**Abstract:**

The growing presence of anthropogenic pollutants in aquatic ecosystems has become one of the most important challenges in contemporary environmental science. Among these pollutants, microplastics (MPs) and contaminants of emerging concern (CECs) have attracted increasing attention due to their widespread occurrence, persistence and potential impacts on ecosystems and human health. Their determination is currently based mainly on chromatographic and spectrometric techniques, which provide excellent analytical performance but often require extensive sample preparation, high solvent consumption and laboratory-based instrumentation. Consequently, there is a growing need for complementary analytical methodologies capable of providing rapid, non-destructive and environmentally sustainable alternatives for environmental monitoring. Within this framework, the present thesis investigates the integration of advanced spectroscopic techniques and chemometric methods for the characterization of complex environmental matrices, following a progressive research pathway that moves from model systems towards real environmental samples. The first part of the work, carried out at the University of the Basque Country (UPV/EHU, Bilbao), focused on the spectroscopic characterization of microplastics using laboratory Raman microscopy, portable Raman spectroscopy and MicroNIR spectroscopy. The objective was to evaluate the capability of vibrational spectroscopy, combined with multivariate statistical analysis, to discriminate different polymeric materials and establish robust chemometric workflows suitable for rapid and sustainable plastic identification. Building upon the knowledge acquired during this stage, the second part of the research, developed at the University of Modena and Reggio Emilia (UNIMORE), extended the same analytical philosophy to considerably more complex environmental matrices containing contaminants of emerging concern. The study involved Polar Organic Chemical Integrative Samplers (POCIS) collected at the influent and effluent of the Modena wastewater treatment plant together with simulated solid-phase extraction (SPE) matrices prepared according to a Design of Experiments (DoE) using Oasis HLB sorbent and five pharmaceutical standards (caffeine, hydrochlorothiazide, diclofenac, metoprolol and atenolol). These samples were investigated using laboratory Raman microscopy, hyperspectral imaging (HSI), portable Raman spectroscopy and MicroNIR spectroscopy in order to evaluate the analytical capabilities and limitations of each platform for the investigation of environmentally relevant contaminants. A comprehensive chemometric workflow was developed in MATLAB to maximise the extraction of chemical information from the acquired spectroscopic datasets. Particular attention was devoted to Multivariate Curve Resolution–Alternating Least Squares (MCR-ALS), implemented through different modelling strategies according to the characteristics of each analytical platform. For Raman mapping and hyperspectral imaging, an original workflow combining Principal Component Analysis (PCA), data geometry analysis through Convex Hull selection and MCR-ALS was developed to reconstruct two-dimensional chemical distribution maps directly comparable with the acquired spectroscopic images. Conversely, portable Raman and MicroNIR datasets were analysed through conventional, multiset and equality-constrained

MCR-ALS approaches, incorporating prior chemical information from reference standards to improve spectral resolution and facilitate contaminant recognition.

The obtained results demonstrated that laboratory Raman microscopy provides the highest chemical specificity, successfully distinguishing the spectral contribution of the Oasis HLB sorbent from additional spectral features potentially associated with retained contaminants while preserving their spatial distribution across the POCIS surface. Hyperspectral imaging proved to be an efficient tool for the rapid visualization of spatial heterogeneity within the analysed samples, although its limited spectral resolution hindered the discrimination of individual contaminants in highly complex matrices, highlighting the trade-off between spatial and chemical resolution. Portable Raman and MicroNIR spectroscopy, when combined with constrained MCR-ALS approaches, successfully resolved several characteristic spectral signatures of the investigated pharmaceuticals. In particular, the equality-constrained strategy improved the recognition of individual standards by exploiting characteristic spectral regions, whereas the multiset approach facilitated the incorporation of prior chemical knowledge into the optimisation process. Nevertheless, both techniques revealed that the reconstruction of concentration profiles remains one of the principal challenges for spectroscopic analysis of complex sorbent matrices, mainly due to the dominant spectral contribution of the resin and the extensive overlap between analyte and matrix signals. Overall, this work demonstrates that the principal limitation of spectroscopic techniques for the investigation of complex environmental samples is not exclusively instrumental but largely chemometric. The results underline the crucial role of advanced multivariate modelling in enhancing the analytical performance of sustainable spectroscopic methodologies and show how the integration of complementary techniques, prior chemical knowledge and constrained optimisation strategies can significantly improve the interpretation of complex datasets. The developed workflows provide a robust methodological framework for future investigations on both microplastics and contaminants of emerging concern, while establishing the foundations for the development of integrated, rapid and environmentally sustainable analytical strategies for aquatic environmental monitoring.

# 1. Introduction

This thesis project - which also included a three-month research period abroad at the University of the Basque Country (UPV/EHU) - originates from the growing need to strengthen environmental prevention strategies, with a particular focus on the monitoring of wastewater and the performance of treatment plants.
The study is part of a broader research initiative, known as EXWASTER, which aims to develop analytical approaches (both target and non-target) for the surveillance of wastewater streams, with the ultimate goal of promoting more sustainable and circular water-management practices. Within this framework, the project seeks to define analytical strategies that integrate chromatographic and spectroscopic techniques with multivariate data analysis, in order to monitor both inorganic contaminants (e.g., heavy metals) and organic micropollutants classified as emerging contaminants. For the latter, the research involves the use of Polar Organic Chemical Integrative Samplers (POCIS) - a type of passive sampler consisting of two superimposed disks containing a microporous membrane and a sorbent resin. These devices gradually accumulate polar and moderately polar compounds over time, allowing their subsequent qualitative and quantitative assessment and making them particularly suitable for long-term monitoring in wastewater systems.
With regard to the period spent abroad, the research project included the study of wastewater with the aim of qualitatively determining the presence of microplastics and, consequently, assessing the current condition of the Crispijana wastewater treatment plant (EDAR). The IBeA research group – Department of Analytical Chemistry at UPV/EHU – has been working for several years on the development of analytical approaches applicable to the Galindo treatment plant, in collaboration with the Aguas Bilbao Bizkaia Consortium. These activities form part of the PlasticEDAR project and the ongoing PlastCare project. So far, the work carried out at the Galindo facility has led to the development of a method for detecting and quantifying MPs in water samples using Raman spectroscopy, with a detection capability down to roughly 200 µm. However, robust procedures for sludge analysis still need to be established, as this matrix is far more complex and lacks certified standards or reference materials. The expertise gained over the years has prompted the proposal of this new research line, which aims to deepen the study of sludge samples and extend the diagnostic assessment to a second major treatment plant, namely the Crispijana EDAR.

## *1.1. Scientific context and motivation*

Environmental prevention is a broad framework of strategies, policies, and regulatory measures aimed at identifying and eliminating environmental damage - or at least mitigating its impact - by acting on the underlying causes before adverse effects can occur. Within this context, the treatment and management of wastewater play a crucial role, as wastewater represents one of the main pathways through which microplastics and emerging contaminants

enter ecosystems and reach living organisms[1]. Wastewater streams, which include both industrial and municipal effluents, convey a wide variety of biologically active and persistent chemical substances toward treatment facilities.

Wastewater treatment plants (WWTPs) were originally designed to remove the highest possible fraction of pollutants and suspended solids; however, over the last decade, numerous studies have shown that, despite their high efficiency in eliminating organic and microbiological contaminants, these systems do not consistently achieve effective and reproducible removal of emerging molecules or solid particles such as microplastics and nanoplastics. Their limited performance is primarily related to the small size, tendency to form complex aggregates, and high stability of these substances - features that enable them to pass through treatment barriers and ultimately accumulate in sewage sludge or persist in treated effluents. For these reasons, their presence in "recycled" or treated waters has become increasingly critical for ecosystem protection. In summary, monitoring wastewater is a key step both for assessing the ecotoxicological impact associated with MPs and ECs, and for ensuring a more sustainable management of water resources - an aspect that is becoming essential in a context of progressive water scarcity.

Microplastics (MPs) are commonly described as plastic particles smaller than 5 mm, extending down to the micrometre scale; particles below this range are usually referred to as nanoplastics[2]. Based on their origin, microplastics are generally divided into two groups. Primary microplastics enter the environment directly as small particles, often as a result of industrial activities, whereas secondary microplastics form when larger plastic items break down through weathering and other degradation processes[3]. Concerns about the environmental consequences of plastic pollution - driven by the accumulation of plastic debris in ecosystems and organisms - have been discussed for several decades, but the issue has become particularly urgent in recent years. Between 2010 and 2020, scientific interest grew sharply, with a notable rise in studies examining the sources of microplastics, their potential effects on human and environmental health, and the analytical approaches used to identify and classify them. Recent research suggests that microplastics do not typically cause acute, lethal effects in exposed organisms. However, chronic impacts remain a major concern, especially in the context of long-term exposure, and may involve oxidative stress, bioaccumulation, or metabolic disruption[4].

The polymers most frequently reported as microplastics include polyethylene (PE), polypropylene (PP), polyvinyl chloride (PVC), polyethylene terephthalate (PET), polystyrene (PS), and polycarbonate (PC)[5].

---

[1] Browne et al., «Accumulation of Microplastic on Shorelines Worldwide».

[2] Leslie et al., «Microplastics En Route».

[3] Santos et al., «Marine Debris Contamination along Undeveloped Tropical Beaches from Northeast Brazil».

[4] Li et al., «Microplastics in Freshwater Systems».

[5] Hidalgo-Ruz et al., «Microplastics in the Marine Environment».

Emerging contaminants (ECs) refer to a broad and heterogeneous group of natural or synthetic chemical substances used across industrial, agricultural, domestic, and medical sectors, which have recently attracted growing scientific attention. Their relevance has increased in parallel with studies on microplastics, mainly because many of these compounds exhibit noteworthy ecotoxicological properties. The term *emerging* does not imply that these substances are newly discovered; rather, it reflects a renewed scientific interest stemming from their ability to contaminate environmental compartments - such as surface waters, wastewater, sediments, soils, and biota - and from the recognition of adverse effects that were initially overlooked. Conversely, the word *contaminants* highlights the characteristics that make them particularly concerning[6].

Many of these pollutants are biologically active even at very low concentrations (in the ng/L to µg/L range), to the point of affecting the physiological, reproductive, or behavioural functions of exposed organisms[7]. In addition, a considerable fraction of them is not effectively removed by wastewater treatment plants (WWTPs), as these facilities were not designed to retain complex molecules such as pharmaceuticals and personal care products (PPCPs), flame retardants, perfluoroalkyl substances (PFAS), or fragments of plastic materials (MPs and NPs)[8]. Their continuous production and widespread release into the environment therefore contribute to chronic exposure scenarios that are often difficult to monitor and evaluate comprehensively[9].

In this context, the growing global interest in environmental monitoring has driven the scientific community to develop advanced analytical methodologies capable of identifying, quantifying, and tracing the environmental fate of these contaminants.

In view of these considerations, it becomes evident that the effective monitoring of MPs and ECs depends not only on understanding their environmental behaviour, but also on the availability of analytical techniques capable of capturing their complexity. This naturally leads to the discussion presented in the next section, which focuses on the challenges and limitations of conventional detection methods.

### *1.2. Analytical challenges*

Given the relevance and urgency of the topic, a substantial body of literature is devoted to the investigation of microplastics (MPs) and emerging contaminants (ECs) in wastewater treatment plants (WWTPs). Nevertheless, a careful review of current research shows that the analytical characterization of MPs and ECs is shaped by a methodological divide closely linked to the objectives of each study. On the one hand, many investigations prioritise the acquisition

[6] Schwarzenbach et al., «The Challenge of Micropollutants in Aquatic Systems».

[7] Wu et al., «Wastewater Treatment Plants Act as Essential Sources of Microplastic Formation in Aquatic Environments».

[8] Ofrydopoulou et al., «Assessment of a Wide Array of Organic Micropollutants of Emerging Concern in Wastewater Treatment Plants in Greece».

[9] Finckh et al., «A Risk Based Assessment Approach for Chemical Mixtures from Wastewater Treatment Plant Effluents».

of reliable chemical information - such as polymer identity, additives, and adsorbed pollutants - while on the other, a growing demand concerns high-throughput analyses that remain representative across a broad particle-size spectrum, from the micrometre to the millimetre scale. These requirements become even more challenging when considering the intrinsic complexity of WW matrices, where MPs and ECs occur within dense mixtures of organic debris, inorganic particles, and biofilm-like material. Such heterogeneity imposes practical constraints on both destructive and spectroscopic approaches, affecting their sensitivity, selectivity, and reproducibility. Within this context, several analytical strategies have been proposed, ranging from thermo-analytical methods - such as pyrolysis–GC or thermal desorption–GC, often hyphenated with mass spectrometry - to vibrational spectroscopies including Raman, Fourier Transform Infrared Spectroscopy (FT-IR), and Near Infrared (NIR)[10]. The selection of one method over another depends on the specific information sought and on the scale of the investigation. For instance, pyrolysis–GC–MS is extensively employed for bulk quantification of polymer mass; by generating characteristic thermal fragments, it provides polymer-specific fingerprints and allows the detection of materials such as carbon-black-rich plastics, which are difficult to analyse spectroscopically. Its destructive nature, however, prevents any morphological or particle-level assessment, and the technique is associated with long analysis times[11]. Fourier-transform infrared spectroscopy, in contrast, provides a non-destructive means of identifying polymeric materials through their characteristic vibrational absorption bands. Automated mapping enables the spatial characterization of particles typically between 20 and 500 µm, offering an advantage over pyrolysis-based workflows. Yet its spatial resolution remains insufficient for particles smaller than approximately 50–100 µm, representing a significant limitation when studying fine MPs or nanoplastics[12]. Raman spectroscopy addresses part of this gap by delivering chemically specific spectra with micrometric spatial resolution—often around 1 µm—thereby enabling detailed particle-by-particle analysis. Nonetheless, the technique is sensitive to fluorescence, which is frequently induced by organic-rich matrices typical of WW samples; as a result, suitable pretreatments are often required to obtain interpretable spectra[13]. Hyperspectral imaging (HSI) operates in the visible, near-infrared or mid-infrared ranges and provides spatially resolved spectral information, as each pixel of the acquired image corresponds to a characteristic spectrum. This approach enables rapid screening of large sample areas and facilitates the classification

[10] Talvitie et al., «How Well Is Microlitter Purified from Wastewater?»

[11] Talvitie et al., «How Well Is Microlitter Purified from Wastewater?»

[12] Corami et al., «Small Microplastics (<100 Mm), Plasticizers and Additives in Seawater and Sediments»; Xu et al., «FTIR and Raman Imaging for Microplastics Analysis»; Leung et al., «Improved Raman Spectroscopy-Based Approach to Assess Microplastics in Seafood».

[13] «Characterization of Microplastics by Raman Spectroscopy»; Elert et al., «Comparison of Different Methods for MP Detection»; Dong et al., «Overview of Analytical Methods for the Determination of Microplastics»; Ye et al., «The Development and Application of Advanced Analytical Methods in Microplastics Contamination Detection».

of polymer types with minimal sample preparation. Its main limitation is spatial resolution, which restricts the identification of particles smaller than roughly 100–150 μm[14].
Relying on a single technique is therefore insufficient to capture all descriptors considered essential for comprehensive MPs and ECs characterization. According to a study published by Ye et al. (2022), these descriptors include particle count, total mass, chemical composition, shape, size, size distribution, and colour[15]. For this reason, contemporary research increasingly advocates for the integration of complementary methodologies. In particular, combining the high chemical specificity and micrometric resolution of Raman spectroscopy[16] with the rapid, large-scale screening capabilities of HSI provides a robust and reproducible framework for both qualitative and quantitative assessment of complex WW matrices[17].

### *1.3. Aims and objectives of the research*

The primary objective of this thesis project is to evaluate the effectiveness of a combined approach based on Raman spectroscopy and Hyperspectral Imaging (HSI) for monitoring microplastics (from Crispinjana) and emerging contaminants (from Modena) in wastewater samples. The work aims to assess the extent to which the integration of these two techniques ensures reproducibility, analytical robustness, and accuracy in the characterization of complex matrices typical of wastewater treatment plants. Furthermore, the project seeks to critically compare the spectroscopic methodology adopted, evaluating operational limitations, discriminative power, level of automation, and suitability for monitoring at different scales.
The ultimate goal is to enhance the understanding and implementation of more reliable and reproducible analytical strategies for the qualitative control of persistent pollutants within wastewater treatment processes.

[14] Abimbola et al., «In-Situ Detection of Microplastics in the Aquatic Environment».

[15] Ye et al., «The Development and Application of Advanced Analytical Methods in Microplastics Contamination Detection».

[16] «Characterization of Microplastics by Raman Spectroscopy».

[17] Rocha-Santos e Duarte, «A Critical Overview of the Analytical Approaches to the Occurrence, the Fate and the Behavior of Microplastics in the Environment».

## 2. Scientific Background and Research Context

### 2.1. *Environmental relevance of emerging contaminants and microplastics*

In the introductory chapter, the definitions of microplastics and emerging contaminants were presented and contextualised, outlining the environmental impact associated with the long-term accumulation of these pollutants. However, the mechanisms driving such accumulation, the substances most frequently identified, the related toxicological implications, and the directions currently pursued by scientific research have not yet been addressed[18]. First, it is important to emphasise that the accumulation of plastic materials and contaminant particles is closely interconnected, as MPs can act as a vector capable of adsorbing a wide range of organic substances onto their surfaces[19]. In this context, the term "accumulation" does not strictly refer to the progressive build-up of materials in large quantities[20], but rather describes how the large-scale industrial production of plastic materials and contaminants leads to an exponential increase in their environmental presence, particularly within aquatic systems, which act as the main reservoirs and transport pathways. As a consequence of this now well-documented occurrence, the issue has also extended to marine fauna, prompting scientific research to investigate strategic solutions aimed at protecting wildlife from contamination[21]. As described by Yu et al. in a 2019 study, absorption phenomena are generally influenced by environmental parameters such as pH, salinity, and the presence of organic matter and suspended particulates (OM)[22]. This implies that the interaction between MPs and ECs, particularly in the context of wastewater, represents a growing environmental concern. The key challenge lies in maximising the efficiency of wastewater treatments plants,considering that they are not designed to ensure the complete removal of microplastics, but rather to reduce their concentrations within regulatory limits introduced in 2018, and that they are largely ineffective at retaining organic and inorganic substances classified as emerging contaminants, except for those partially associated with plastic materials through adsorption processes[23].
Until 2018, European legislation did not include specific provisions addressing microplastics and nanoplastics in environmental matrices, particularly within the context of wastewater. Subsequently, the European Commission adopted the *European Strategy for Plastics,* aimed at reducing plastic pollution while simultaneously promoting recycling practices. In addition, this policy materials are designed, produced, used, and recycled across the European Union[24]. In 2019, the European Chemicals Agency (ECHA) introduced specific guidelines to oversee the

[18] Wright et al., «The Physical Impacts of Microplastics on Marine Organisms».

[19] Cavazzoli et al., «Analysis of Micro- and Nanoplastics in Wastewater Treatment Plants».

[20] «assumere significato».

[21] Zhang et al., «Occurrence, Removal, and Ecological Risk Assessment of Emerging Organic Contaminants in an Industrial WWTP».

[22] Yu et al., «Adsorption Behavior of Organic Pollutants and Metals on Micro/Nanoplastics in the Aquatic Environment».

[23] Cavazzoli et al., «Analysis of Micro- and Nanoplastics in Wastewater Treatment Plants».

[24] «plastics-strategy_en».

risks associated with the presence of micro- and nanoplastics materials in commercial products, proposing analytical strategies for their identification across a wide range of matrices. However, despite the fact as early as 2016 the European Commission published the report entitled "Marine Litter and Microplastics: Global Lessons and Research to Inspire Action and Guide Policy Change", there still no specific regulation in place for wastewater treatment plants. At present, regulatory efforts are limited to guidance documents, such as the "Guidance document on the fate and behaviour of microplastics in the aquatic environment", which provide general recommendations rather than binding requirements[25].

Emerging contaminants represent a particularly challenging issue. As previously noted, conventional wastewater treatment plants (WWTPs) are not designed to efficiently remove complex pharmaceuticals, pesticides, or personal care products (PPCPs), as their treatment processes are primarily tailored to the elimination of biodegradable compounds containing carbon, nitrogen, and phosphorus. As a result, these substances frequently traverse the treatment systems largely unaltered effectively rendering WWTPs reservoirs of pollutants[26], which include parent compounds, metabolites, and transformation products, such as thermoplastic styrenics or silicones[27]. Consequently, the identification and classification of PPCPs residues represent primary objectives to support the development of future the environmental impact of these substances[28].

In order to assess the ecotoxicological risks posed by emerging contaminants, given the vast array of compounds encompassed by this category, the European Commission has proposed two objective parameters to estimate their impact: measured environmental concentrations (MECs) and predicted no-effect concentrations (PNECs)[29]. To date, although European scientific research is increasingly addressing this issue, removal rates in WWTPs have been estimated for only a small fraction of the total number of pharmaceuticals present[30].

Proceeding systematically, it is possible to identify which compounds fall within the two broad categories, or at least to define the most common ones, highlighting certain physicochemical properties and their ecotoxicological impact. Starting with microplastics, the compounds most frequently encountered in wastewater treatment plants include: polyethylene (PE),

[25] Cavazzoli et al., «Analysis of Micro- and Nanoplastics in Wastewater Treatment Plants».

[26] Talvitie et al., «Do Wastewater Treatment Plants Act as a Potential Point Source of Microplastics?»

[27] Lopez et al., «Removal Efficiency for Emerging Contaminants in a WWTP from Madrid (Spain) after Secondary and Tertiary Treatment and Environmental Impact on the Manzanares River».

[28] Burns et al., «Temporal and Spatial Variation in Pharmaceutical Concentrations in an Urban River System».

[29] «Decision, E. C. "Commission Implementing Decision (EU) 2020/1161 of 4 August 2020 establishing a watch list of substances for Union-wide monitoring in the field of water policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council." European Comission (2020).»

[30] Gracia-Lor et al., «Occurrence and Removal of Pharmaceuticals in Wastewater Treatment Plants at the Spanish Mediterranean Area of Valencia»; Botero-Coy et al., «'An Investigation into the Occurrence and Removal of Pharmaceuticals in Colombian Wastewater'»; Burns et al., «Temporal and Spatial Variation in Pharmaceutical Concentrations in an Urban River System»; Boxall et al., «Exploiting Monitoring Data in Environmental Exposure Modelling and Risk Assessment of Pharmaceuticals».

polypropylene (PP), polyvinyl chloride (PVC), polyethylene terephthalate (PET), polystyrene (PS), and polycarbonate (PC).

- Polyethylene (PE): PE is the most widely produced polymer worldwide. It has a density slightly lower than that of water, which allows it to float; additionally, it is highly hydrophobic and resistant to chemical agents. In the environment, PE accumulates persistent organic pollutants (POPs) on its surface through adsorption processes. Moreover, it can cause intestinal blockages in marine fauna.
- Polypropylene (PP): PP is mainly used in the production of caps, food containers, and fishing nets. Its low density makes it the lightest among common polymers, and it exhibits high resistance to heat. Ecotoxicologically, UV radiation fragments PP into secondary plastics, and its fibers are among the most abundant found in the intestines of mammals.
- Polyvinyl chloride (PVC): PVC is widely employed in construction, particularly for pipes and cables, and in the medical sector. It is one of the heaviest common plastics, often due to the addition of plasticizers such as phthalates. Environmental accumulation can lead to the release of toxic additives and chlorine into deep sediments, causing endocrine disruption in organisms living in close contact with the seabed.
- Polyethylene terephthalate (PET): PET is a key material for the production of plastic bottles and synthetic fibers. It has a high density, approximately 1.30 g/cm$^3$, as well as high tensile strength (the maximum stress a material can withstand) and resistance to biodegradation. In wastewater, PET fibers are the most abundant and are primarily released through textile washing. Being readily ingested by filter-feeding organisms, PET easily enters the human food chain. Prolonged exposure can cause chronic damage to the digestive system.
- Polystyrene (PS): PS is commonly used as thermal insulation and protective packaging. Its physicochemical properties are distinctive: when expanded, it is extremely lightweight, while in solid form, it has a density slightly higher than that of water. Its ecotoxicological impact is of particular concern compared to other common MPs, as it releases styrene, a substance highly toxic to the central nervous system. PS is also highly prone to fragmentation, promoting plankton ingestion and thereby disrupting the base of the food chain.
- Polycarbonate (PC): PC is widely used in electronics and the medical sector, as well as in the production of eyeglass lenses. It is characterized by extreme rigidity and transparency, combined with high mechanical and thermal resistance, and a density greater than that of water. The primary environmental risk associated with PC accumulation lies in the release of Bisphenol A (BPA), a monomer used in its synthesis. BPA is a potent endocrine disruptor capable of affecting embryonic development and the reproductive systems of fish and amphibians, even at very low concentrations. By sinking, PC microplastics accumulate on the seabed, contaminating marine sediments

and deep waters. Consequently, the impact of Bisphenol A is closely monitored by both health and environmental authorities[31].

