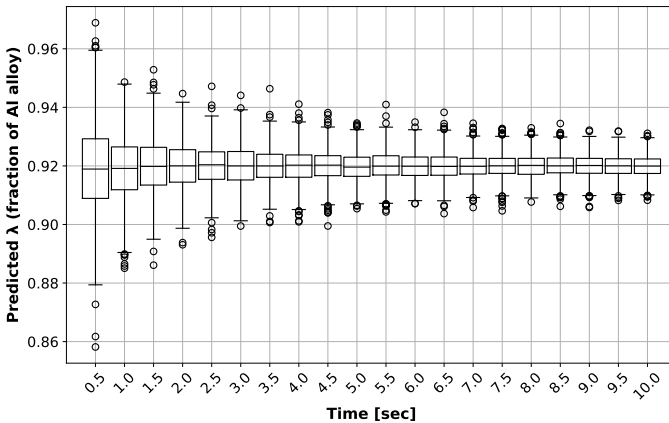


Fig. 2: Predicted aluminium alloy fraction for simulated short-time Al<sub>2</sub>-Cu<sub>5</sub> mixture data (preset aluminium fraction  $\lambda = 0.92$ ). Boxplots show 1000 predictions per measurement time, showing decreasing dispersion and increasing accuracy with longer acquisition times.

lowing, the prediction accuracy of the regressive approach is investigated for different measurement times. Figure 2 shows the predicted Aluminium alloy fraction in simulated short-time measurements of a Al<sub>2</sub>-Cu<sub>5</sub> mixture with a preset aluminium fraction of  $\lambda = 0.92$  as a function of measurement time. The boxplots represent the distribution of 1000 predictions for each respective measurement time.

All boxplots exhibit a median close to the ground truth. As the measurement time increases, the dispersion of the predictions decreases significantly, which is reflected in the narrowing of the boxes. After only 0.5 s, 95% of the estimates lie within a deviation of  $\pm 3\%$  from the true value, while after 1.5 s, 98% of the estimates are already within  $\pm 2\%$ . This demonstrates the increasing prediction accuracy achieved with longer acquisition times.

Finally, we examine mixtures composed of various copper alloys and different Aluminium alloys. Each mixture consists of 8% Aluminium alloy and 92% copper alloy, with a simulated measurement time of 1 s. The predictive performance is evaluated across these material combinations, and the dispersion of the estimates for each case is presented as a boxplot (Fig. 3).

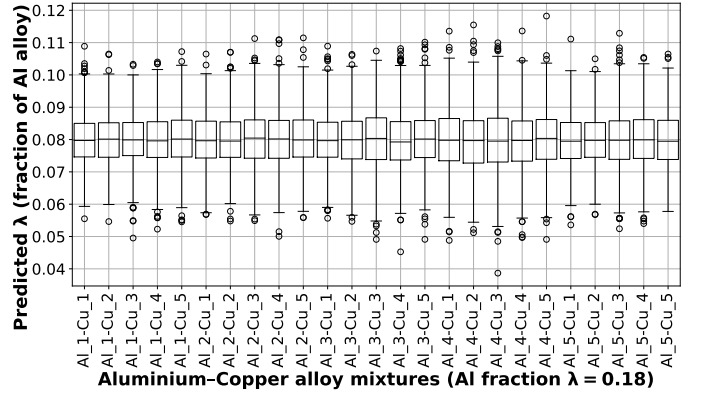


Fig. 3: Predicted aluminium and copper alloy fractions in simulated 1 s Cu-Al mixtures (preset aluminium fraction  $\lambda = 0.08$ ). Boxplots show 1000 predictions per mixture, showing variability across alloy combinations and reliable accuracy even at low aluminium fractions.

## V. CONCLUSION

In this work, we introduced a regression-based approach for PGNAA spectroscopy that enables the quantitative determination of mixture ratios in metal alloy compositions. By formulating the analysis as a regression problem, the method addresses the fundamental limitations of conventional classification approaches, which are unable to resolve intermediate compositions.

Despite the significant statistical noise inherent in short-time measurements, the proposed Maximum Log-Likelihood framework achieves high prediction accuracy. Even at a measurement time of only 1.5 s, 98% of the predictions deviate by less than 2% from the true mixture fraction. The framework provides a technical foundation for future real-time monitoring systems in the field of alloy recycling.

This work represents an important step toward the precise, non-destructive online analysis of heterogeneous metal flows. It establishes a foundation for future developments in real-time alloy composition monitoring, which is highly relevant for the advancement of automated metal recycling processes.

## VI. ACKNOWLEDGMENTS

This work was funded by EU & NRW (EFRE-20800105)

## REFERENCES

- [1] M. R. Abdelnour, et al., 'Prompt gamma neutron activation analysis: A review of applications, design, analytics, challenges, and prospects' in *Radiation Physics and Chemistry* (2025, S. 112693).
- [2] H. Shayan, et al., 'PGNAA Spectral Classification of Metal with Density Estimations' *IEEE Transactions on Nuclear Science* ( Volume: 70, Issue: 6, June 2023).