Machine learning automated data analysis of microplastics measurements and its implications on advancing microplastics analysis to routine analysis

Michaela Radl*[1], Aurelia Liechtenstein†[1], and Benedikt Hufnagl‡[1]

[1]Purency GmbH – Walfischgasse 8/34 1010 Vienna, Austria

Abstract

To understand the occurrence and origin of microplastics (MP) in our environment, a reliable and scalable measurement method is needed. Promising methods for MP analysis are based on micro-spectroscopy since it allows to determine the particle shape as well as the chemical structure, using Raman and micro-FTIR (imaging). However, the generated amounts of complex data pose a challenge:
The conventional approach for analysing spectral data is by using spectral library searches. However, studies show that this approach has limitations in the context of MP analyses. Manual interventions and various parameter settings (reference spectra, similarity, operator, etc.) lead to reduced reproducibility and comparability. Also, the limited amount of reference spectra in a library leads to results of questionable quality (false positives and false negatives). Meanwhile the analysis is time-consuming and, thus, costly.

These problems led to new developments in the use of spectral library search and finding alternative solutions, such as the introduction of machine learning (ML) technology. ML: To solve the MP classification problem, a model was trained to decide which pixel in a hyperspectral image is a polymer and which one is not. It was also taught to distinguish between different types of polymers.
In contrast, library-based analysis solutions consists of hundred IR spectra in their library. The cross-checking of each pixel in the library is slow. A more robust classification requires a bigger library which will slow down the process of the classification even further. The advantage with ML is that there is no trade-off between analytical quality and speed. Models can be trained with sufficient data to produce robust results for a wide range of different samples while decreasing the time needed to perform the data analysis. The broad applicability of only one model and the reduction of manual intervention increases reproducibility and comparability.

Keywords: Microplastics analysis, Machine learning, micro fourier transform interferometer (μFTIR), Raman, hyperspectral image

*Speaker
†Corresponding author: Aurelia.Liechtenstein@purency.ai
‡Corresponding author: Benedikt.Hufnagl@purency.ai