With regard to emerging contaminants, their classification is not only broader in terms of the range of compounds included but also highly complex. By definition, emerging contaminants are unregulated, novel, or re-emerging substances, detectable at trace levels, which raise concerns for human health and the environmental due to their persistence, bioactivity, and potential toxicity[32]. Nevertheless it is possible to define broad internal categories based on their physicochemical properties, field of application, and origin:

- Per- and polyfluoroalkyl substances (PFAS): PFAS have been industrially produced since the 1940s and are primarily used in the manufacture of firefighting foams, non-stick cookware coatings, technical textiles, and, to a lesser extent, food packaging. The carbon-fluorine (C–F) bond - the strongest in organic chemistry - confers extreme chemical stability. Additionally, PFAS are amphiphilic, being both water- and oil-repellent, highly persistent, and exhibiting high mobility in water. Ecotoxicologically, PFAS tend to bioaccumulate along the food chain, from crustaceans to top predators. In mammals, they can induce hepatotoxicity, immune system disruption, and reproductive interference.
- Pharmaceuticals and hormones (PPCPs): These compounds are primarily produced for domestic and hospital use and enter marine ecosystems through wastewater effluents and agricultural runoff from intensive livestock operations. They are generally hydrophilic and may carry an electrical charge at physiological pH. They are also designed to maintain high biological activity even at concentrations on the order of ng/L. While their direct ecotoxicological impact is relatively moderate, prolonged exposure can cause chronic disturbances.
- UV filters (Oxybenzone): Found mainly in cosmetic products such as sunscreens and skin care formulations, these compounds are lipophilic and capable of absorbing and dissipating ultraviolet radiation. Their accumulation is the primary cause of coral “bleaching,” and they induce oxidative stress, ultimately leading to the death of coral reef ecosystems.
- Brominated flame retardants (BDEs): These additives are incorporated in electronic circuits, computer plastics, and furniture upholstery to reduce flammability. Among their physicochemical properties, high lipophilicity stands out, promoting accumulation in the adipose tissues of organisms. Consequently, BDEs are neurotoxic, causing developmental neurological dysfunction, particularly in children.

Overall, the high chemical heterogeneity of emerging contaminants, combined with the role of microplastics as vectors and reservoirs of bioactive substances, makes environmental monitoring particularly urgent. Wastewater treatment plants, in particular, play a key role

[31] Jovanović, «Ingestion of Microplastics by Fish and Its Potential Consequences from a Physical Perspective».

[32] Zhang et al., «Occurrence, Removal, and Ecological Risk Assessment of Emerging Organic Contaminants in an Industrial WWTP».

both as monitoring points and as potential secondary sources of pollution, highlighting the limitations of traditional supervision approaches. Consequently, understanding the ecotoxicological impact and environmental fate of these contaminants requires the application of advanced analytical methodologies capable of providing reliable chemical - especially concentration - and structural information, while ensuring high reproducibility in the analysis of complex matrices.

### 2.2. *Wastewater as a complex analytical matrix (overview of current monitoring strategies and analytical challenges related to contaminant detection in wastewater).*

The study of pollutant substances in complex matrices for the monitoring of wastewater treatment plants (WWTPs) represents, as highlighted in the previous sections, a contemporary challenge that is strongly driving scientific research. To date, owing to the growing global interest in this topic, the number of publications addressing the determination of microplastics (MPs) and emerging contaminants (ECs) in aquatic environments has increased exponentially; however, research specifically focused on WWTPs remains relatively recent and still underdeveloped.

Since approximately 2015, numerous scientific research groups have focused on the investigation of matrices derived from wastewater treatment plants, with the aim of optimizing the various stages of analytical determination. These include sampling, sample treatment - such as filtration, density separation, and digestion - and the selection of the most appropriate analytical technique. This choice strongly depends on the specific objectives of the study and, as extensively reported in the literature, generally involves either spectroscopic techniques, most notably Raman and Fourier.transform infrared (FTIR) spectroscopy, or chromatographic approaches, particularly pyrolysis - gas chromatography - mass spectrometry (Py-GC-MS). To date, the most widely adopted strategy consists of a combined approach,starting with visual inspection by optical microscopy, followed by chemical analysis of the polymeric composition using spectroscopic or thermo-analytical techniques[33].

Starting from the observation of samples by optical microscopy, this stage is inherently prone to error, as it does not provide a fully reliable or reproducible characterization and should not be regarded as a critical step for achieving the final analytical objective. Nevertheless, it represents an important preliminary step, as it significantly facilitates the subsequent point-by-point spectroscopic investigation in terms of analytical speed. Optical microscopy allows for the approximate identification of potential microplastics based on their shape and color[34]. Clearly, during this phase, any precaution adopted or additional parameter considered may enhance the quality of the observation, thereby improving overall accuracy.

Visual identification is an appropriate approach when dealing with large sample volumes or in situations where access to costly analytical instrumentation is limited. However, a potential bias is always present during the observation stage; moreover, the quality of the identification

[33] Wang e Wang, «Investigation of Microplastics in Aquatic Environments».

[34] Crawford e Quinn, «Microplastic Identification Techniques» (Elsevier, 2017).

results depends on several factors, including operator subjectivity, matrix composition, and the shape and size of the particles[35].

Turning to spectroscopic techniques, the literature reports numerous studies in which Raman and FTIR spectroscopy are cited as the primary analytical methods for the characterization of microplastics and emerging contaminants, including within the context of wastewater matrices.

FTIR spectroscopy generates a characteristic infrared spectrum for a specific chemical bond. Since different materials exhibit distinct bond compositions, unknown substances can be identified through comparison with reference spectra of known or pure materials. Owing to its high reliability, FTIR has become one of the most widely employed techniques for the characterization of microplastics in environmental contexts [36].

In monitoring programs for MPs and ECs, FTIR is employed in two main ways: either by scanning all suspected particles[37] or by analyzing a set of subsamples to validate the results of the visual identification previously performed using optical microscopy[38]. FTIR is not only capable of accurately identifying polymer types, but it also provides additional information on the physicochemical alteration of MPs by assessing their degree of oxidations, which is useful for evaluating their environmental impact in terms of susceptibility to fragmentation[39]. However, FTIR is capable of identifying polymer composition only for microplastics with sizes larger than approximately 10–20 µm, and its applicability may be lost when the target particles are smaller than this threshold[40]. Nevertheless, FTIR remains a promising technique for polymer identification and is also applied to the analysis of emerging contaminants, particularly in the context of wastewater treatment plant monitoring. Optimized FTIR-based technologies, such as micro–FTIR spectroscopy (µ-FTIR), attenuated total reflectance FTIR (ATR-FTIR), and focal plane array FTIR spectroscopy (FPA-FTIR), have been increasingly employed worldwide in research on plastic contaminants. In particular, µ-FTIR facilitates the detection of smaller particles (>10 µm), while ATR-FTIR allows the direct analysis of larger particles (>500 µm) with irregular shapes, without the need for extensive sample preparation[41]. Finally, FPA-FTIR enables high-throughput, unbiased analysis of all plastic particles (>20 µm) through scanning with a high degree of lateral resolution[42]. Transmission-mode FTIR imaging allows for the simultaneous physical characterization of analyzed particles, and is therefore increasingly

---

[35] Song et al., «A Comparison of Microscopic and Spectroscopic Identification Methods for Analysis of Microplastics in Environmental Samples»; Bergmann et al., *Marine Anthropogenic Litter*.

[36] Silva et al., «Microplastics in the Environment»; Shim et al., «Identification Methods in Microplastic Analysis», 2017.

[37] Browne et al., «Accumulation of Microplastic on Shorelines Worldwide», 2011.

[38] Wang et al., «Microplastics Pollution in Inland Freshwaters of China».

[39] Corcoran et al., «Plastics and Beaches».

[40] Huppertsberg e Knepper, «Instrumental Analysis of Microplastics—Benefits and Challenges».

[41] Löder e Gerdts, «Methodology Used for the Detection and Identification of Microplastics—A Critical Appraisal».

[42] Tagg et al., «Identification and Quantification of Microplastics in Wastewater Using Focal Plane Array-Based Reflectance Micro-FT-IR Imaging».

being applied to the detection and investigation of MPs in wastewater treatment plants samples[43]. Compared to conventional FTIR, FPA-FTIR is significantly faster in detecting microplastics; however, the instrumentation is more expensive and demands high computational power[44].

Raman spectroscopy, like FTIR, is widely employed for monitoring polymers and emerging contaminants in various environmental matrices[45]. Theoretically, Raman spectroscopy involves irradiating the target sample with a monochromatic beam, which generates backscattered light at a different frequency as a result of absorption, scattering, or reflection by the molecular structure and atomic composition of the sample[46]. This frequency shift produces a vibrational spectrum characteristic of each sample. Consequently, Raman spectroscopy enables non-destructive chemical characterization of MPs[47]. However, its applicability to ECs is more limited due to structural complexity, extremely low contaminants, and the scarce availability of reference spectra in the literature. Like FTIR, Raman spectroscopy offers high reliability, the possibility for high-throughput screening, and minimal sample pretreatment[48]. Additionally, it provides advantages in terms of higher spatial resolution, a border spectral range, and lower sensitivity to water interference - an essential feature in the context of wastewater treatment plant monitoring[49]. Furthermore, combining Raman spectroscopy with microscopy (micr-Raman) allows for the identification of MPs down to size of 1 μm[50]. It is feasible to obtain spatially resolved chemical images of entire samples with sub-micrometer resolution using micro-Raman spectroscopy coupled with Raman spectral resolution using micro-Raman spectroscopy coupled with Raman spectral imaging systems[51]. The main limitation of the techniques arises from interferences due to additives, pigments, or chemical substances associated with polymers and pollutant residues, which can negatively affect accuracy[52]. However, in the context of MPs analysis, this issue is largely mitigated by the wide availability of reference spectra in the literature - particularly for the most common polymers previously discussed. Moreover, Raman spectroscopy inherently exhibits a low signal-to-noise ratio, which can complicate spectral analysis[53]. Nevertheless, these

[43] Crawford e Quinn, «Microplastic Identification Techniques» (Elsevier, 2017); Elert et al., «Comparison of Different Methods for MP Detection».

[44] Huppertsberg e Knepper, «Instrumental Analysis of Microplastics—Benefits and Challenges».

[45] Araujo et al., «Identification of Microplastics Using Raman Spectroscopy».

[46] Crawford e Quinn, «Microplastic Identification Techniques» (Elsevier, 2017).

[47] Shim et al., «Identification Methods in Microplastic Analysis», 2017.

[48] Araujo et al., «Identification of Microplastics Using Raman Spectroscopy».

[49] Käppler et al., «Analysis of Environmental Microplastics by Vibrational Microspectroscopy»; Araujo et al., «Identification of Microplastics Using Raman Spectroscopy».

[50] Crawford e Quinn, «Microplastic Identification Techniques» (Elsevier, 2017).

[51] Löder e Gerdts, «Methodology Used for the Detection and Identification of Microplastics—A Critical Appraisal».

[52] Huppertsberg e Knepper, «Instrumental Analysis of Microplastics—Benefits and Challenges».

[53] Araujo et al., «Identification of Microplastics Using Raman Spectroscopy».

challenges do not preclude Raman spectroscopy from being a powerful analytical technique in microplastic and emerging contaminants research.

Pyrolysis - gas chromatography - mass spectrometry (Py-GC-MS) represented the chromatographic alternative for the anaòysis of environmental matrices through the characterization of thermal degradation products. It was widely employed prior to the advent of spectroscopic techniques for contaminant investigation[54]. The technique offers practical advantages that allow trace-level analysis of polymers while minimizing sample pretreatment. Moreover, it enables the simultaneous acquisition of detailed information on the chemical composition of both polymers and associated organic additives, without interference from inorganic matter in the generation of pyrograms[55]. Another major advantage of the technique lies in its ability to operate with a small sample amount, approximately 100–500 µg, while ensuring high reproducibility and accuracy of the resulting chromatograms. These chromatograms, when compared with reference materials reported in the literature, allow for straightforward polymer identification[56].

The major limitations of chromatographic techniques are associated with both the long analysis time (30/100 minutes) and the possibility of analyzing only a single particle per run, in addition to being a destructive method that does not allow recovery of the solid sample inserted into the tube[57]. IFurthermore, practical constraints exist during sample introduction, as only sufficiently large analytes (> 100 µm) can be loaded. For this reasons, a less restrictive variant is sometimes preferred, such as thermal extraction-desorption gas chromatography - mass spectrometry (TED-GC-MS). This technique combines thermal extraction with thermogravimetric analysis and thermal desorption with gas chromatography coupled to mass spectrometry, allowing, however, the characterization of a limited set of polymers or contaminants[58]. These analytical limitations make chromatography a complementary technique to spectroscopy for achieving a comprehensive and integrated analysis.

Nowadays, thanks to the development of new technologies and the aim of maximizing the number of particles analyzed in the shortest possible time, scientific research is testing an alternative analytical approach to the two spectroscopic techniques previously mentioned: hyperspectral imaging (HSI). Hyperspectral imaging involves the study of light-material interactions and represents an analytical process that combines spectroscopy with imaging. The hyperspectral cube, which contains both spectral and spatial information for each pixel captured by a two-dimensional (2D) detector array[59]. This spatial information identities the

[54] Fries et al., «Identification of Polymer Types and Additives in Marine Microplastic Particles Using Pyrolysis-GC/MS and Scanning Electron Microscopy».

[55] Käppler et al., «Comparison of µ-ATR-FTIR Spectroscopy and Py-GCMS as Identification Tools for Microplastic Particles and Fibers Isolated from River Sediments».

[56] Dümichen et al., «Fast Identification of Microplastics in Complex Environmental Samples by a Thermal Degradation Method».

[57] Wang e Wang, «Investigation of Microplastics in Aquatic Environments».

[58] Käppler et al., «Comparison of µ-ATR-FTIR Spectroscopy and Py-GCMS as Identification Tools for Microplastic Particles and Fibers Isolated from River Sediments».

[59] Bhargava et al., «Hyperspectral Imaging and Its Applications».

source of each spectrum within the samples, enabling a more accurate assessment of the ambient lighting conditions. Moreover, HSI can cover a continuous spectral range of light with multiple spectral bands while maintaining high spectral resolution. For this reason, HSI is capable of capturing spectral differences under varying two-dimensional environmental conditions. The distinction between hyperspectral cubes and RGB images lies in the fact that, whereas a camera - and thus an RGB image - has only three bands at known wavelengths (red, green, and blue), hyperspectral imaging acquires images across a much broader range of wavelengths. Additionally, the narrower the frequency bands, the better the spectral resolution. The hyperspectral cube represents a three-dimensional database of two-dimensional images at each wavelength[60].

[60] Bhargava et al., «Hyperspectral Imaging and Its Applications».

## 3. Materials and methods

### *3.1. Data Analysis and chemometrics in spectroscopy.*

The spectroscopic techniques employed in this study, namely Raman spectroscopy, portable Raman spectroscopy, near-infrared (NIR) hyperspectral imaging and micro-NIR spectroscopy, generate high-dimensional datasets characterised by substantial informational complexity, as well as by non-negligible contributions from noise and variability arising from both instrumental factors and sample matrix effects. In the context of wastewater analysis, this complexity is further amplified by the pronounced heterogeneity of the samples, which typically include organic residues, inorganic particulate matter and components of biological origin. Under these conditions, a direct interpretation of raw spectral data is not sufficient to obtain reliable and reproducible information on the chemical composition of the analysed samples.

For this reason, data analysis was carried out using a chemometric approach, adopted as a fundamental tool for the systematic processing and interpretation of multivariate spectroscopic datasets. In particular, the application of chemometric techniques allows an improvement in the informational quality of the data, a reduction in dimensionality, and the identification of patterns and chemical differences that would not be readily observable through conventional univariate analysis. Prior to the application of spectral preprocessing procedures and multivariate analyses, the data acquired by Raman spectroscopy, portable Raman spectroscopy, near-infrared (NIR) hyperspectral imaging and micro-NIR spectroscopy were subjected to a preliminary phase of data organisation and structuring within the MATLAB (MathWorks, Inc., Massachusetts USA) environment. This step did not involve any prior selection or exclusion of spectra, but was exclusively aimed at converting the raw data files into a mathematically tractable format compatible with subsequent chemometric analyses, particularly through the use of the HyperTools3 (Free Graphical User Interface Hyperspectral and Multispectral Image Analysis) and PLS-Toolbox toolboxes (Eigenvector Research, Inc.,WA 98831, USA).

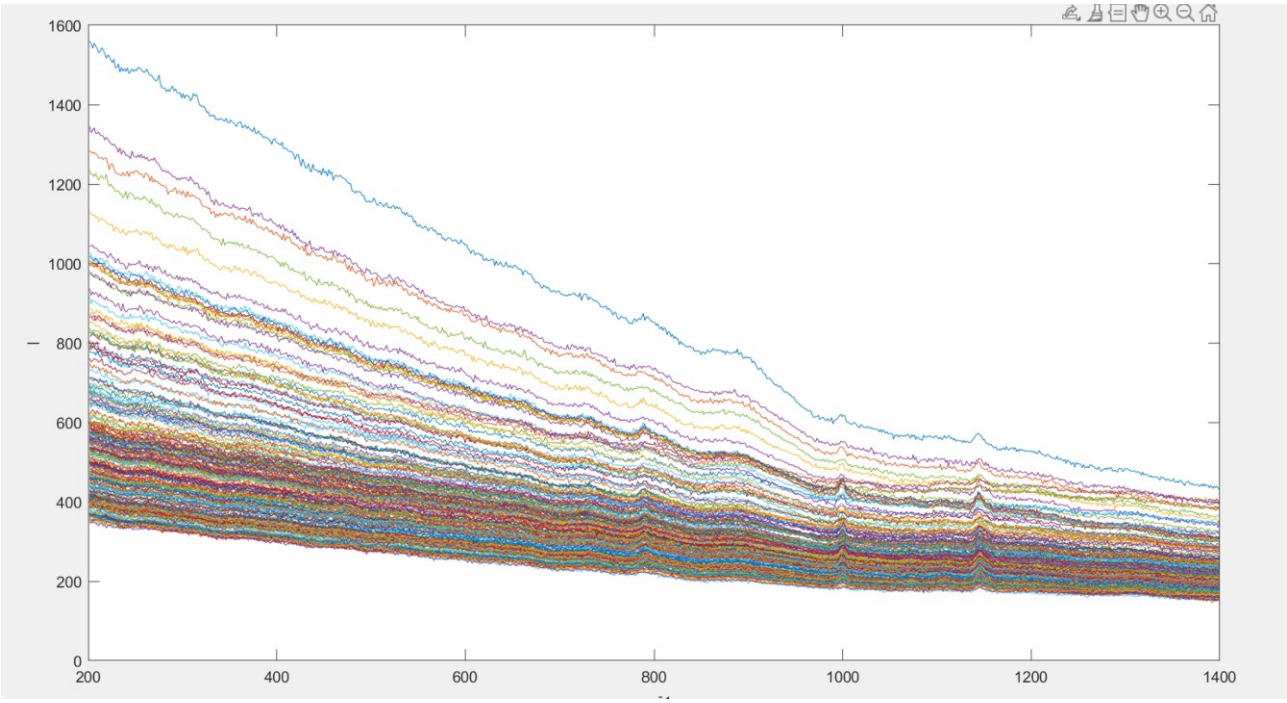
In the case of Raman spectra, the data acquired and stored in textual formats (.txt or .csv) were imported into MATLAB using dedicated scripts specifically developed for the automated file reading and extraction of spectral variables (wavenumber and intensity). The spectra were subsequently reorganised into two-dimensional matrices, where rows correspond to individual acquired spectra and columns represent the spectral variables. This data structuring enabled a coherent and reproducible treatment of the datasets, avoiding subjective interventions on the original data and ensuring full traceability throughout the entire analytical workflow. The same data organisation strategy was applied to the portable Raman and micro-NIR datasets, whose spectra were imported, structured, and processed in MATLAB prior to chemometric analysis. Similarly, near-infrared hyperspectral data were initially handled as three-dimensional datasets (x, y, λ) before being reorganised into two-dimensional matrices suitable for subsequent multivariate analysis. The hyperspectral cubes, consisting of

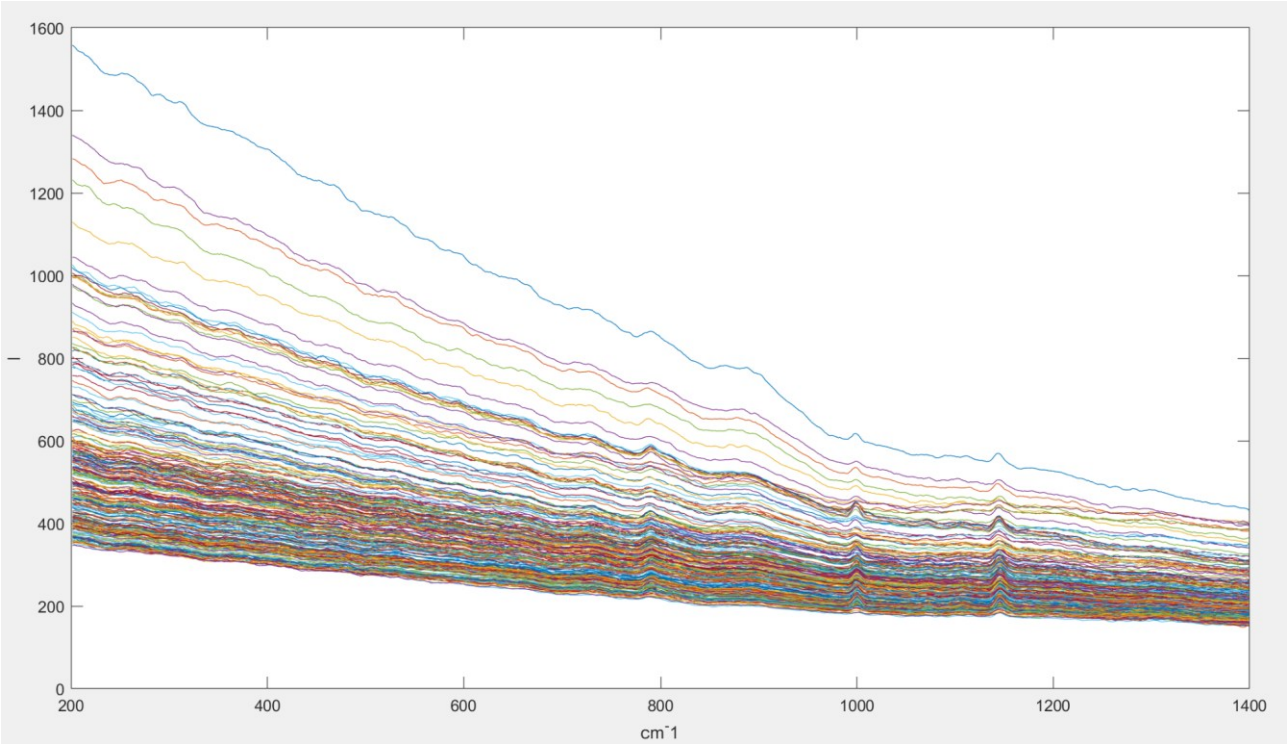
spatial pixels each associated with an NIR spectrum, were imported into MATLAB and initially handled as three-dimensional data structures (x, y, λ). Subsequently, in accordance with the operational requirements of HyperTools3, the data were organised into two-dimensional matrices, where each row corresponds to the spectrum of a single pixel and each column to a spectral variable. This transformation enabled the uniform application of preprocessing procedures and multivariate analyses across the entire dataset, while preserving the link to the original spatial information. Once the data structuring phase was completed, all datasets were subjected to spectral preprocessing aimed at reducing the influence of non-informative contributions of instrumental or matrix origin and at improving the comparability among spectra.

*3.1.1. Modena treatment data*

For the Raman data acquired using the Horiba instrument, the preprocessing workflow included a denoising step performed through a Savitzky–Golay filter (polynomial order = 2, window width = 9), followed by baseline correction in order to minimise the contribution of background fluorescence typically associated with environmental samples. Baseline correction was carried out using the asymmetric least squares (ALS) method with the following parameters: $\lambda = 1 \times 10^5$, $p = 0.01$ and $d = 2$. A specific feature of the Horiba acquisition system is the division of the investigated Raman range into two independent spectral regions, corresponding to 200–1400 $cm^{-1}$ and 1400–2000 $cm^{-1}$. Consequently, the two spectral intervals were treated as separate datasets throughout the entire data-processing workflow. The original Raman map consisted of 265 acquisition points; however, prior to preprocessing, the first spectrum of each spectral region was removed because it exhibited anomalous behaviour and a markedly higher contribution from background noise compared with the remaining spectra. Therefore, all subsequent analyses were performed on datasets composed of 264 spectra. Each spectral interval was independently subjected to denoising, baseline correction and probabilistic quotient normalization (PQN), which was selected instead of L2 normalization in order to reduce intensity variability and improve spectral comparability. Following preprocessing, the two spectral regions were maintained as independent datasets and separately analysed by Multivariate Curve Resolution (MCR)[61], allowing the spectral information contained within each Raman interval to be evaluated independently and avoiding potential artefacts associated with spectral merging. Representative examples of the preprocessing workflow and the resulting spectral datasets are reported in Figure 1 and Figure 2.

[61] Jaumot et al., «MCR-ALS GUI 2.0».


1600
1400
1200
1000
I
800
600
400
200
0
200
400
600
800
1000
1200
1400


1600
1400
1200
1000
I
800
600
400
200
0
200
400
600
800
1000
1200
1400
cm-1

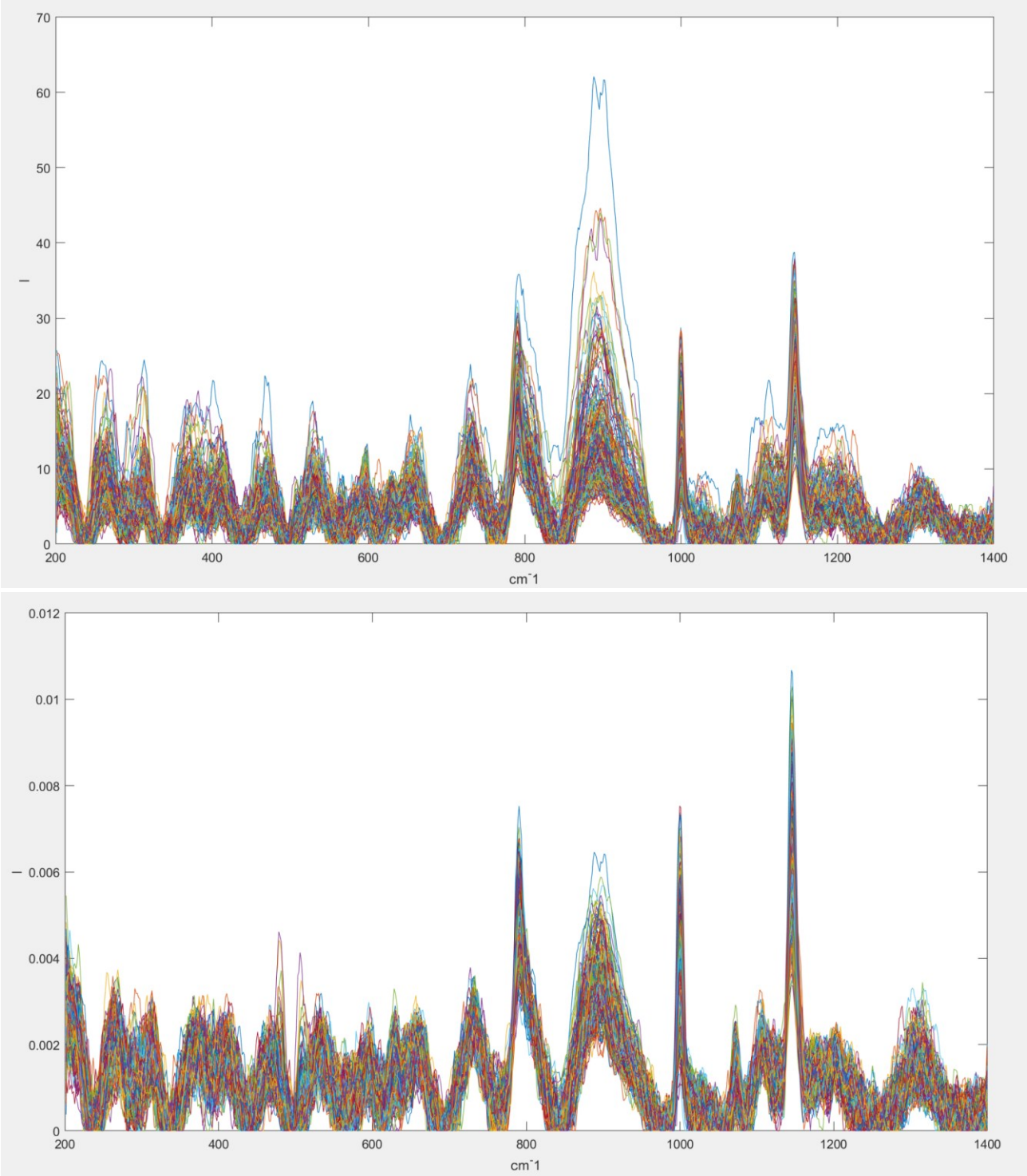
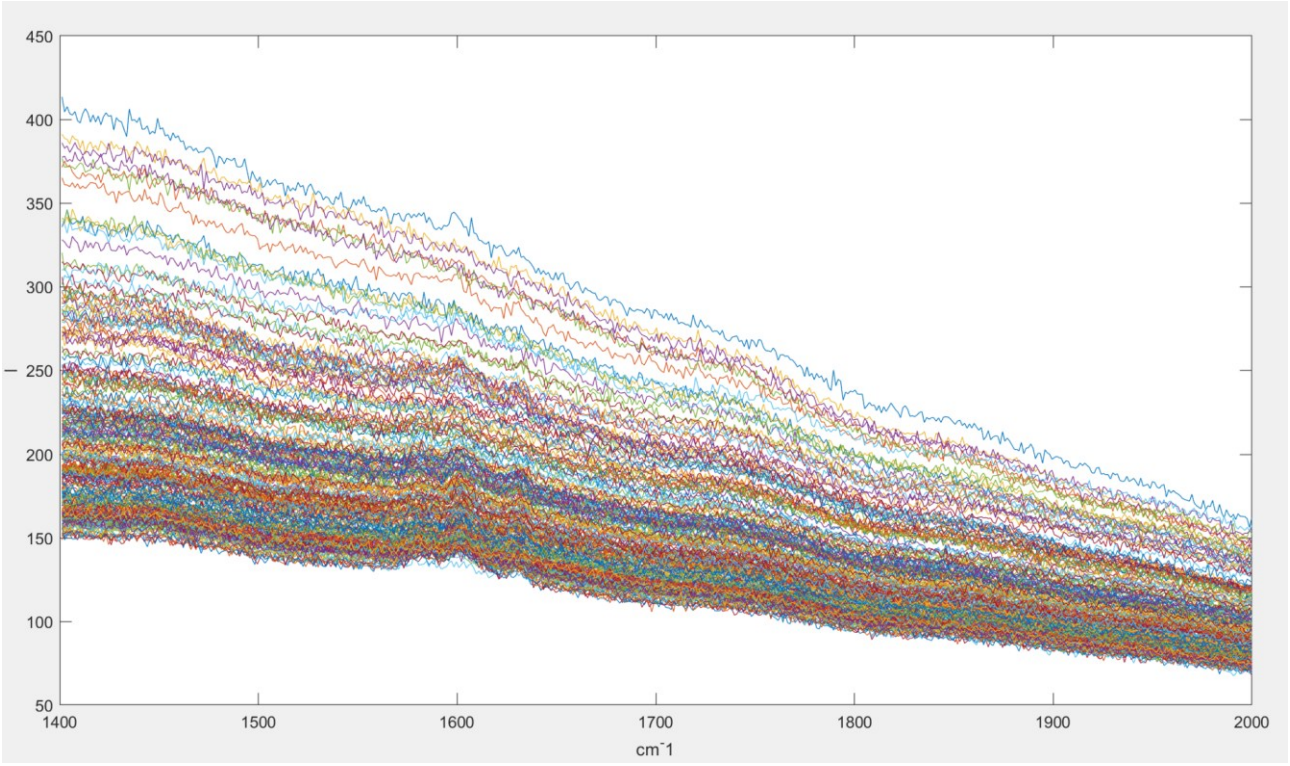
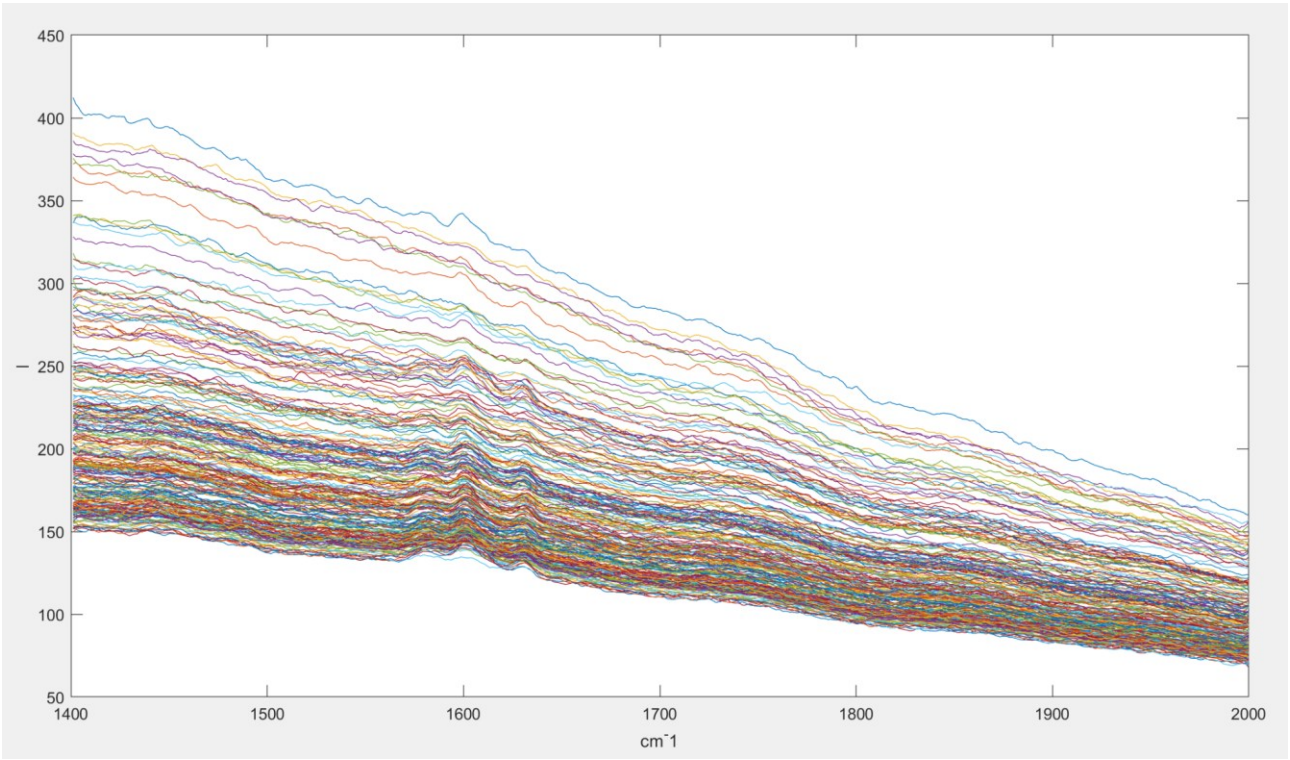



Figure 1. Preprocessing first spectral range 200 $cm^{-1}$ -1400 $cm^{-1}$: the first is the unpreprocessed spectra, followed by denoised spectra, then the baseline correction and after all the normalized spectra.


450
400
350
300
250
200
150
100
50
I
1400
1500
1600
1700
1800
1900
2000
cm^-1


450
400
350
300
250
200
150
100
50
I
1400
1500
1600
1700
1800
1900
2000
cm^-1

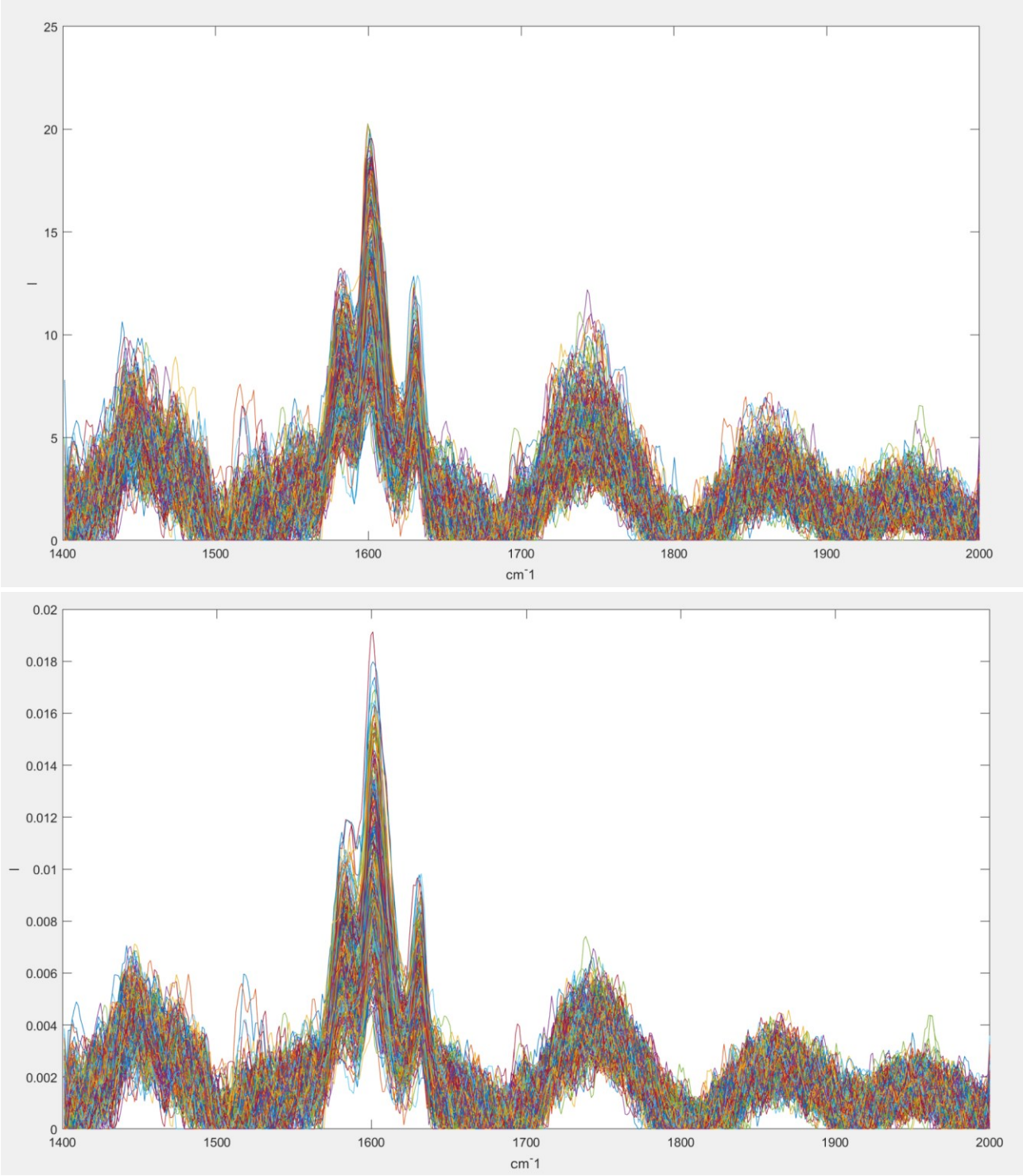
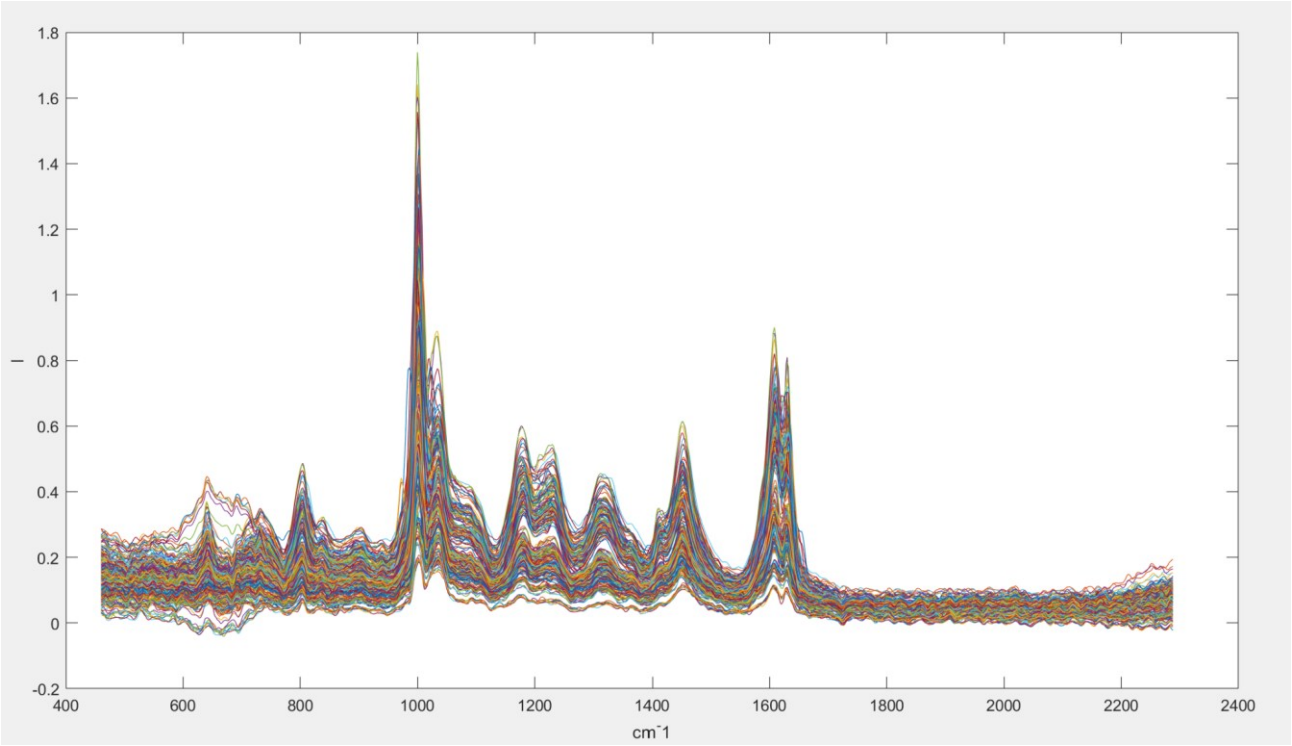
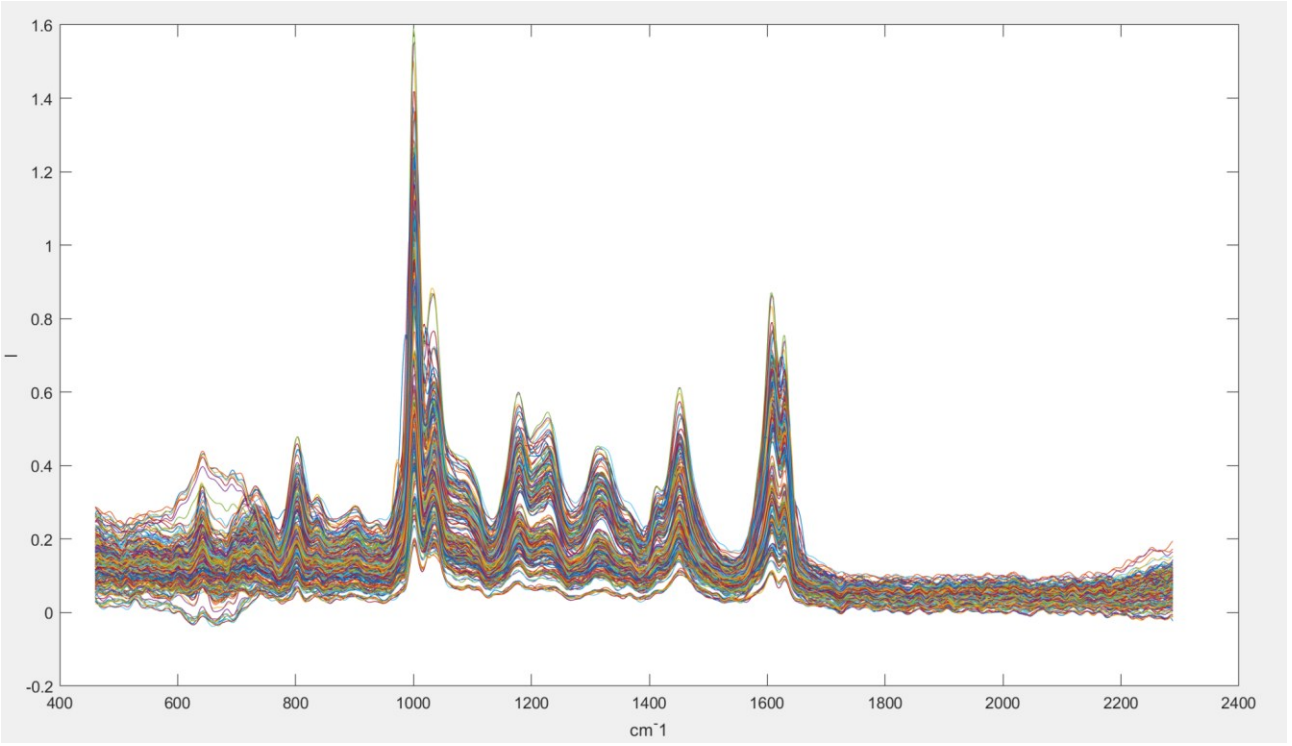



Figure 2. Preprocessing second spectral range 1400 $cm^{-1}$ - 2000 $cm^{-1}$: the first is the unpreprocessed spectra, followed by denoised spectra, then the baseline correction and after all the normalized spectra.

For the portable Raman datasets, spectral acquisitions were performed on sealed plastic bags containing the same sorbent material used in the POCIS devices, previously spiked with five analytical standards at known concentrations according to an experimental design developed on five concentration levels (2, 5, 8, 11, 14 ppm). In order to improve the reproducibility of the measurements and account for local heterogeneity within the samples, spectra were collected

from three fixed positions on each sachet, using three replicates for each point. An important aspect of the acquisition procedure concerns the background spectra automatically recorded by the portable Raman instrument during the start-up phase. Since the measurements were carried out in three different acquisition sessions, three independent background spectra were generated. During preprocessing, each background spectrum was specifically subtracted from the corresponding acquisition block in order to minimize instrumental contributions and ensure greater consistency among the datasets collected at different times. After background subtraction, the individual spectral matrices were merged into a single dataset and subjected to the same preprocessing workflow previously applied to the Horiba Raman maps. In particular, preprocessing included a denoising step using a Savitzky–Golay filter, baseline correction and Probabilistic Quotient Normalization (PQN). This approach allowed a consistent treatment of all Raman datasets while improving spectral comparability and reducing non-informative variability associated with instrumental drift or acquisition conditions. Representative examples of the preprocessing procedure and the resulting spectra are reported in the following Figure 3.


1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
I
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
cm^-1


1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
I
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
cm^-1

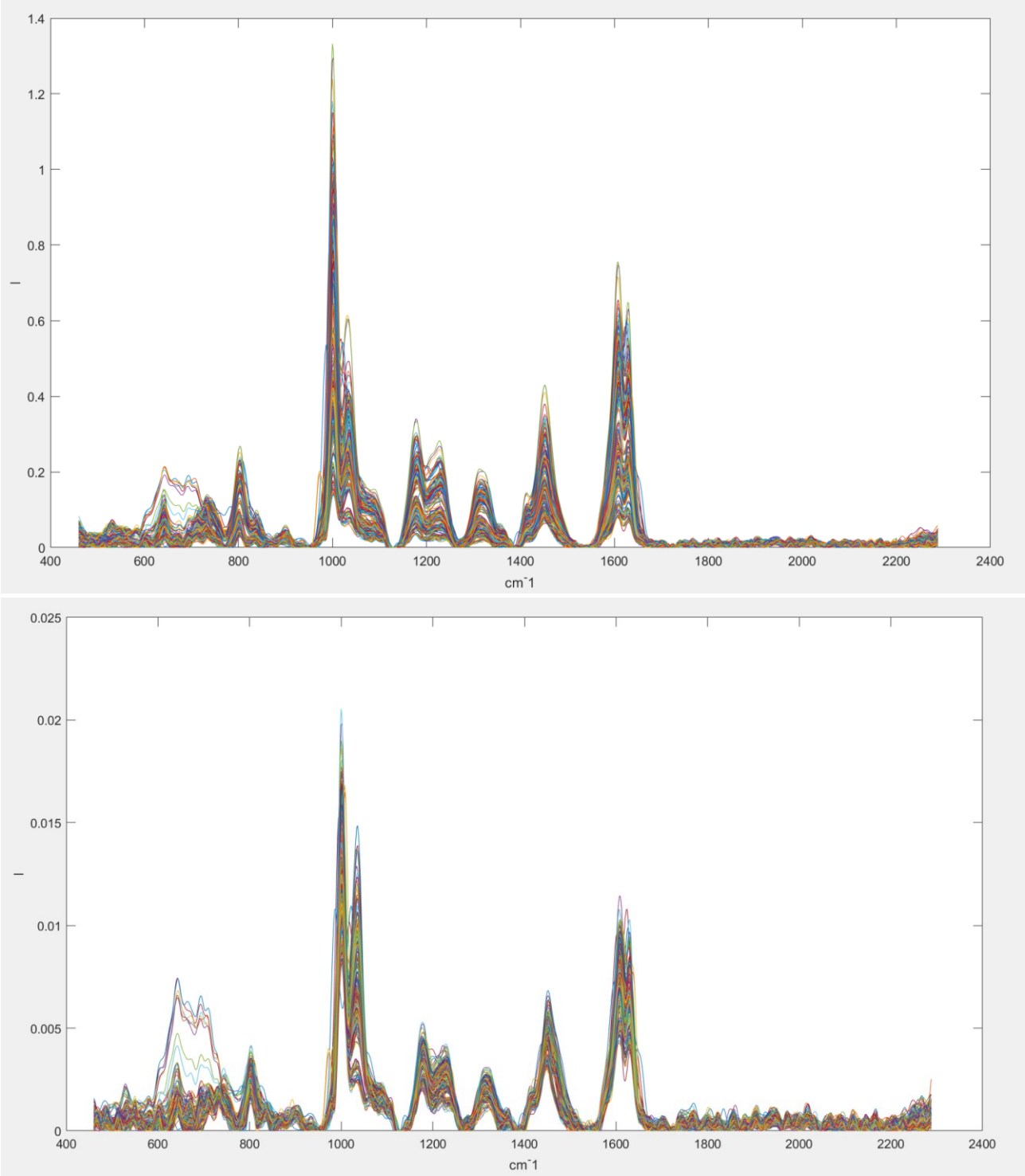



Figure 3. Background subtraction and preprocessing portable Raman spectra:background subtraction (first photo), smoothing SG (second photo), baseline correction (third photo), PQN normalization (last photo).

For the microNIR datasets, spectral preprocessing was performed after data acquisition in order to reduce non-informative variability and improve the comparability among samples. The instrumental setup and acquisition parameters are discussed in the dedicated section. Spectral acquisitions were carried out on two different sample surfaces: plastic sachets containing the sorbent resin spiked with the five analytical standards and Petri dishes prepared

for complementary analyses. In both cases, five spectra were acquired for each sample, followed by the acquisition of five blank spectra consisting of the sorbent resin only. The preprocessing workflow (Figure 4 and Figure 5) included an initial baseline correction performed using the weighted thresholding method with the following parameters: $\lambda = 100$ and $p = 0.001$. This step was necessary to minimise baseline distortions and broad background contributions commonly observed in NIR spectra. Subsequently, Standard Normal Variate (SNV) normalisation was applied in order to compensate for intensity variations associated with scattering effects and small differences in sample surface morphology. For exploratory multivariate analysis, the preprocessed datasets were additionally subjected to mean centering prior to Principal Component Analysis (PCA). This operation allowed the data variance to be centred around the average spectral profile, improving the interpretation of score and loading distributions during chemometric analysis.

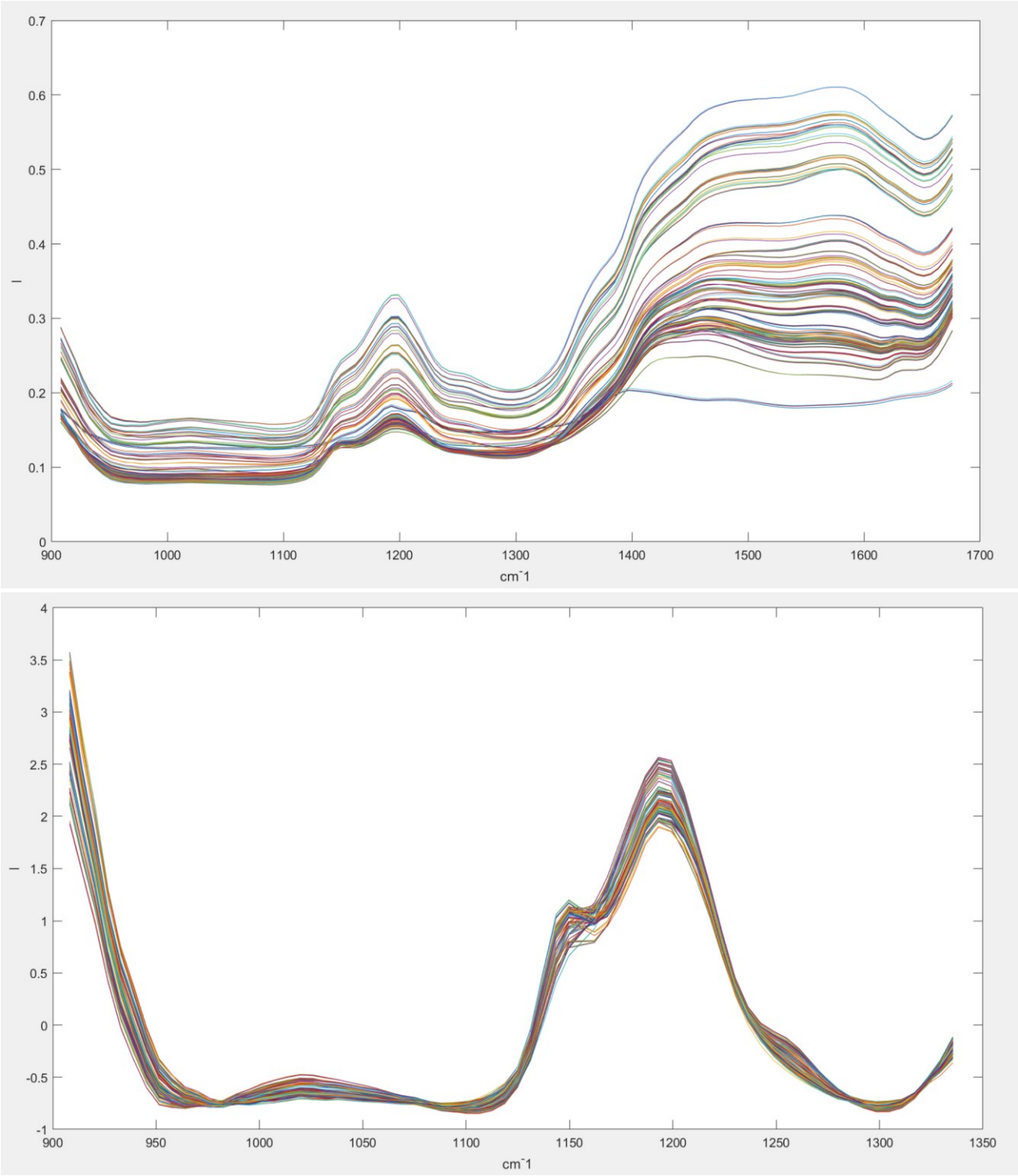



Figure 4. Preprocessing portable Raman spectra (sachets): above the unpreprocessed spectra and then preprocessed spectra.

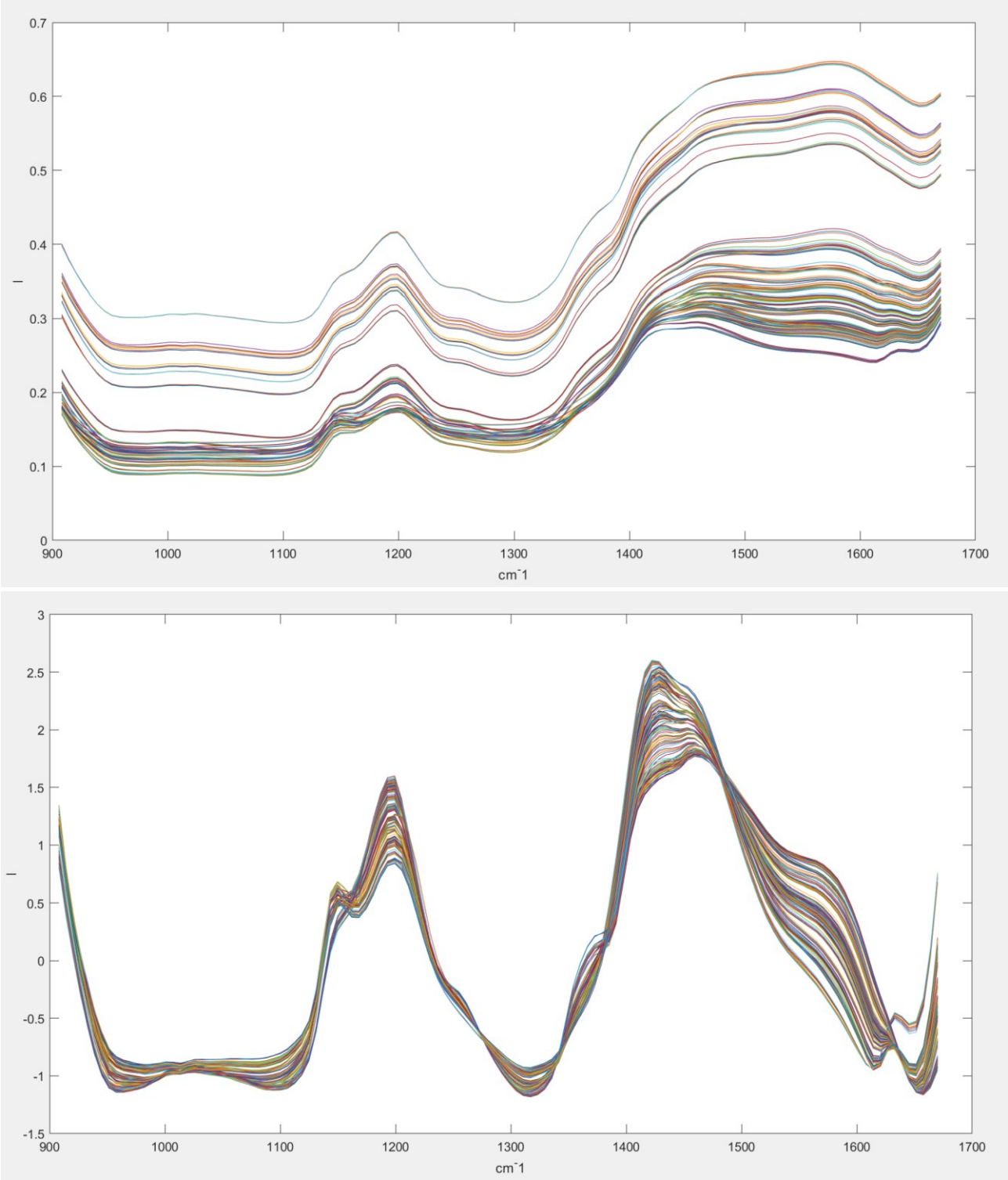



Figure 5. Preprocessing portable Raman spectra (petri dishes): above the unpreprocessed spectra and then preprocessed spectra.

Similarly, the NIR hyperspectral data were subjected to a dedicated preprocessing workflow aimed at reducing non-informative spectral contributions and improving the comparability among spectra. Baseline correction was first performed using the Wavelet Transform method with $\lambda = 100$ and $p = 0.001$, in order to minimize baseline distortions and low-frequency

variations unrelated to the chemical composition of the samples. Subsequently, Standard Normal Variate (SNV) normalization was applied to compensate for multiplicative scattering effects and intensity variations arising from differences in sample morphology and surface heterogeneity. The combination of baseline correction and SNV normalization improved the overall quality of the spectral data and facilitated the subsequent chemometric analyses (Figure 6).

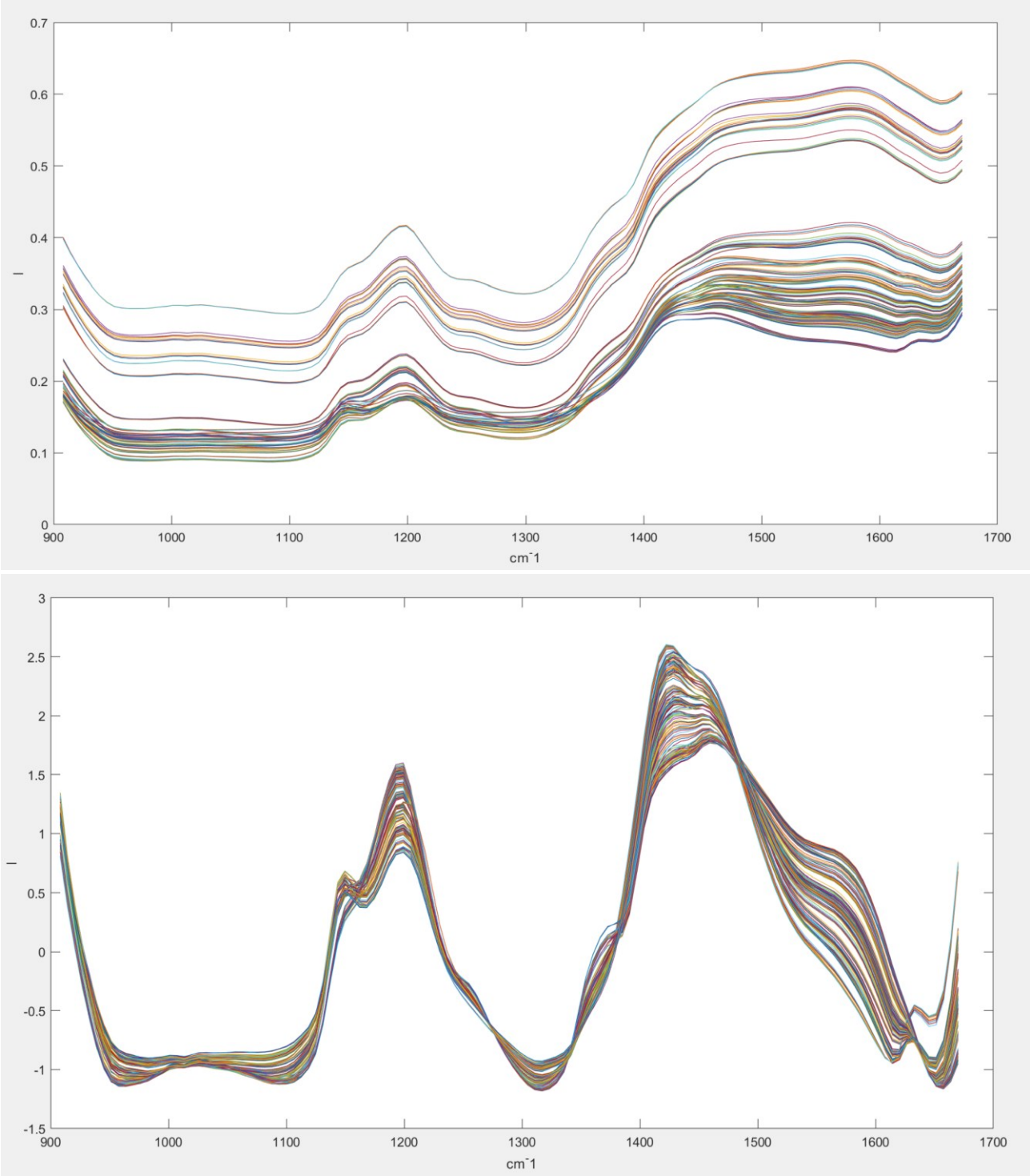



Figure 6. NIR-HSI spectra POCIS Inlet 27_10 (100_Dx) before (first photo) and after preprocessing (second photo).

To extract chemically meaningful information from the spectroscopic datasets and to investigate the contribution of different spectral components within the analysed samples, Multivariate Curve Resolution–Alternating Least Squares (MCR-ALS) was applied as the main chemometric approach. The implementation of MCR-ALS was adapted according to the characteristics of each spectroscopic platform and to the specific analytical objective of the corresponding dataset. For the Raman mapping datasets acquired with the Horiba Raman microscope and for the HSI datasets acquired on POCIS samples, MCR-ALS was primarily employed as a tool for resolving the main spectral contributions and reconstructing their spatial distribution across the analysed surface. Prior to MCR-ALS optimisation, a data geometry approach was applied in order to identify representative spectral profiles to be used as initial estimates for the spectral matrix ($S_0$). A dedicated MATLAB routine implementing this strategy was employed (APPENDIX A Routine MatLab)[62]. The preprocessed spectral matrix was initially imported into MATLAB and arranged according to the original spatial structure of the dataset. A truncated singular value decomposition (SVD)-based PCA was then performed by retaining forty principal components. The obtained score vectors were normalised with respect to the first principal component in order to reduce scale differences among the principal directions and enhance the visualization of the spectral distribution within the score space. The normalised PCA score space was subsequently analysed through a Convex Hull approach. The convex hull vertices were considered representative of the most spectrally distinct observations within the dataset and were therefore extracted as candidate initial estimates for the subsequent MCR-ALS optimisation. This strategy allowed the selection of experimentally meaningful spectral profiles directly from the dataset without requiring manual identification of pure components. For the Horiba Raman datasets, the two independently acquired spectral windows (200–1400 $cm^{-1}$ and 1400–2000 $cm^{-1}$) were processed separately. Therefore, the PCA-based data geometry analysis, spectral selection and subsequent MCR-ALS optimisation were independently performed for each spectral region. This approach allowed the chemical information contained within the two acquisition windows to be evaluated independently while avoiding possible distortions caused by combining regions characterised by different spectral features. The selected spectral profiles were subsequently used as initial estimates for MCR-ALS optimisation performed through the MCR-ALS toolbox command-line implementation. During the iterative optimisation procedure, the Fast Non-Negative Least Squares (FFNLS) algorithm was applied by imposing non-negativity constraints on both concentration profiles (C) and spectral profiles (S), ensuring chemically meaningful solutions. Following optimisation, the resolved concentration profiles were projected back onto the original spatial coordinates of the Raman maps. This procedure enabled the reconstruction of two-dimensional chemical distribution maps associated with each resolved component, allowing direct comparison between the chemometric results and the optical images acquired during the Raman mapping procedure. The same computational workflow was subsequently applied

[62] Löder e Gerdts, «Methodology Used for the Detection and Identification of Microplastics—A Critical Appraisal».

to the hyperspectral imaging datasets acquired from the POCIS samples. Due to the considerably larger dataset size, containing approximately 170,000 spectra per analysed image, the same PCA-assisted spectral selection strategy and MCR-ALS optimisation approach were adopted. The main difference compared with the Raman mapping datasets concerned the constraints applied during the ALS optimisation. For the HSI datasets, non-negativity constraints were not applied during the MCR-ALS optimisation. This choice was related to the characteristics of the preprocessed NIR spectra, which contained negative intensity values generated by the applied spectral correction procedure. Under these conditions, the application of non-negativity constraints on the spectral profiles resulted in optimisation errors and prevented the correct convergence of the ALS algorithm. Therefore, the HSI datasets were analysed without non-negativity constraints in order to ensure the stability and convergence of the MCR-ALS optimisation procedure, while maintaining the possibility of resolving the main spectral contributions associated with the hyperspectral images. For the datasets acquired using the portable Raman spectrometer and the MicroNIR instrument, MCR-ALS was applied with the objective of evaluating the capability of portable spectroscopic platforms to resolve the spectral contributions associated with the simulated SPE samples prepared according to the Design of Experiments (DoE). Unlike the Raman mapping datasets, the main objective was not the spatial reconstruction of chemical components but the identification and discrimination of the investigated emerging contaminants within the complex sorbent matrix. Three different modelling strategies were investigated: conventional MCR-ALS, multiset MCR-ALS and equality-constrained MCR-ALS. In the conventional approach, the preprocessed experimental spectra were analysed as a single data matrix without additional external information. In the multiset approach, the experimental datasets and the corresponding reference spectra of the investigated standards were combined within a common MCR-ALS framework. Additional information regarding the expected presence of specific components within the different datasets was introduced through appropriate constraint matrices, allowing the optimisation procedure to exploit the structure of the combined dataset. The equality-constrained MCR-ALS approach was implemented by introducing prior spectral information derived from the pure standard spectra. Characteristic spectral regions associated with each investigated compound were identified by analysing the reference spectra. A constraint matrix was subsequently constructed in which the selected spectral regions were defined as unconstrained variables (NaN), whereas non-informative spectral regions were fixed to zero. This approach forced the optimisation procedure to search for the contribution of each compound within chemically relevant spectral regions while maintaining flexibility in the remaining parts of the spectrum. For all MCR-ALS models applied to portable Raman and MicroNIR datasets, the FFNLS algorithm was used with non-negativity constraints applied to both concentration and spectral profiles. The resolved spectral profiles (Sopt) and concentration profiles (Copt) obtained from the different modelling strategies were subsequently compared to evaluate the influence of the applied constraints on component resolution and contaminant discrimination. The complete set of optimisation parameters

resolved components and their analytical interpretation are reported and discussed in the Results and Discussion section.

### *3.1.2. Bilbao treatment data*

For the Raman data, the preprocessing workflow included a denoising step using a Savitzky–Golay filter (polynomial order = 2, window width = 9), followed by baseline correction to minimize the contribution of background fluorescence typical of environmental samples. Baseline correction was performed using the asymmetric least squares method with the following parameters: $\lambda = 1 \times 10^5$, $p = 0.01$, and $d = 2$. Finally, Probabilistic Quotient Normalization (PQN) was applied to correct for systematic intensity variations between spectra. PQN is a normalization method widely used in chemometric analysis to compensate for dilution or overall intensity differences by scaling each spectrum relative to a reference spectrum, typically the median spectrum of the dataset. The normalization factor is calculated as the median of the quotients between each spectrum and the reference, allowing robust correction of multiplicative effects while preserving the relative intensity relationships among spectral features. This approach is particularly suitable for complex environmental samples, where variations in signal intensity may arise from differences in particle size, laser–sample interaction, or local surface heterogeneity (Figure 7).

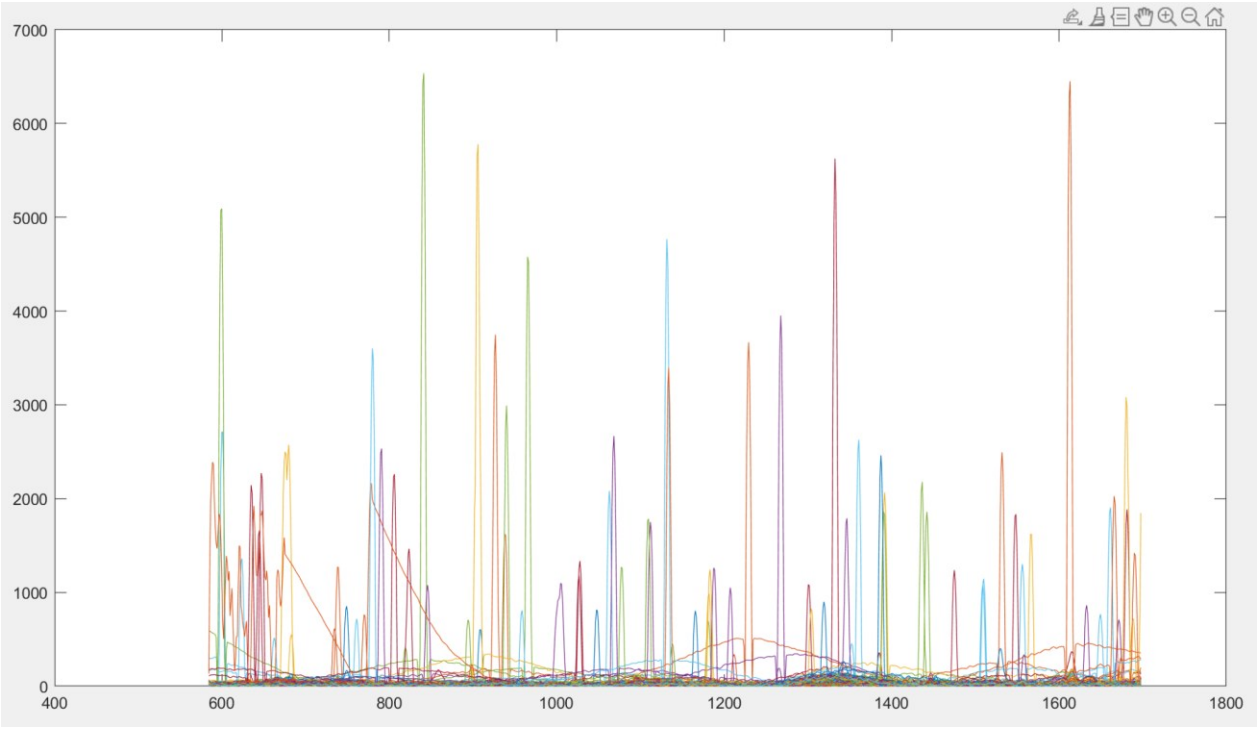
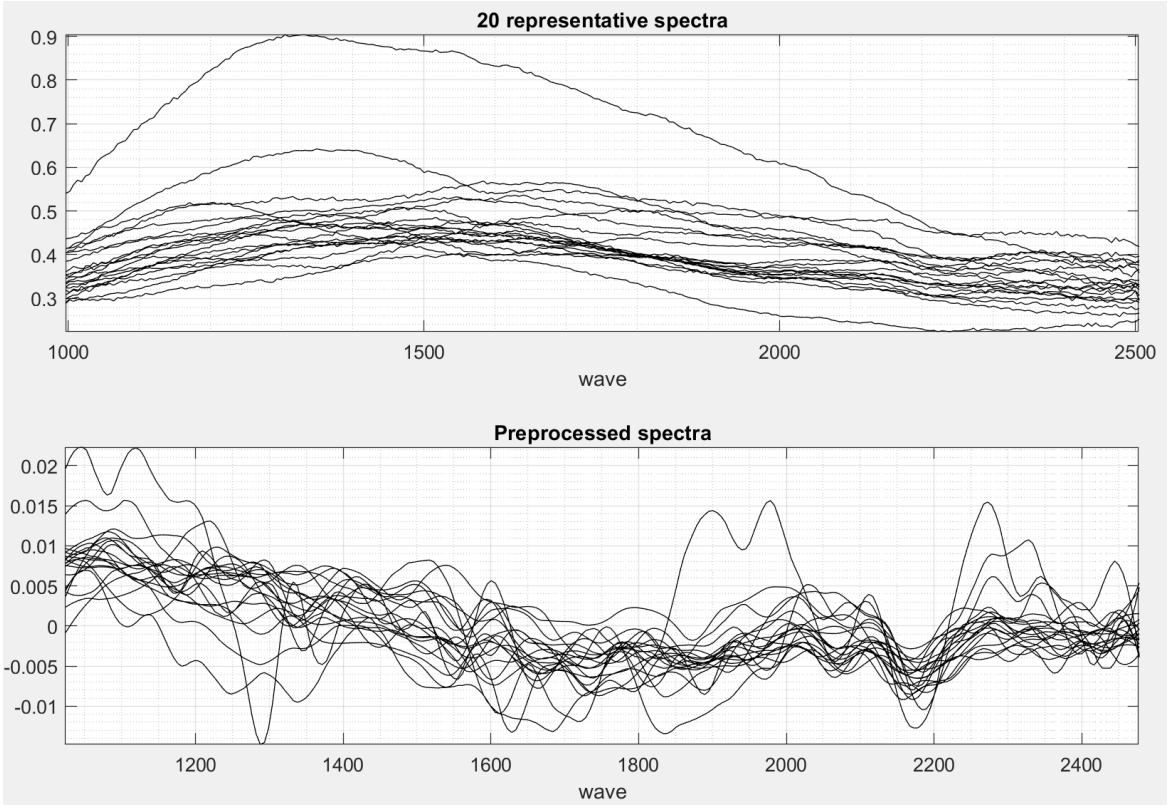


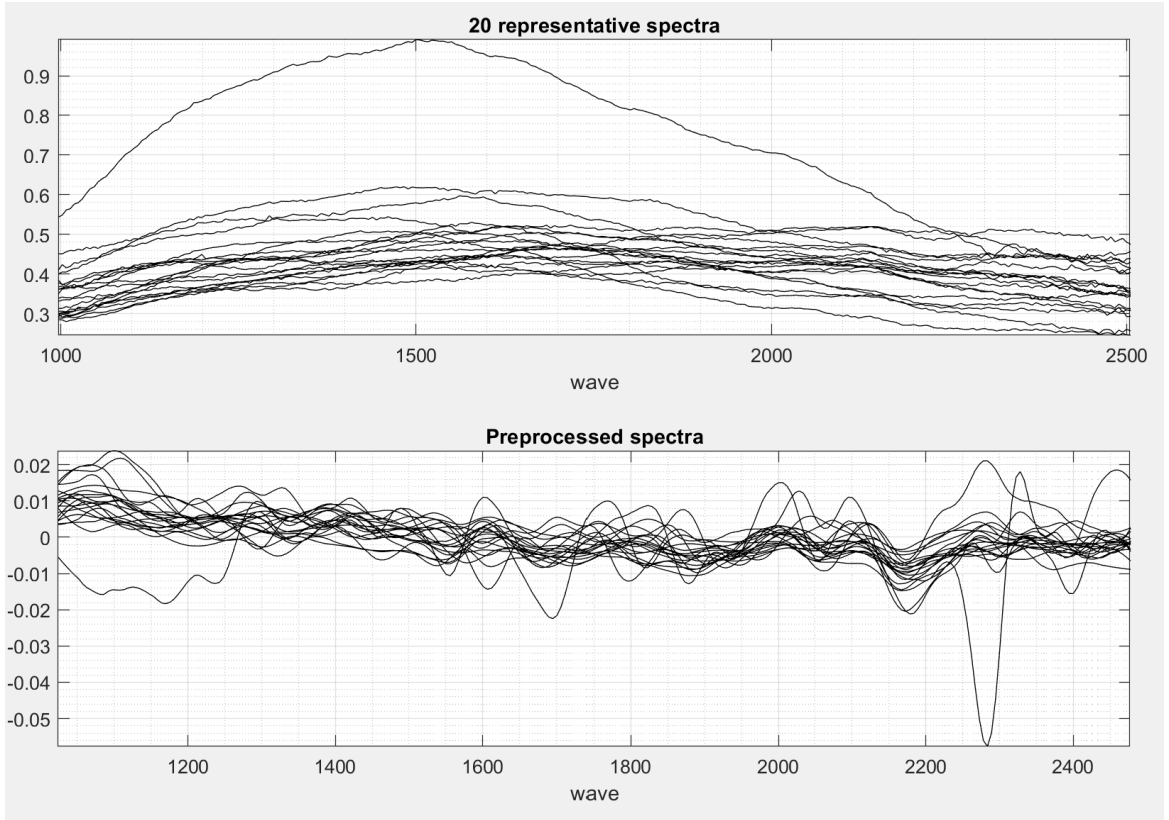

Figure 7. Preprocessed Raman spectra of Tert_tre mapping.

Similarly, the NIR hyperspectral data were processed using the HyperTools3 environment through a dedicated preprocessing workflow designed to enhance spectral interpretability and reduce variability unrelated to chemical composition. As a first step, each hyperspectral image was spatially cropped using the *spatial crop* function in order to isolate one filter at a time and remove irrelevant areas of the image. Subsequently, a mask was generated using the PCA

scores surface selective approach, which allowed the discrimination of the filter region from the background and the selection of the pixels corresponding to the sample for further spectral analysis.

The extracted spectra were then subjected to Standard Normal Variate (SNV) normalization. This preprocessing step is particularly suitable for solid and heterogeneous matrices, as it corrects multiplicative scattering effects and path length variations arising from differences in particle size, surface roughness, and local packing density. By centering and scaling each spectrum individually, SNV reduces intensity variations caused by physical scattering phenomena and improves the comparability of spectra across the dataset. Subsequently, a first derivative transformation was calculated using a Savitzky–Golay filter (window size = 11, polynomial order = 2, derivative order = 1). This operation helps remove baseline offsets and slow baseline drifts, while enhancing the resolution of overlapping absorption bands and emphasizing subtle spectral features associated with specific chemical components. Finally, an additional Savitzky–Golay smoothing step (window size = 11, polynomial order = 2) was applied to reduce high-frequency noise and improve the signal-to-noise ratio while preserving the shape and position of the spectral bands, thereby facilitating subsequent multivariate analysis (Figure 8).


20 representative spectra
0.9
0.8
0.7
0.6
0.5
0.4
0.3
1000
1500
2000
2500
wave
Preprocessed spectra
0.02
0.015
0.01
0.005
0
-0.005
-0.01
1200
1400
1600
1800
2000
2200
2400
wave


20 representative spectra
0.9
0.8
0.7
0.6
0.5
0.4
0.3
1000
1500
2000
2500
wave
Preprocessed spectra
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
1200
1400
1600
1800
2000
2200
2400
wave

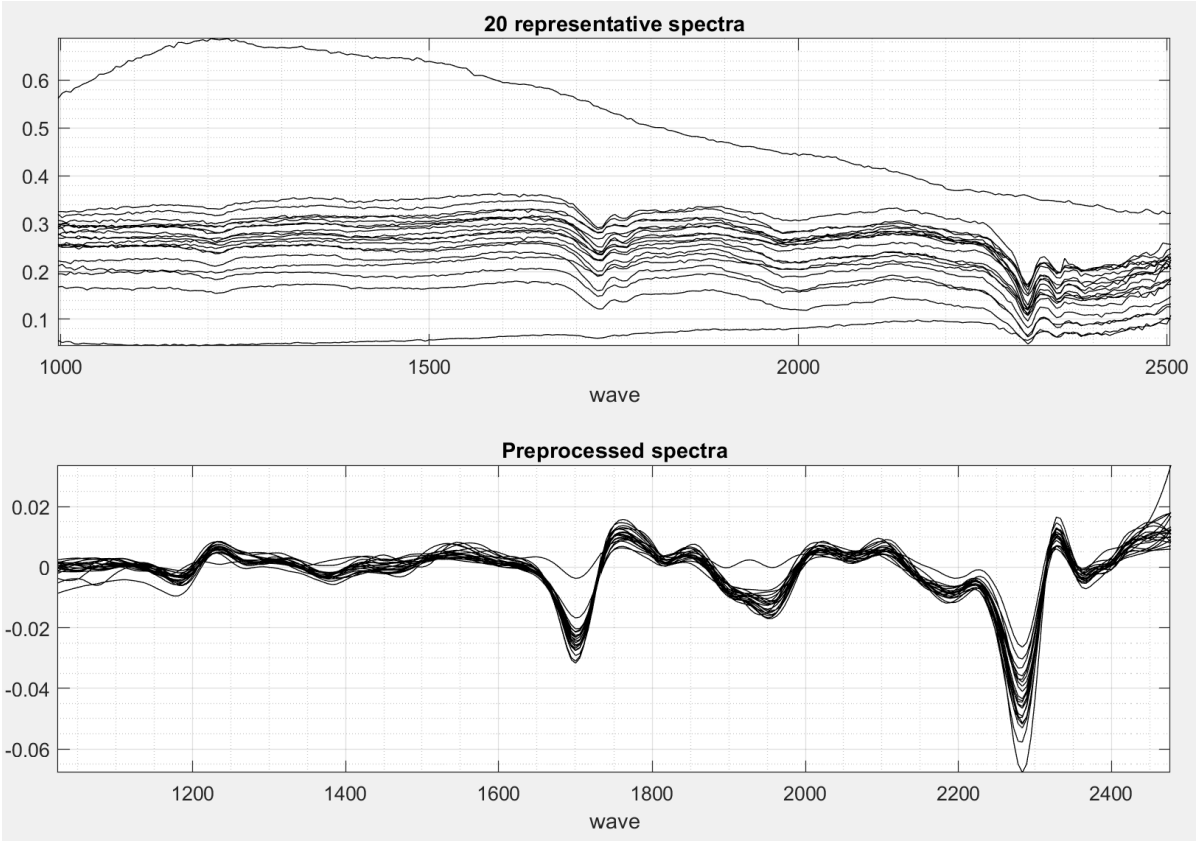



Figure 8. Preprocessing NIR-HSI spectra of three fractions: Biol fraction (first photo), Deka fraction (second photo) and Eraz fraction (last photo).

Prior to the application of MCR-ALS, an exploratory analysis of the dataset was carried out using Principal Component Analysis (PCA). PCA is an unsupervised multivariate technique widely used in chemometrics to reduce data dimensionality while preserving the main sources of variance within the dataset. By projecting the original variables onto a reduced set of orthogonal latent variables (principal components), PCA enables the identification of underlying patterns, similarities, and potential clustering among samples, as well as the detection of outliers and dominant spectral trends. The PCA model was calculated using the HyperTools3 toolbox, which provides an interactive chemometric environment for the analysis of spectroscopic and hyperspectral datasets. Although its graphical interface differs from that of the PLS_Toolbox, HyperTools3 implements the same fundamental PCA algorithms and functionalities commonly used in chemometric analysis, including visualization of scores, loadings, and explained variance. An initial PCA model was first computed by retaining twelve principal components. Following inspection of the scree plot and the cumulative explained variance, and applying both the elbow criterion and the clutter rule, the number of retained components was reduced to ten principal components. These criteria help identify the point at which additional components contribute only marginally to the explained variance and mainly describe residual noise or minor variability rather than meaningful systematic information.
Subsequently, MCR-ALS analysis was performed within the same software environment. The number of components was set to ten, as suggested by the previous PCA analysis. The

algorithm was initialized using random initial estimates, while non-negativity constraints were imposed on both the concentration and spectral profiles, in accordance with the physical meaning of spectroscopic data. This chemometric analysis was applied exclusively to the Raman maps, where the large number of spectra and the spatially resolved nature of the dataset required multivariate analysis to extract chemically meaningful information. For the other micrometric fractions, the point-by-point Raman analysis was considered sufficient to identify the microplastics (MPs) present in the samples. Finally, the results obtained from the Raman chemometric analysis will be compared with those derived from the NIR spectra of the same fractions, in order to evaluate the potential advantages of the two spectroscopic techniques, particularly in terms of analytical reproducibility.
The same chemometric workflow, consisting of PCA followed by MCR-ALS, was also applied to the NIR spectral datasets in order to allow a consistent comparison between the two spectroscopic techniques.

### *3.2. Spectroscopic and imaging techniques.*

Raman spectroscopy analyses performed during the permanence in Bilbao were carried out using two different instrumental platforms, selected according to the specific objectives and sample types investigated throughout the experimental work. An InVia Raman microscope (Figure 9) was employed for the characterisation of microplastics retained on metallic filters collected from wastewater treatment plants in the Bilbao area. This instrument was particularly suited for high-resolution, particle-level analysis, enabling the identification of polymeric materials directly on the filtration supports. Conversely, a Horiba Raman spectrometer (Figure 9) was used exclusively for the analysis of emerging contaminants (ECs) in samples processed at the University of Modena and Reggio Emilia. In addition to the benchtop Raman system, portable Raman and microNIR instruments were also investigated as complementary experimental techniques. The use of portable instrumentation was aimed at evaluating whether less performing and more compact analytical devices, compared to conventional laboratory-based systems, could still provide meaningful spectral information related to the presence of ECs within the analysed samples. This approach was particularly relevant in view of the potential development of faster and more accessible screening methodologies for environmental monitoring applications. The choice of the Horiba Raman system ensured consistency with the analytical protocols routinely adopted in the Modena laboratory, while the integration of portable techniques allowed a preliminary assessment of their applicability to the investigated matrices. Although different spectroscopic systems and analytical targets were involved, comparable data quality criteria and preprocessing strategies were adopted to ensure methodological coherence across the generated datasets.


Figure 9. Raman InVia (left) and Raman Horiba (right).

In contrast to Raman spectroscopy, which was applied selectively depending on the analytical target and sample origin, hyperspectral imaging in the NIR range (Figure 10) was systematically employed for all investigated samples including those collected and analysed in both Modena and Bilbao. This approach enabled a homogeneous spectroscopic screening of the entire sample set, irrespective of the specific matrix or analytical focus. The application of HSI to all samples allowed a consistent comparison across different wastewater treatment plants and sample types, providing complementary information to Raman spectroscopy in terms of spatial coverage, throughput and suitability for large-area analysis. In this context, HSI served as a unifying analytical platform, supporting the comparative evaluation of spectroscopic methodologies for the study of both microplastics and emerging contaminants.

Figure 10. Hyperspectral camera SPECIM.

### *3.1.1 Spectroscopic parameters for the analysis of samples from Modena*

For the analysis of selected portions of the POCIS samples, Raman spectral maps were acquired using a Horiba Raman microscope equipped with a 100× objective, in order to achieve high spatial resolution and improve the detection of local spectral variations associated with the adsorbed compounds. The analyses were performed directly on restricted areas of the sorbent material, selected after preliminary optical observation of the sample surface. The acquired Raman maps (Figure 11) consisted of 265 individual spectra (acquisition points), collected over a spectral range approximately extending from 400 to 1800 $cm^{-1}$, corresponding to the region containing the main vibrational bands of both the sorbent material and the investigated contaminants. Spectral acquisitions were performed using a laser power set at 10%, with an acquisition time of 5 s and 2 accumulations for each analysed point, in order to obtain an adequate signal-to-noise ratio while minimising possible thermal effects on the sample surface. A relevant characteristic of the Horiba system concerns the acquisition mode of the Raman spectra. The instrument automatically divides the full spectral range into two separate acquisition windows, generating two independent spectral regions for each analysed point. In the acquired datasets, these regions corresponded to variables 1–659 and 660–1147, respectively. This instrumental configuration was maintained for all mapped samples in order to ensure spectral consistency and adequate signal quality across the entire investigated Raman range.

Figure 11. Acquired area.

Portable Raman analyses were performed using a handheld Raman spectrometer under fixed acquisition conditions for all analysed samples. Spectra were collected with an exposure time of 2000 ms, using 15 captures for each spectrum and a laser power set at the maximum instrumental value (99.89%, in order to compensate for the lower sensitivity of the portable system compared to benchtop instrumentation, Figure 12). During acquisition, the distance between the sample surface and the laser source was maintained at approximately 1 cm in order to ensure consistent focusing conditions and improve spectral reproducibility. Before the acquisition of the samples, a background spectrum was recorded by positioning the Raman probe at its maximum height and acquiring the spectrum of ambient air. This procedure was repeated at the beginning of each acquisition session in order to account for possible instrumental variations occurring during the experimental work. Subsequently, spectral acquisitions were carried out on the plastic sachets containing the sorbent material and the analytical standards. For each sachet, spectra were acquired from three fixed measurement points, collecting three replicates for each point in order to improve the representativeness and reproducibility of the measurements. At the end of each analytical session, blank measurements consisting of the sorbent material without added standards were acquired following the same acquisition procedure adopted for the samples.

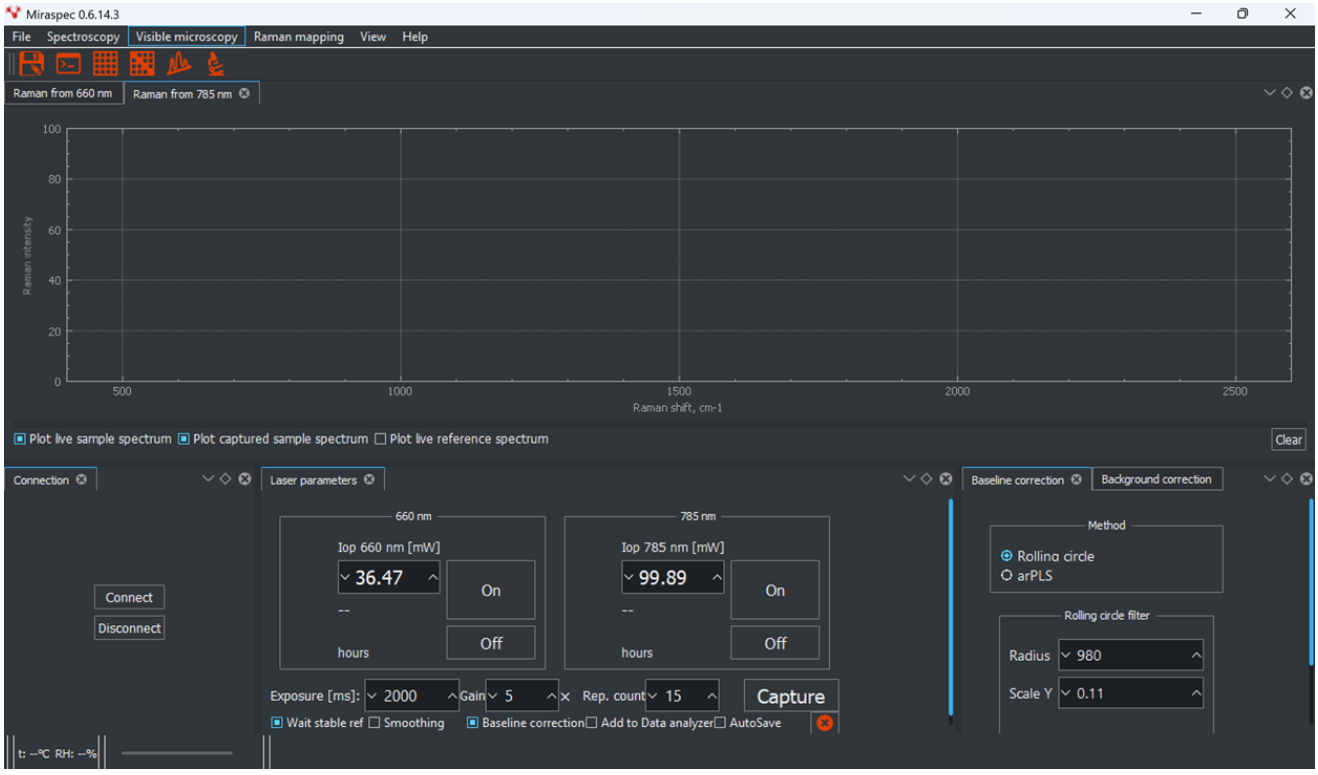



Figure 12. Portable Raman software interface.

MicroNIR analyses (Figure 13) were performed in diffuse reflectance mode, selected to maximise the interaction between the incident radiation and the analysed samples while maintaining a rapid and non-destructive acquisition procedure. Spectra were acquired using an integration time of 9.7 ms and a scan count of 100 for each measurement, allowing signal

averaging and improving the overall signal-to-noise ratio of the collected spectra. Before sample acquisition, the instrumental calibration procedure included the acquisition of a dark spectrum by covering the lamp with the dedicated optical shutter provided by the instrument. Subsequently, a reference spectrum was acquired using a circular white reference platform supplied for reflectance measurements. After the acquisition of the dark and reference spectra, five blank spectra were collected in order to verify the stability and reproducibility of the instrumental signal prior to sample analysis. Following this preliminary calibration phase, spectral acquisitions were performed directly on the analysed samples, including both the plastic sachets containing the sorbent material and the Petri dish samples. The acquisition parameters were maintained constant for all measurements in order to ensure spectral consistency and improve the comparability among datasets obtained from different sample surfaces and concentration levels.

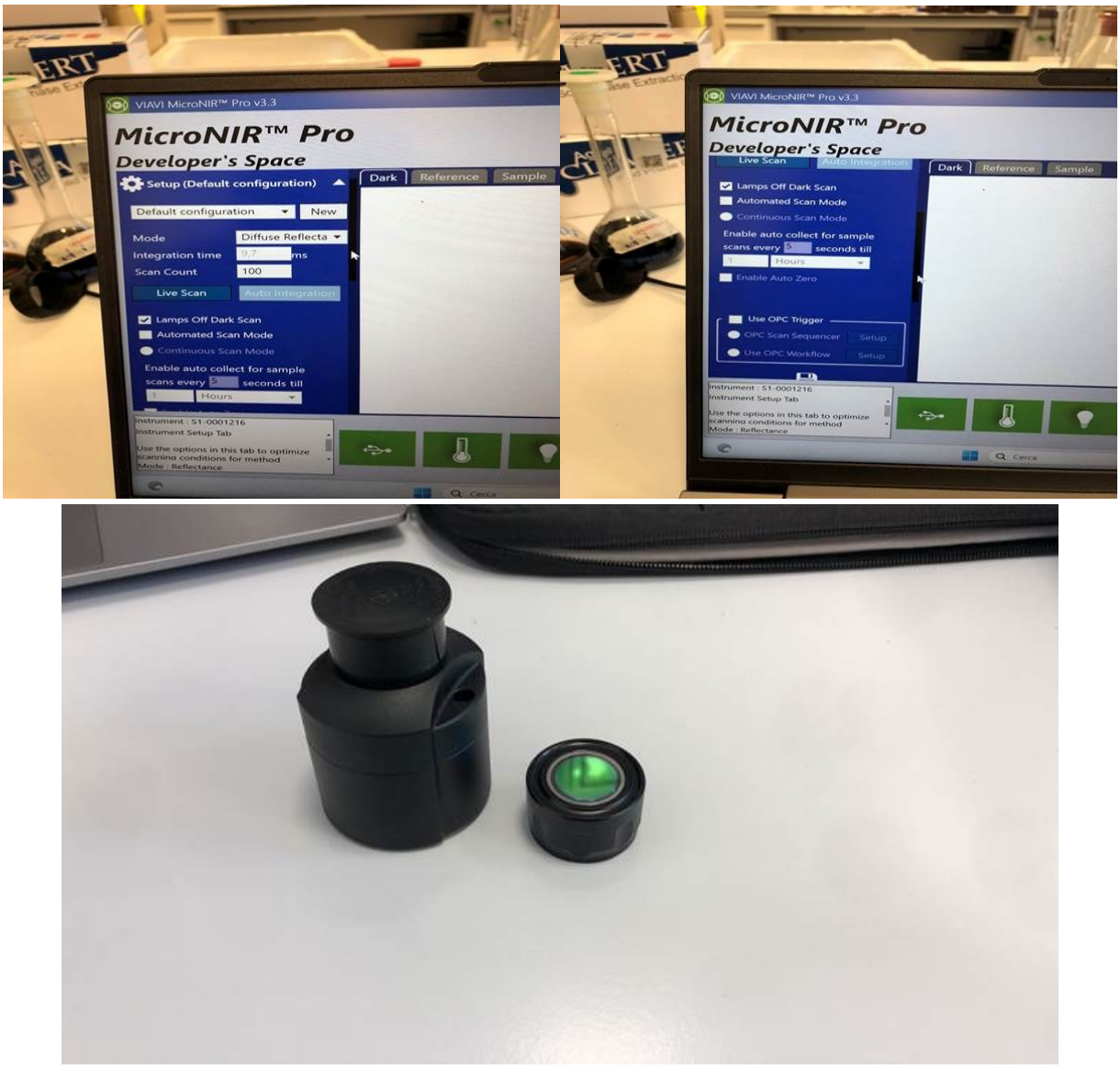



Figure 13. NIR acquisition parameters (above) and NIR lamp.

*3.2.2. Spectroscopic parameters for the analysis of samples from Bilbao*

Prior to spectroscopic analysis, the metallic filters, obtained after a mild pre-treatment of the wastewater samples, were examined under an optical microscope in order to preliminarily identify potential microplastic particles. This screening step (Figure 14, left) was based on morphological criteria, including colour, shape and particle size, and was aimed at facilitating and accelerating the subsequent Raman point-by-point analysis by restricting measurements to regions of interest.

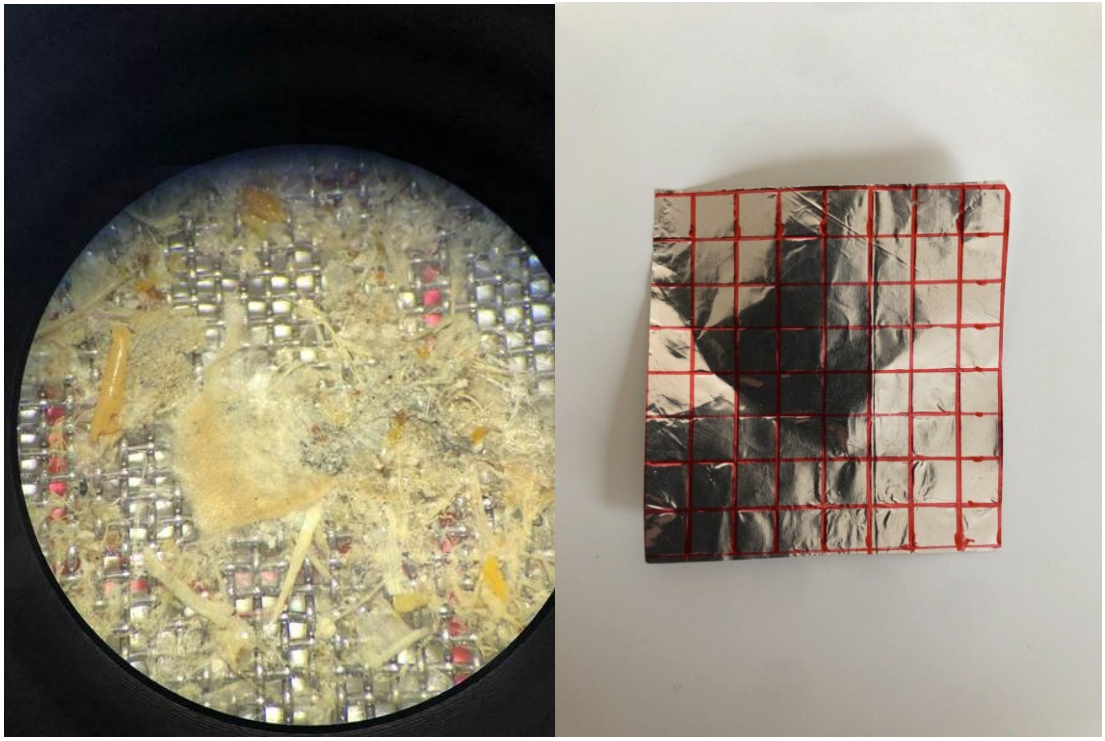


Figure 14. Microscope observation of one interesting spot on the metallic filter (left) and reference grid (right).

To systematically inspect the entire filter surface, each metallic filter was placed on a reference grid of approximately 4 cm x 4 cm, subdivided into square areas of 5 mm x 5 mm (Figure 14, right). In analogy with a grid-based mapping approach, each individual spot was visually inspected, and only those regions showing features compatible with MPs were selected for Raman analysis. This procedure allowed a structured and reproducible exploration of the filters while minimising unnecessary spectral acquisitions.

Raman measurements were performed using a red laser (1200 l/mm) excitation wavelength of 785 nm. For most particles, spectra were acquired using a 20x objective, with a laser power set between 5% and 10%, depending on the acquisition time and the number of accumulations. The spectral range typically extended from 400 to 1800 $cm^{-1}$, covering the characteristic vibrational bands of common polymeric materials. In standard conditions, spectra were collected using 5 accumulations with acquisition times of 5/10 s per accumulation and a laser power of 10%. For smaller particles requiring higher spatial resolution, a 50x objective was

employed. In these cases, parameters were selected as a compromise between signal-to-noise ratio and preservation of particle integrity.

Polymer identification was performed by comparing the acquired Raman spectra with reference spectra of commonly detected microplastics, including PP, PE, PET, PS, PVC and PA. Spectral matching focused on the presence and relative intensities of diagnostic Raman bands characteristic of each polymer type.

The metallic filters analyzed by Raman spectroscopy were found to be representative of four distinct micrometric fractions, each corresponding to five sampling points along the treatment line of the wastewater treatment plant. For each sampling point, the first two micrometric fractions were analyzed using a point-by-point Raman approach, following a preliminary observation phase under an optical microscope. As described previously, this step allowed the selective identification of particles potentially attributable to MPs based on morphological criteria, thereby optimizing the number of Raman acquisitions and restricting the analysis to regions of actual interest. For the third micrometric fraction, a Raman mapping approach was employed to achieve a more extensive and systematic spatial characterization of the filter surface. Maps were acquired using an integration time of 2 s per acquisition, with 2 accumulations per point, a laser power set to 50%, and a low magnification objective (5x), chosen to cover large sample areas. The final map image was subsequently reconstructed by mosaicking (Figure 15) the individual acquisitions, allowing a continuous and coherent visualization of the analyzed regions (Figure 16). The combined use of point-by-point analysis for the finer fractions and Raman mapping for the third fraction enabled the analytical strategy to be tailored to particle size and distribution, balancing spatial resolution, areal coverage, and acquisition time. This approach provided complementary information on the presence and type of MPs in the metallic filters, enhancing the overall robustness of the spectroscopic characterization.

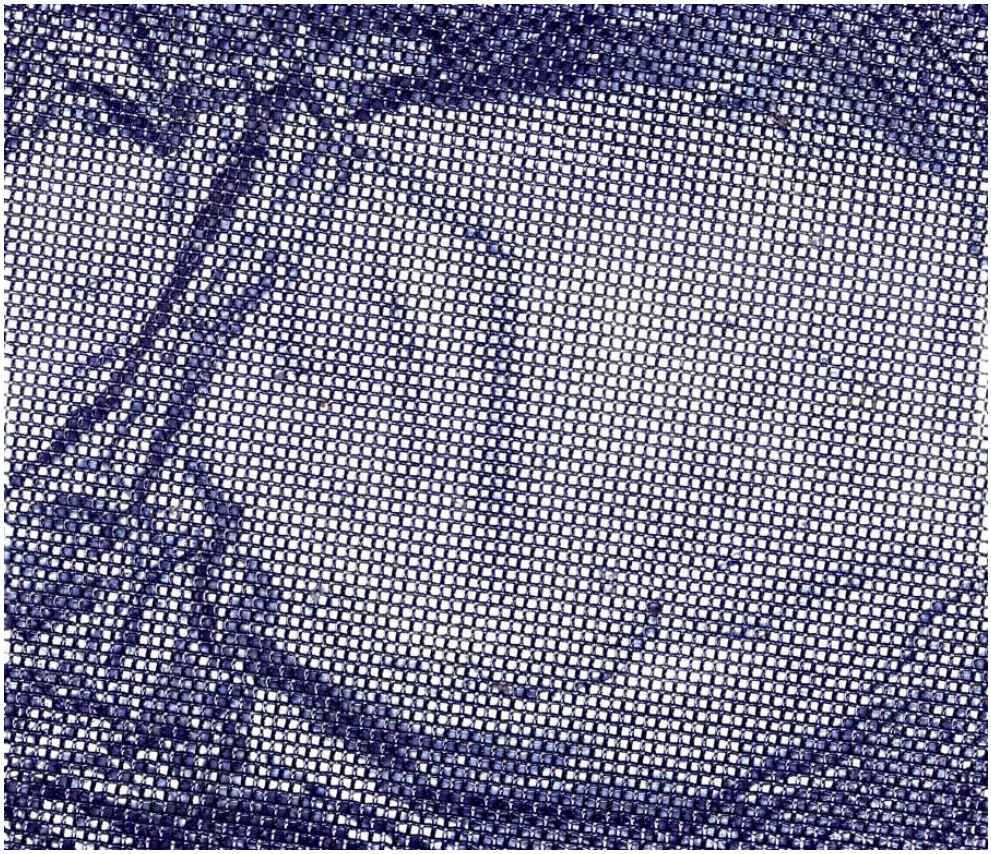


Figure 15. Mosaic of 0.2-0.5 fraction of Biol_Irt sampling point

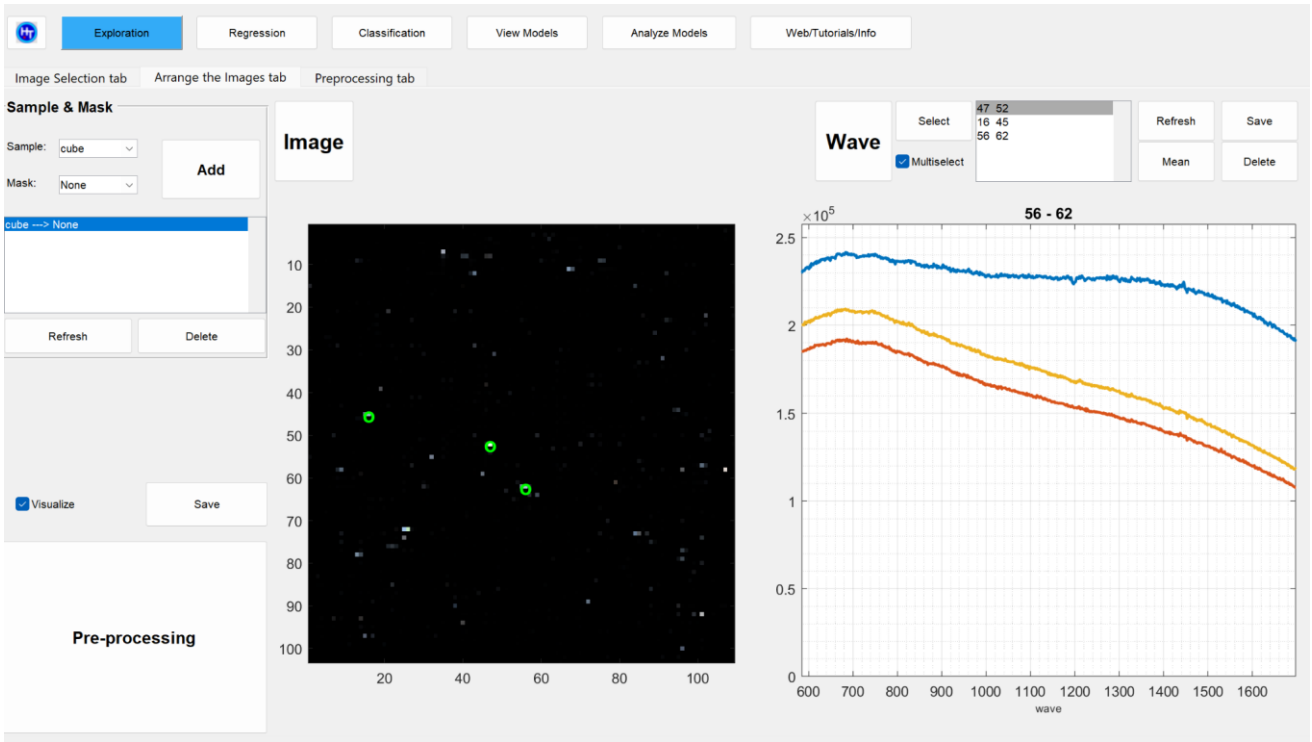



Figure 16. Three random Raman spectra of 0.5-0.2mm fraction without preprocessing on HyperTools environment.

In addition to the Raman analyses, hyperspectral images (HSI) in the short-wave infrared (SWIR) region were also acquired for the smallest micrometric fraction of the same samples (Figure 17), with the aim of obtaining complementary spatial and spectral information on the distribution of particles on the surface of the metallic filters. While Raman spectroscopy allows the molecular identification of individual particles, hyperspectral imaging enables the simultaneous acquisition of spectral information for each pixel of the image, providing a spatially resolved representation of the sample composition over relatively large areas.

HSI acquisitions were performed using a hyperspectral camera operating in the SWIR region (Lumo Scanner v2023-631), configured in a *pushbroom* acquisition mode. In this configuration, spectral information for all wavelengths is simultaneously recorded along a single line of the sample, while the relative motion between the sensor and the sample allows the progressive reconstruction of a three-dimensional hyperspectral datacube composed of two spatial dimensions and one spectral dimension. The resulting hyperspectral dataset consisted of 384 spatial pixels in the transverse direction (samples), 4864 scan lines (lines), and 272 spectral bands (bands), generating a datacube with overall dimensions of 384 × 4864 × 272. The spectral range covered by the sensor extended approximately from 996 to 2505 nm, corresponding to the SWIR region, where absorption bands associated with overtones and combination bands of vibrational modes of several functional groups present in polymeric materials can typically be observed. The spectral resolution of the system, expressed as full width at half maximum (FWHM), was approximately 5.6 nm across the entire spectral range.

Data acquisition was carried out using an integration time of 3.9 ms and an acquisition rate of approximately 140 frames per second. The detector was operated without spatial binning (1 × 1), thus preserving the full spatial resolution of the sensor. The system was configured in internal trigger mode with active synchronization in order to ensure temporal stability during the scanning of the filter surface. The hyperspectral data were stored in ENVI format using a band-interleaved-by-line (BIL) structure and a 12-bit data depth. During acquisition, the SWIR sensor was maintained at an operating temperature of approximately 152 K to minimize thermal noise and ensure stable measurement conditions. These acquisition parameters enabled the collection of hyperspectral images with high spectral resolution in the SWIR region, supporting the detection and spectral characterization of particles potentially attributable to microplastics on the surface of the analyzed filters.

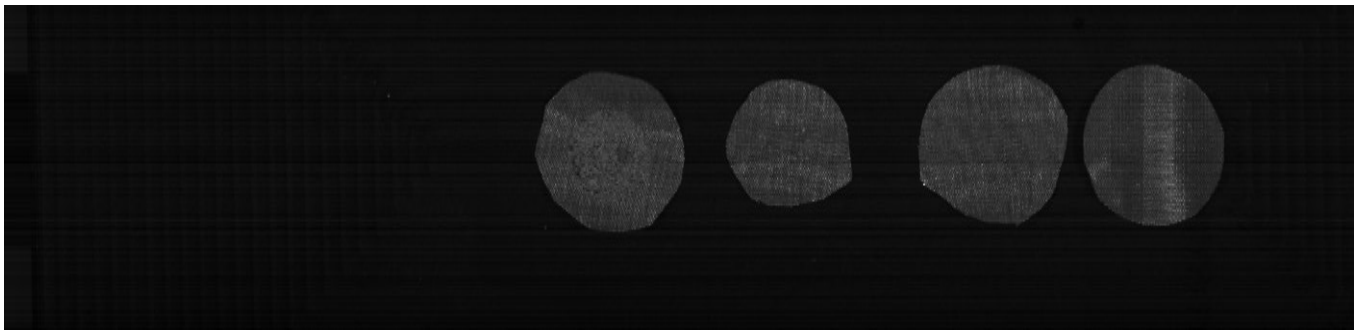


Figure 17. Metallic filters acquired by HSI camera.

### *3.3 Samples collections and treatments*

The experimental activity developed during this thesis involved the analysis of two distinct groups of environmental samples, collected and processed according to different objectives and analytical strategies. The first group consisted of wastewater samples collected from wastewater treatment plants located in the Basque Country (Spain), which were investigated for the identification and characterisation of microplastics along the treatment line. The second group comprised passive sampling materials and laboratory-prepared sorbent samples from the Modena wastewater treatment plant, used for the investigation of emerging contaminants through different spectroscopic approaches. Given the different nature of the target analytes and matrices involved, specific sampling, preparation and analytical procedures were adopted for each case study. The following sections describe in detail the methodologies applied to the Bilbao and Modena samples prior to the spectroscopic analyses and chemometric data treatment.

#### *3.3.1 Sampling and treatment of environmental matrices from Modena*

The environmental samples investigated in this study consisted of Polar Organic Chemical Integrative Samplers (POCIS, Figure 18) deployed at the wastewater treatment plant of Modena. The sampling strategy included samplers exposed at two representative locations along the treatment line, corresponding to the influent (Inlet) and effluent (Outlet) streams of the plant. In addition, an unexposed POCIS was analysed as a blank reference in order to assess the spectral characteristics of the sorbent material in the absence of contaminant accumulation. Due to the exploratory nature of the study and the limited time available for instrumental analyses, spectroscopic investigations were performed only on the blank POCIS and on one sampler collected at the influent section of the treatment plant (Inlet). The selected samplers were analysed using benchtop Raman spectroscopy and near-infrared hyperspectral imaging (HSI). After retrieval, the POCIS devices were opened by separating the two membrane discs, allowing direct access to the internal sorbent phase. No extraction, purification or additional pretreatment procedures were applied prior to analysis. Instead, spectroscopic measurements were performed directly on the exposed sorbent material in its original state. This approach was adopted to evaluate the capability of Raman and hyperspectral techniques to detect chemical differences associated with contaminant accumulation while preserving the integrity of the samples and minimising potential analyte losses. The datasets obtained from the POCIS samples were subsequently processed through the chemometric workflow described in the previous section. In contrast, portable Raman and MicroNIR spectroscopy were not applied to the environmental POCIS samples but were exclusively employed for the analysis of laboratory-prepared samples containing the sorbent material spiked with selected standards at known concentration levels. This complementary experimental design allowed the performance of portable spectroscopic platforms to be evaluated under controlled conditions before their potential application to more complex environmental matrices.

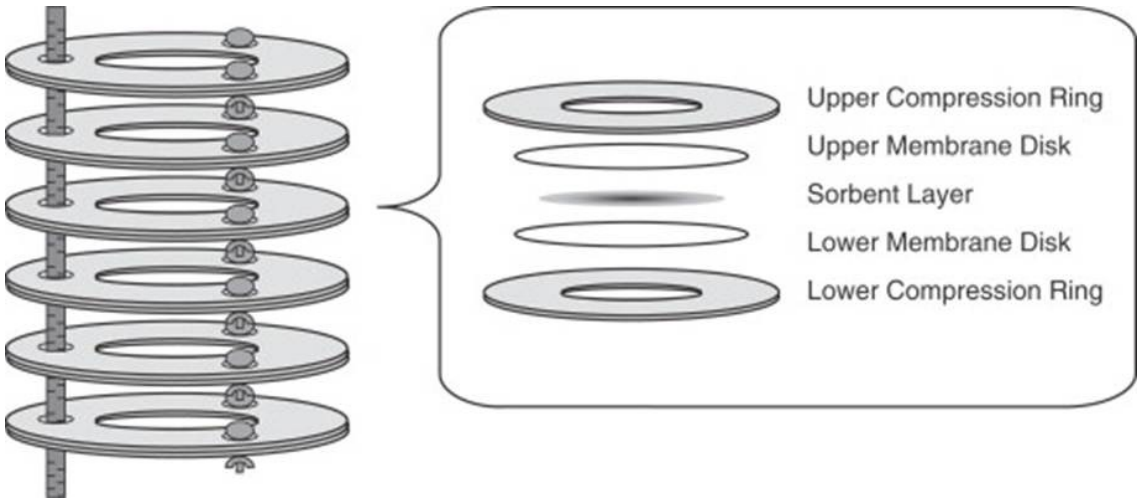



Figure 18. POCIS device for sampling ECs

Following the analysis of the environmental POCIS samples, an additional set of laboratory-prepared sorbent materials was generated in order to evaluate the performance of portable Raman and MicroNIR spectroscopy under controlled experimental conditions. The samples were prepared using the same sorbent phase employed in the POCIS devices, thus providing a matrix representative of real solid-phase extraction (SPE) materials while allowing the controlled addition of target analytes. Five pharmaceutical compounds commonly classified as emerging contaminants were selected as model analytes: caffeine, hydrochlorothiazide, diclofenac, metoprolol and atenolol. The standards were added directly to the sorbent material according to a dedicated Design of Experiments (DoE), specifically developed to investigate the spectroscopic response over a wide range of concentration combinations (Table 1). Each experimental sample was prepared by introducing 300 mg of sorbent material into a 10 mL glass vial. The selected standards were then added according to the concentration levels defined by the DoE together with 1 mL of methanol, used as solvent to promote the homogeneous distribution of the analytes throughout the sorbent phase. The experimental design consisted of thirty different formulations, in which the concentration of each contaminant was systematically varied across five concentration levels, generating a dataset characterised by controlled chemical variability and allowing potential interactions among the target compounds to be investigated. After preparation, the vials were left to dry for approximately 24 h to allow complete evaporation of the methanol. Once dried, the sorbent material was transferred into dedicated plastic sachets, which provided a reproducible and practical sample configuration for subsequent portable Raman analyses. The same formulations were also analysed by MicroNIR spectroscopy, both within the sachets and after transfer into Petri dishes, in order to evaluate the influence of the sample support on the acquired spectra. The resulting dataset provided a controlled experimental framework for assessing the capability of portable spectroscopic techniques to detect spectral variations associated with emerging contaminants adsorbed onto sorbent materials. This approach represents a preliminary step towards the application of portable spectroscopic platforms for the rapid screening of environmental samples and passive sampling materials.

Table 1. DoE for SPE simulation preparation

| Exp No | Exp Name | Run Order | Incl/Excl | Caffeine CAFF | Metoprolol MTPL | Atenolol ATN | Diclofenac DCF | Hydroclorothiazide HCTZ |
|---|---|---|---|---|---|---|---|---|
| 1 | N1 | 10 | Incl | 14 | 8 | 8 | 2 | 2 |
| 2 | N2 | 26 | Incl | 5 | 5 | 5 | 5 | 2 |
| 3 | N3 | 23 | Incl | 11 | 5 | 11 | 8 | 2 |
| 4 | N4 | 30 | Incl | 5 | 14 | 2 | 11 | 2 |
| 5 | N5 | 9 | Incl | 2 | 2 | 14 | 14 | 2 |
| 6 | N6 | 3 | Incl | 8 | 11 | 11 | 2 | 2 |
| 7 | N7 | 4 | Incl | 11 | 11 | 14 | 5 | 5 |
| 8 | N8 | 2 | Incl | 8 | 2 | 5 | 8 | 5 |
| 9 | N9 | 7 | Incl | 2 | 11 | 2 | 11 | 5 |
| 10 | N10 | 13 | Incl | 14 | 14 | 11 | 14 | 5 |
| 11 | N11 | 6 | Incl | 2 | 5 | 14 | 2 | 5 |
| 12 | N12 | 25 | Incl | 5 | 8 | 8 | 5 | 5 |
| 13 | N13 | 16 | Incl | 14 | 2 | 2 | 8 | 8 |
| 14 | N14 | 20 | Incl | 11 | 8 | 5 | 11 | 8 |
| 15 | N15 | 17 | Incl | 5 | 11 | 8 | 14 | 8 |
| 16 | N16 | 19 | Incl | 14 | 5 | 14 | 2 | 8 |
| 17 | N17 | 1 | Incl | 2 | 14 | 11 | 5 | 8 |
| 18 | N18 | 29 | Incl | 8 | 14 | 8 | 8 | 8 |
| 19 | N19 | 5 | Incl | 14 | 8 | 11 | 11 | 11 |
| 20 | N20 | 14 | Incl | 11 | 5 | 2 | 14 | 11 |
| 21 | N21 | 28 | Incl | 5 | 2 | 2 | 2 | 11 |
| 22 | N22 | 24 | Incl | 11 | 2 | 8 | 5 | 11 |
| 23 | N23 | 18 | Incl | 2 | 11 | 5 | 8 | 11 |
| 24 | N24 | 22 | Incl | 8 | 14 | 14 | 11 | 11 |
| 25 | N25 | 15 | Incl | 8 | 2 | 11 | 14 | 14 |
| 26 | N26 | 12 | Incl | 11 | 14 | 5 | 2 | 14 |
| 27 | N27 | 11 | Incl | 8 | 8 | 2 | 5 | 14 |
| 28 | N28 | 21 | Incl | 5 | 8 | 14 | 8 | 14 |
| 29 | N29 | 27 | Incl | 2 | 5 | 8 | 11 | 14 |
| 30 | N30 | 8 | Incl | 14 | 11 | 5 | 14 | 14 |

### *3.3.2 Sampling of environmental matrices from Bilbao*

The samples analysed in the present study were collected from two wastewater treatment plants located in the Basque Country: the Galindo wastewater treatment plant, a major industrial facility located near the municipalities of Barakaldo and Sestao in the metropolitan area of Bilbao, and the Crispijana wastewater treatment plant, located in the municipality of Vitoria-Gasteiz. In the case of the Galindo plant, which has already been the subject of several studies within previous research projects, sampling was carried out by selecting three representative points along the treatment line of the facility: the influent of the wastewater treatment plant (ERAZ), the inlet of the primary settling stage (DEKA sar), and the outlet of the tertiary treatment (TERT tre). For the first two sampling points, approximately 10 L of wastewater matrix were collected, divided into four 2.5 L bottles, whereas for the tertiary treatment outlet a larger sample volume of approximately 15 L was collected, distributed into six 2.5 L bottles. The samples, collected in mid-September, were mainly used for a preliminary familiarisation phase with the analytical procedure and for the optimisation of the sample preparation protocol.

Sampling at the Crispijana wastewater treatment plant was carried out on 21 October 2025 during the morning hours. In this case, five sampling points were identified, selected to represent the main stages of the wastewater treatment process (Figure 19): the plant influent (ERAZ), the inlet of the settling stage (DEKA sar), the inlet and outlet of the biological treatment (BIOL sar and BIOL irt), and the outlet of the tertiary treatment (TERT tre). Similarly to the sampling strategy adopted for the Galindo plant, approximately 10 L of samples were collected at the first two sampling points and distributed into four 2.5 L bottles, whereas for the remaining points approximately 15 L were collected and divided into six 2.5 L bottles. Sampling was carried out using automatic pumps dedicated to each sampling point, with the target sample volume digitally set for collection. This approach allowed the standardisation of the sampling operations and ensured the collection of adequate and comparable volumes for each analysed matrix. Once collected, the samples were transferred into sampling bottles and subsequently transported to the laboratory without any additional preliminary treatment. For each sampling point, each bottle was subsequently opened and its entire content was subjected to the sample preparation protocol described in the previous sections, which included pre-treatment steps, digestion of the organic matrix, and, when necessary, density separation. This procedure allowed the preparation of metallic filters suitable for subsequent microscopic observations and spectroscopic analyses, aimed at the identification and characterisation of microplastics (MPs) present in the different matrices along the treatment line of the wastewater treatment plants.

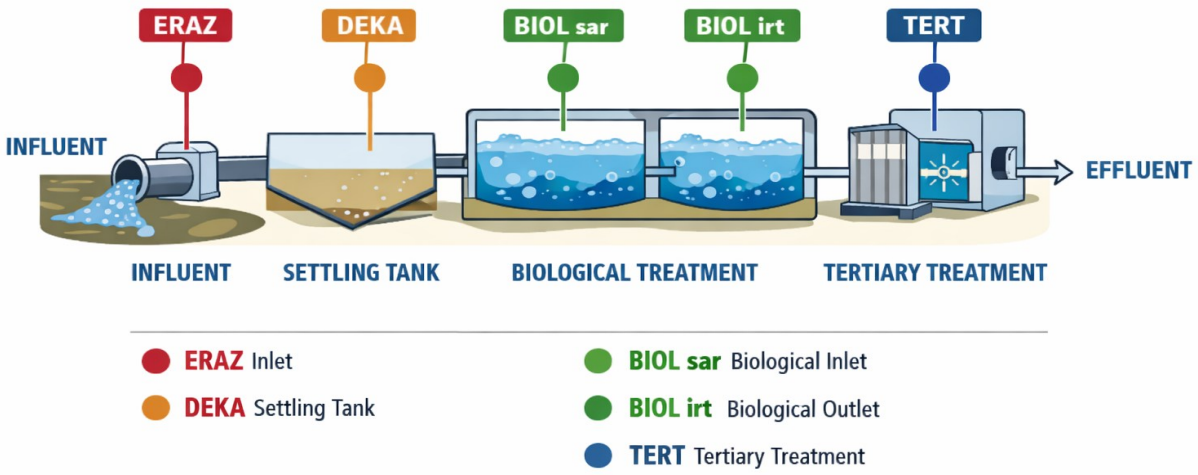



Figure 19. Principal sampling points in Crispijana station.

*3.3.2.1 Samples treatment from Bilbao.*

The samples analysed at the Martina Casiano Technological Platform of the University of the Basque Country (EHU/UPV), in collaboration with the IBeA Research Group, were subjected to a pre-treatment according to a specific laboratory procedure. In particular, following an extensive review of the scientific literature and taking into account methodologies previously adopted by the research group for the determination of microplastics (MPs) in other environmental matrices, an initial protocol was identified for the preparation of metallic filters intended for subsequent spectroscopic analyses (Figure 20)[63]. This methodology was initially subjected to a preliminary optimisation phase, during which I had the opportunity to become familiar with the laboratory experimental procedures. For this purpose, samples collected from the Galindo wastewater treatment plant were used.

[63] «trattamento campioni Bilbao».

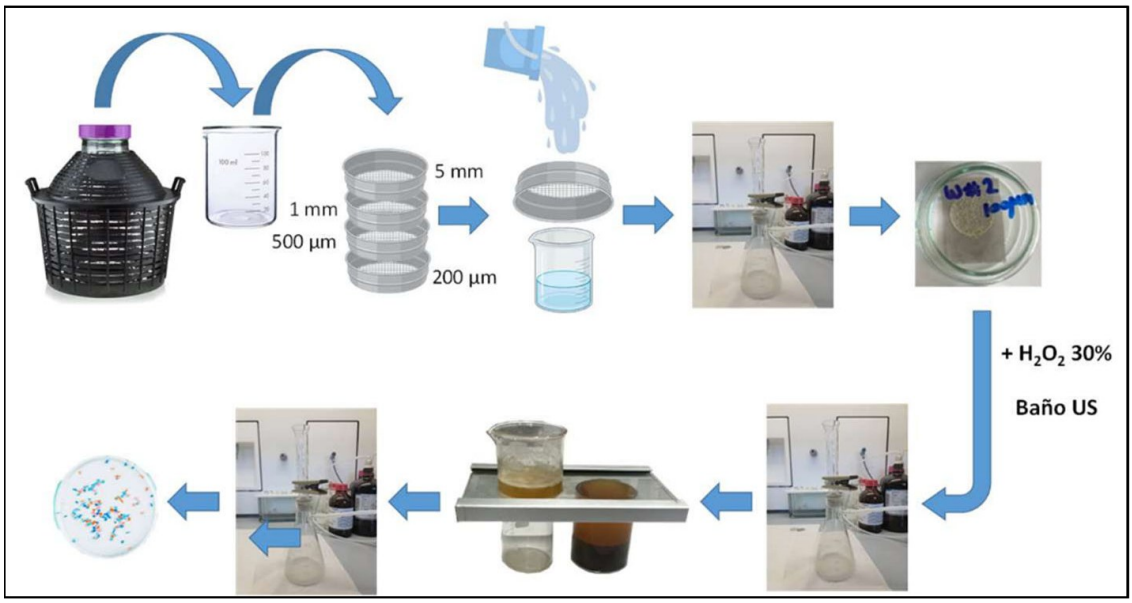



Figure 20. Previous procedure for wastewater treatment.

The liquid sample, representative of a single sampling point, was initially subjected to a particle-size separation step using a series of metallic sieves with different mesh sizes (1 mm, 0.5 mm, 0.2 mm and 0.05 mm). This step allowed the material to be divided into distinct size fractions, thereby reducing the complexity of the matrix and focusing the analysis on the particle size range relevant to microplastic investigation. During sieving, the sample was rinsed with water to facilitate the passage of particles through the meshes and to remove part of the finer matrix and adhering materials. The micrometric fractions were then individually recovered and subjected to vacuum filtration using a system consisting of a filtration funnel and a flask, allowing the solid fraction to be retained on a metallic filter. The resulting filter was subsequently transferred to a Petri dish and treated with hydrogen peroxide (30% $H_2O_2$) in order to perform an oxidative digestion of the organic matter present in the sample. This treatment was carried out in an ultrasonic bath, a condition that promotes particle dispersion and enhances the contact between the oxidising agent and the organic matrix, thereby increasing digestion efficiency. The use of hydrogen peroxide represents a widely adopted strategy in sample preparation protocols for microplastic analysis, as it enables the degradation of a large fraction of organic material without significantly altering the structure of most plastic polymers. At the end of the digestion step, the sample was again subjected to vacuum filtration to separate the liquid phase, containing the reaction products, from the remaining solid residue. In cases where the matrix was particularly rich in sediments or inorganic material, the obtained residue was further subjected to density separation (flotation) through the addition of a high-density sodium iodide (NaI) solution. This step exploits the density difference between plastic particles, which are generally less dense, and the heavier mineral fraction: microplastics tend to remain suspended or float, while sediments settle at the bottom of the container. The supernatant fraction, enriched in potentially plastic particles, was then recovered and filtered again under vacuum. The solid residue collected on the filter,

characterised by a reduced presence of organic material and mineral particulates, was finally used for the subsequent observation and identification of microplastics by spectroscopic techniques.

The application of the initial procedure highlighted some operational limitations, mainly related to the reduced effectiveness of the oxidative digestion step as the amount of interfering material on the filters increased. In particular, in samples characterised by a high content of organic residues—predominantly of cellulosic origin—the hydrogen peroxide treatment was sometimes insufficient to ensure complete removal of the organic matrix, thereby compromising the efficiency of the subsequent microplastic isolation steps.

In light of these observations, alternative strategies were evaluated to improve the selective removal of the cellulosic component. Following a review of the scientific literature, a preliminary chemical treatment using a urea/thiourea/NaOH solution (8%/6.5%/8% w/w, UTN) was introduced. This solution is known for its effectiveness in cellulose solubilisation and removal[64]. The implementation of this preliminary treatment significantly reduced the amount of fibrous organic material retained on the filters, thereby improving the conditions for the subsequent oxidative digestion step (Figure 21). After this stage, the sample was subjected to digestion with 33% hydrogen peroxide, performed on a heating and stirring plate at approximately 60 °C for about 6 hours, followed by a prolonged stirring phase at room temperature for an additional 18–20 hours. In cases where the filters exhibited a particularly high amount of organic residues, the agitation period could be extended up to 48 hours in order to ensure a more complete digestion of the matrix. In the optimised protocol, the use of an ultrasonic bath was avoided, as the introduction of the preliminary treatment with the UTN solution, combined with the prolonged oxidative digestion, proved to be sufficient for the removal of the organic fraction not relevant for the analysis. Finally, if inorganic residues were still present after the digestion step, the sample was subjected to density separation (flotation) using a high-density solution, according to the procedure previously described, in order to separate plastic particles from the heavier mineral fraction.

Overall, the optimised methodology proved to be effective for sample treatment and for the preparation of metallic filters (Figure 22), enabling the obtainment of sufficiently clean surfaces suitable for subsequent microscopic observations and spectroscopic analyses, which are essential for the identification and characterisation of microplastics (MPs).

[64] Olsen et al., «Facilitating Microplastic Quantification through the Introduction of a Cellulose Dissolution Step Prior to Oxidation».

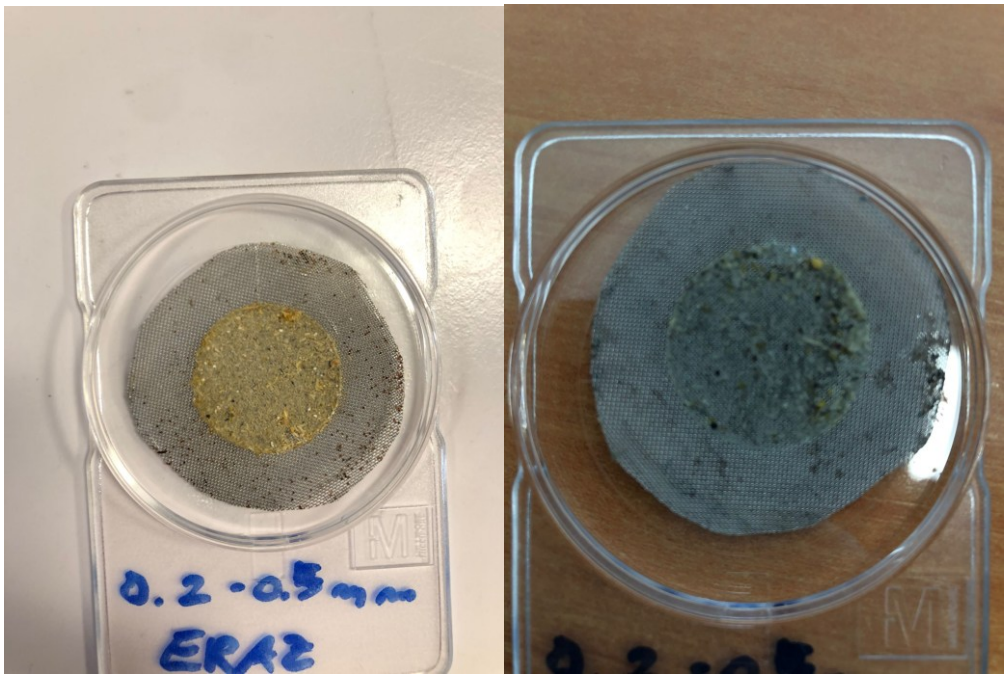



Figure 21. Metallic filter of ERAZ fraction pre (left) and post treatment (right) with UTN solution.

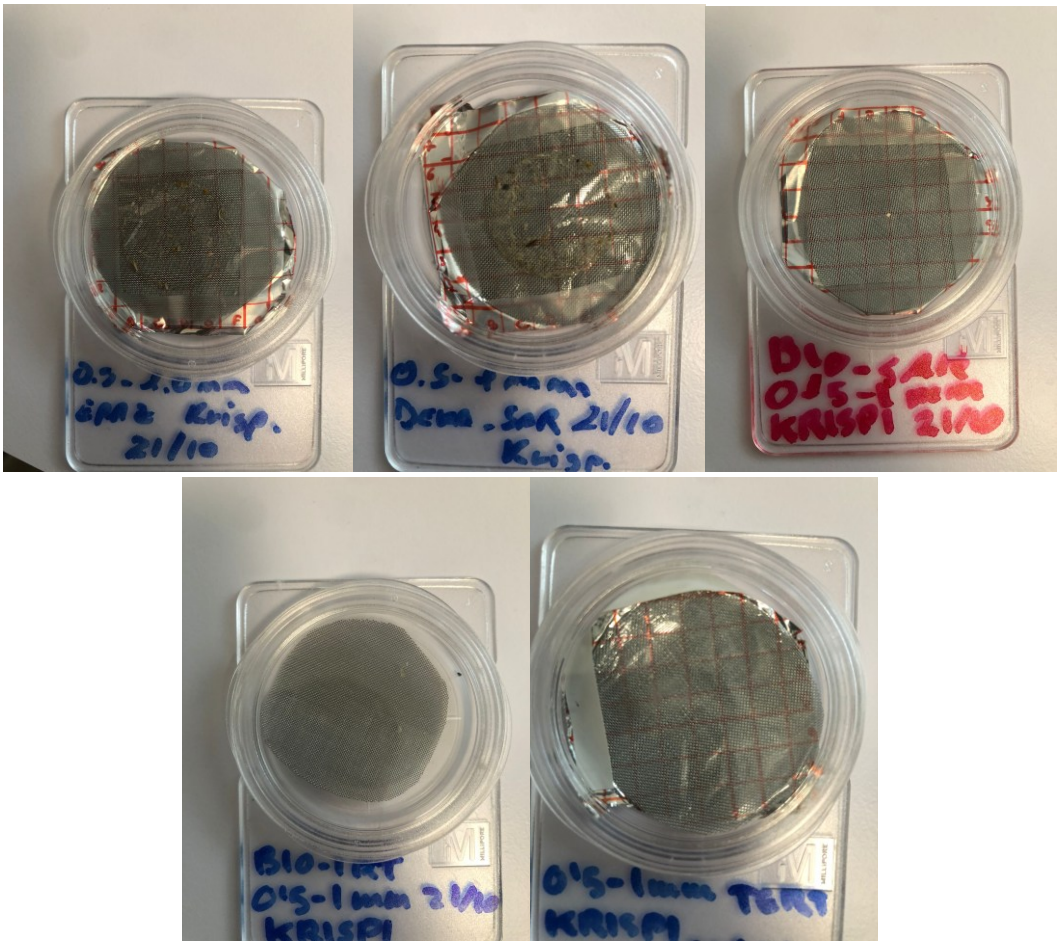


Figure 22. Metallic filters corresponding to one size fraction after the analytical treatment described above: ERAZ, DEKA sar, BIOLsar, BIOL irt and TERT tre.

## 4. Results and Discussion

### *4.1 Results and Discussion of Modena*

The MCR-ALS analysis performed independently on the two Raman spectral regions investigated (200–1400 $cm^{-1}$ and 1400–2000 $cm^{-1}$) provided highly comparable results, confirming the consistency of the chemometric approach across the different acquisition windows of the Horiba Raman system. For both spectral intervals, the optimisation procedure indicated that a four-component model represented the most appropriate solution in terms of balance between spectral resolution, chemical interpretability and model stability. Increasing the number of components beyond this value resulted in the generation of additional spectral and concentration profiles (Sopt and Copt) that did not provide further relevant chemical information. Instead, the additional components showed strong similarities with already resolved profiles, suggesting that the model was progressively describing minor variations or redundant spectral features rather than independent chemical contributions. Therefore, the four-component solution was selected as the most representative description of the main sources of spectral variability present within the analysed Raman maps. The first resolved component represented the dominant contribution within both spectral regions and was characterised by a concentration profile that closely reproduced the spatial distribution of the sorbent material observed in the original Raman mapping images (Figure 23). The corresponding resolved spectral profile exhibited the characteristic Raman features associated with the HLB resin (Figure 24), confirming that the main contribution detected by the instrument was related to the polymeric sorbent matrix itself. This result was expected considering the composition of the analysed POCIS, where the resin represents the major component within the sampled area and therefore generates the most intense and spatially distributed Raman signal. The ability of the MCR-ALS model to reconstruct the morphology of the sorbent demonstrates that the applied workflow was able to correctly extract the dominant spectral contribution and correlate it with its actual spatial localisation. The second resolved component (Figure 25) showed a complementary spatial distribution compared with the first one, being mainly concentrated in the areas surrounding the resin particles rather than within the regions dominated by the sorbent signal. This behaviour suggests that this component describes spectral contributions associated with the fraction of the sample different from the dominant HLB resin contribution. From an environmental perspective, this observation is particularly relevant because the regions surrounding the sorbent particles may represent the areas where adsorbed compounds or minor chemical constituents contribute to the overall Raman response. However, the interpretation of this component requires caution. The corresponding spectral profile did not show a clear and unique fingerprint attributable to a single chemical compound, but rather appeared to contain a combination of overlapping spectral contributions. This behaviour is consistent with the intrinsic complexity of POCIS matrices, where multiple compounds may be simultaneously retained by the sorbent and where their concentration levels may be significantly lower compared with the signal generated

by the resin itself. Therefore, although the second component may contain information related to retained contaminants of emerging concern, the current dataset does not allow an unambiguous identification of individual compounds. Instead, this component should be considered as a mixed spectral contribution potentially associated with minor chemical species present within the analysed area. The third and fourth components (Figure 26 and Figure 27 respectively) provided additional information regarding localised spectral variability within the Raman maps. Their concentration profiles revealed spatial features that were not completely represented by the second component, suggesting the presence of further minor contributions distributed across the analysed region. Nevertheless, these components were progressively affected by an increase in spectral noise and lower signal intensity, which limited their direct chemical interpretation. In particular, the corresponding spectral profiles appeared less defined compared with the first two components, indicating that part of the explained variance may be related to residual instrumental effects, background contributions or minor fluctuations within the sample matrix. For this reason, these components were considered potentially informative for describing additional variability, but their direct association with specific chemical species would require further experimental confirmation. Overall, the obtained results demonstrate the potential of combining Raman mapping and MCR-ALS analysis for investigating the chemical heterogeneity of passive sampling devices. The applied workflow successfully separated the dominant contribution of the HLB sorbent from additional spectral features associated with minor components of the analysed matrix. At the same time, the results highlight one of the main challenges associated with the direct spectroscopic investigation of environmental samples: the strong imbalance between the intense signal generated by the sorbent material and the weaker contributions expected from trace-level contaminants. In this context, the use of chemometric approaches represents an essential tool for enhancing spectral information extraction, although additional validation strategies are required before assigning specific chemical identities to the minor resolved components.

The results obtained in this work also provide important indications for future improvements of the analytical strategy. First, the current evaluation was performed on a limited portion of the POCIS surface and on a restricted number of Raman maps. A more extensive acquisition strategy, including multiple mapping areas on the same sampler and replicate analyses of different POCIS devices, would allow the reproducibility of the resolved components to be assessed. In particular, repeated observations of the second component across different regions and samples would help determine whether the associated spectral contribution is consistently related to retained contaminants or whether it mainly originates from intrinsic variability of the sorbent matrix. Increasing the number of analysed areas would therefore provide a stronger statistical basis for distinguishing meaningful chemical information from random variability or noise. A further improvement could involve the refinement of the spectral selection procedure used for the generation of the initial estimates for MCR-ALS optimisation. Although the PCA-based approach implemented in the developed MATLAB workflow successfully identified representative and spectrally distinct profiles, additional selection criteria could be introduced to further restrict the choice of initial spectra. In particular,

preventing the selection of highly similar or partially redundant spectra could improve the uniqueness of the MCR solution and reduce the possibility of generating mathematically valid but chemically redundant components. This aspect becomes particularly relevant when analysing complex environmental samples, where small spectral differences may arise from noise or matrix effects rather than from distinct chemical species. Another relevant development would be the implementation of an interactive visualisation platform capable of directly integrating the MCR concentration maps with the original Raman images. Such an approach would allow the user to select individual pixels or specific regions of interest directly from the reconstructed chemical maps and immediately visualise the corresponding Raman spectrum. This functionality would significantly facilitate the interpretation of the resolved components, allowing a more direct correlation between the spatial distribution of spectral contributions and their original experimental signal. Moreover, it could support targeted re-analysis of specific regions potentially enriched in minor chemical constituents. Finally, the integration of external Raman spectral databases containing reference spectra of contaminants of emerging concern could represent a valuable step towards improving the chemical interpretation of the resolved profiles. While the current workflow successfully identifies spectral differences between the sorbent matrix and additional contributions, the comparison with curated reference libraries could support the tentative assignment of unknown spectral features and facilitate the recognition of specific contaminant classes. In combination with automated spectral matching algorithms and advanced chemometric approaches, such developments could contribute to the creation of more robust and semi-automated workflows for the rapid screening of passive samplers and environmental matrices.

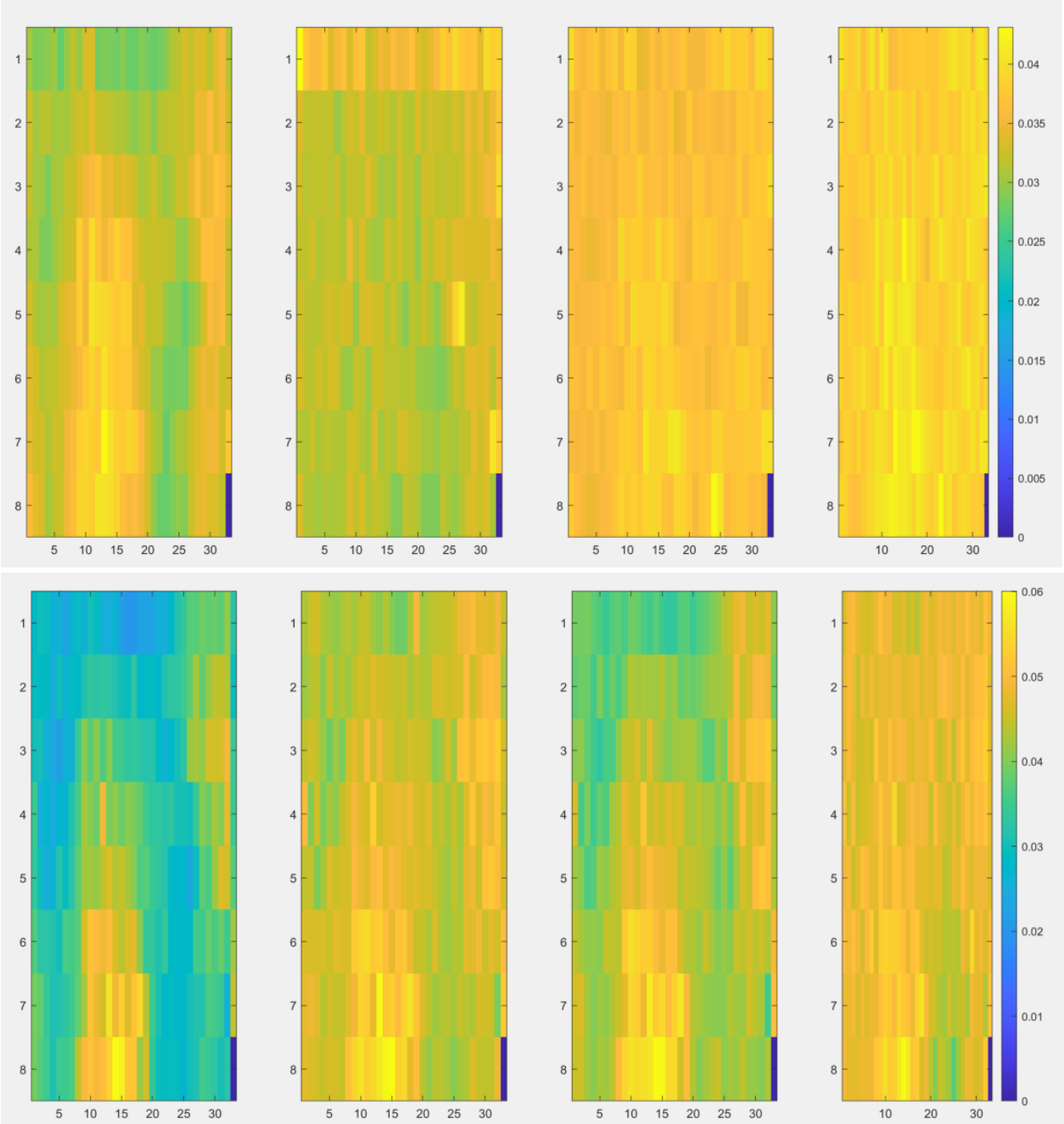



Figure 23. Above the concentration profiles of the first spectral range and then the concentration profiles of the second spectral range.


Figure 24. First resolve spectral profiles: HLB resine Raman spectra.


Figure 25. Second resolve spectral profile.

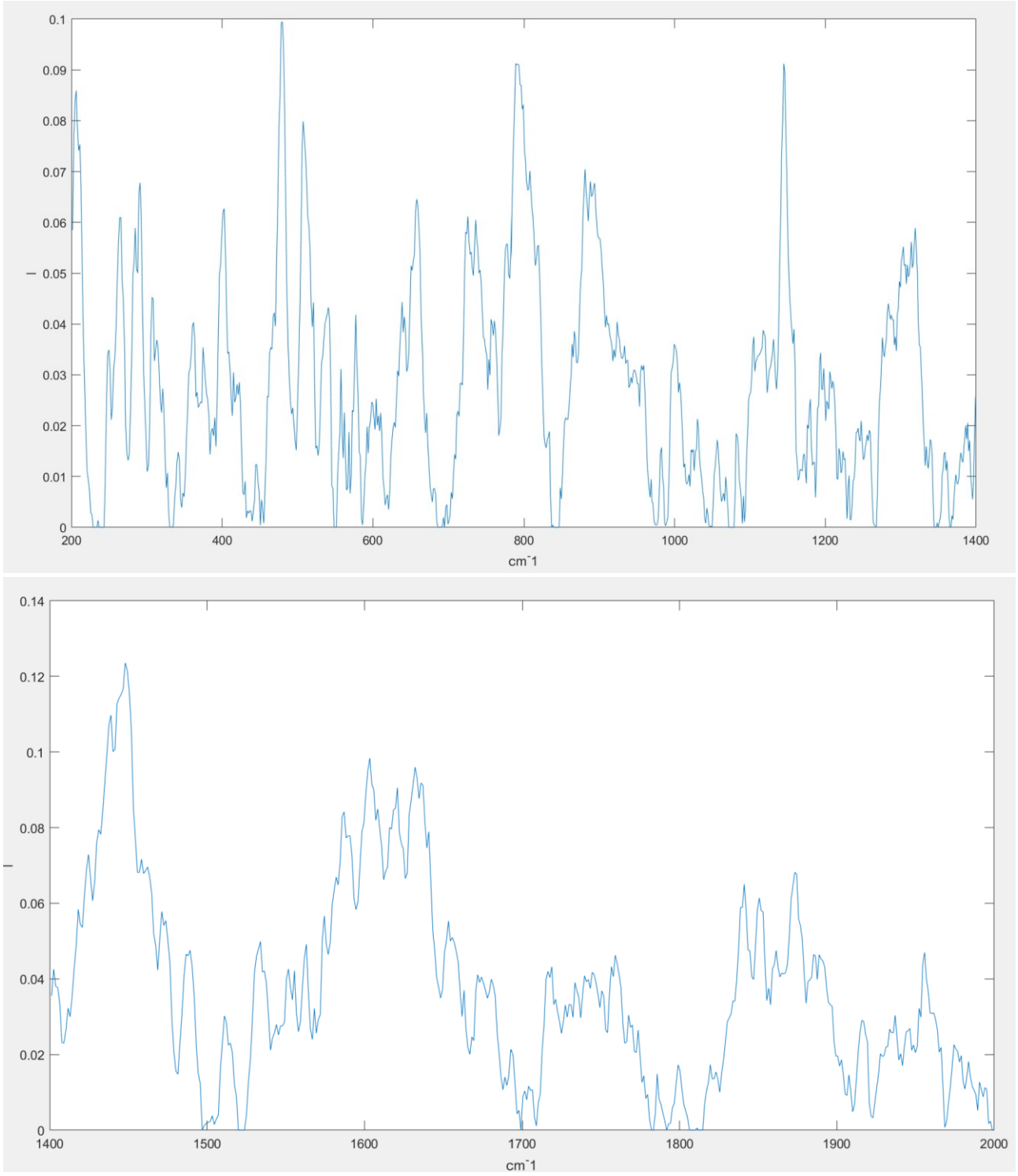



Figure 26. Third resolve spectral profile.


Figure 27. Fourth resolve spectral profile.

The same chemometric workflow previously applied to the Raman datasets was subsequently extended to the hyperspectral images acquired on the POCIS samples. The main difference compared with Raman mapping was represented by the considerably larger size of the hyperspectral datasets, with each analysed image containing approximately 170,000 individual spectra. This substantial increase in the number of observations provided a much more comprehensive description of the spatial variability of the analysed passive sampler

surface. However, the higher amount of spectral information also resulted in increased computational complexity and highlighted the challenges associated with the extraction of chemically meaningful information from highly heterogeneous environmental matrices. The MCR-ALS analysis was successfully performed on the hyperspectral datasets, demonstrating the capability of the developed workflow to handle large hyperspectral matrices. Nevertheless, the interpretation of the obtained results was considerably more challenging compared with the Raman mapping approach. Although the model generated resolved concentration (Copt, Figure 28) and spectral (Sopt, Figure 29) profiles, the extracted components did not provide a clear chemical separation between the dominant contribution of the HLB sorbent and the potential spectral contribution associated with retained contaminants. The main limitation was observed during the interpretation of the resolved spectral profiles. The preliminary evaluation of the dataset through Singular Value Decomposition (SVD) suggested that five components were sufficient to describe the main variance contributions within the hyperspectral images. These components were subsequently considered during the MCR-ALS optimisation. However, the resulting Sopt profiles showed a marked degree of redundancy, with several components exhibiting highly similar spectral characteristics rather than representing chemically independent contributions. Instead of resolving distinct constituents, the model tended to distribute small variations of the dominant spectral contribution among different components. This behaviour indicates that the limitation was not related to the MCR-ALS algorithm itself, but rather to the intrinsic information content of the hyperspectral dataset. Indeed, the capability of MCR-ALS to resolve chemically meaningful profiles is strongly dependent on the presence of sufficiently distinct spectral signatures within the analysed matrix. When different constituents exhibit strongly overlapping spectral responses, the algorithm may generate mathematically valid solutions that are difficult to interpret from a chemical perspective. In the present case, the observed redundancy among the Sopt profiles reflects the limited spectral selectivity available in the original hyperspectral data, rather than an instability of the chemometric model. This behaviour can be explained by considering the intrinsic characteristics of near-infrared spectroscopy. Unlike Raman spectroscopy, which provides narrow and highly specific bands associated with fundamental molecular vibrations, NIR spectroscopy is based on overtone and combination transitions. These absorption features are intrinsically broader, less intense and strongly overlapped, resulting in lower chemical specificity. This limitation becomes particularly relevant when analysing complex environmental matrices such as POCIS, where the retained contaminants are expected to be present at trace concentration levels, while the sorbent material represents the overwhelming fraction of the analysed sample. Consequently, the spectral variance of the dataset is largely dominated by the HLB resin, whereas the contribution of the adsorbed contaminants is expected to be significantly weaker and partially masked by the natural variability of the sorbent matrix. The results obtained also highlight an important distinction between spatial resolution and chemical resolution. The hyperspectral camera provided an exceptionally high number of spatially resolved measurements, allowing the entire POCIS surface to be characterised. However, the increase in the number of acquired pixels did not directly

correspond to a proportional increase in chemical information. In highly homogeneous regions, adjacent pixels may contain almost identical spectral information, leading to a high degree of redundancy within the dataset. Therefore, although hyperspectral imaging improves spatial coverage and enables the identification of local spectral variations, it does not necessarily guarantee enhanced chemical discrimination when the spectral signatures of the analysed compounds are weak and strongly overlapped. Despite these limitations, the resolved concentration profiles provided valuable information regarding the spatial organisation of the analysed POCIS samples. The Copt maps highlighted regions characterised by different spectral responses across the surface, appearing as "topographic-like distributions" rather than simple reproductions of the resin morphology (Figure 28). Areas with higher relative concentration values may correspond to regions where minor chemical contributions, potentially associated with retained contaminants or other matrix components, were preferentially accumulated, whereas areas with lower values were mainly associated with the dominant sorbent contribution. However, due to the lack of sufficiently selective spectral profiles, these regions cannot be unambiguously assigned to specific contaminants or individual chemical classes. Overall, the obtained results indicate that hyperspectral imaging, despite its advantages in terms of rapid acquisition, non-destructive analysis and extensive spatial information, presents significant limitations for the untargeted identification of emerging contaminants in complex passive sampling matrices. In this context, the main analytical challenge is not related to the availability of large spectral datasets, but rather to the extraction of chemically meaningful information from measurements dominated by the sorbent matrix. Increasing the spatial resolution alone is therefore unlikely to overcome the limitations associated with insufficient chemical selectivity. Several future developments may improve the applicability of HSI for this type of analysis. From an instrumental perspective, the use of higher magnification objectives or optical configurations characterised by a reduced field of view could decrease spatial averaging effects and allow the investigation of more localised regions of the passive sampler surface. By reducing the sampled area associated with each pixel, it may become possible to enhance the detection of localised chemical heterogeneities and improve the probability of identifying regions enriched in retained contaminants. From a data analysis perspective, the implementation of intelligent pixel-selection strategies may represent a valuable improvement. Instead of processing the entire hyperspectral cube, algorithms based on exploratory chemometric approaches, clustering methods or machine-learning techniques could be used to identify the most informative regions before MCR-ALS optimisation. Such strategies would reduce computational requirements while increasing the probability of extracting chemically meaningful components from the dataset. Furthermore, the integration of hyperspectral imaging with complementary spectroscopic techniques represents a promising strategy for overcoming the limitations observed in the present work. While HSI provides rapid spatial screening and allows the identification of regions characterised by different spectral behaviours, Raman spectroscopy offers higher molecular specificity and may provide the detailed chemical information required for the characterisation of selected areas. The combination of both

approaches through data-fusion strategies could therefore exploit the complementary advantages of each technique. Additionally, the development of dedicated spectral libraries containing reference spectra of emerging contaminants could support future identification workflows through similarity matching and advanced classification models. The present results demonstrate that, in the context of passive sampler analysis, the main analytical challenge is not simply the acquisition of a large amount of spectral data, but the ability to extract reliable chemical information from highly complex and heterogeneous environmental matrices. Therefore, future improvements should focus on the combination of enhanced instrumental selectivity, advanced chemometric approaches and multimodal analytical strategies rather than solely increasing spatial resolution or dataset size.

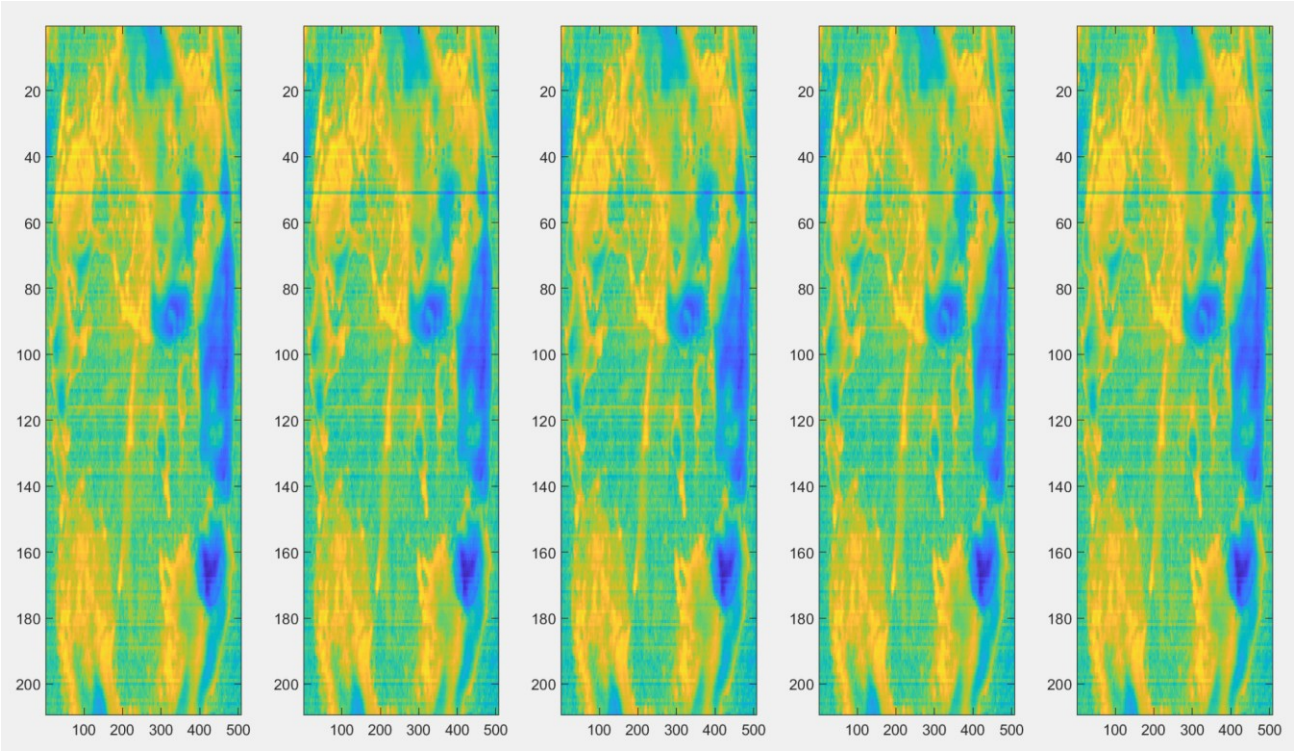
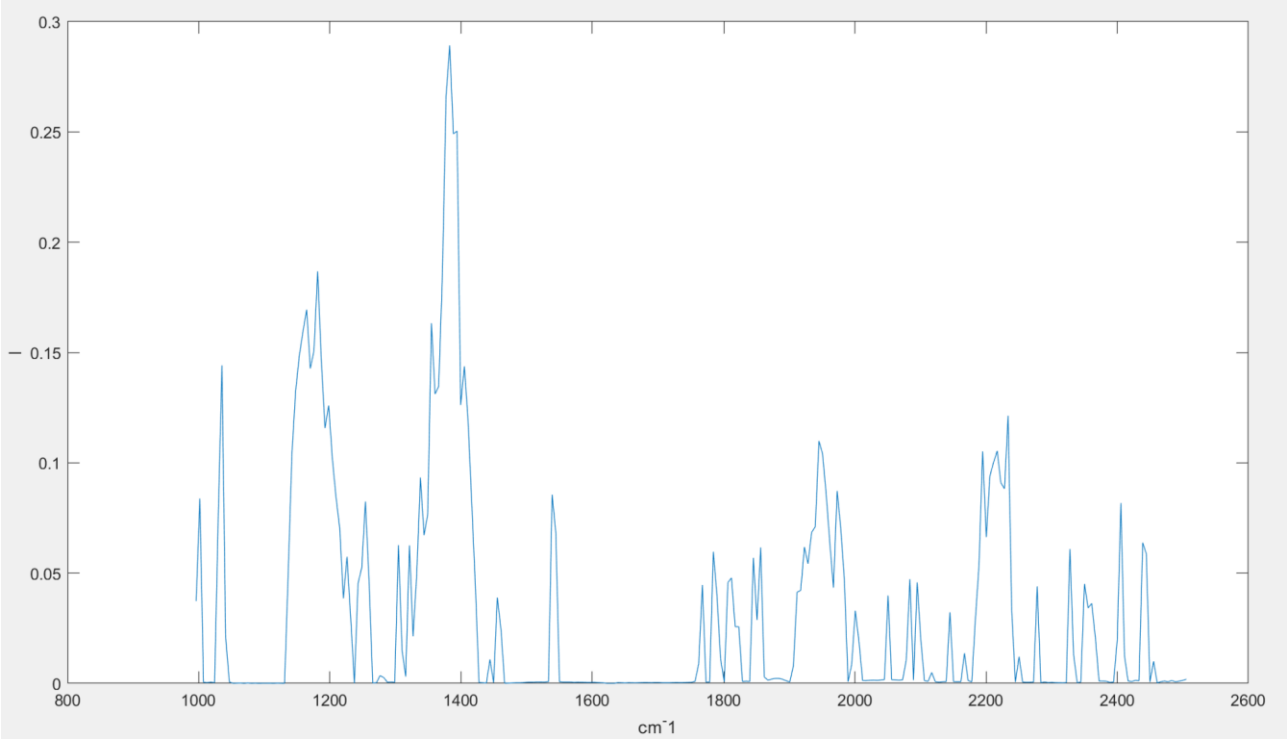
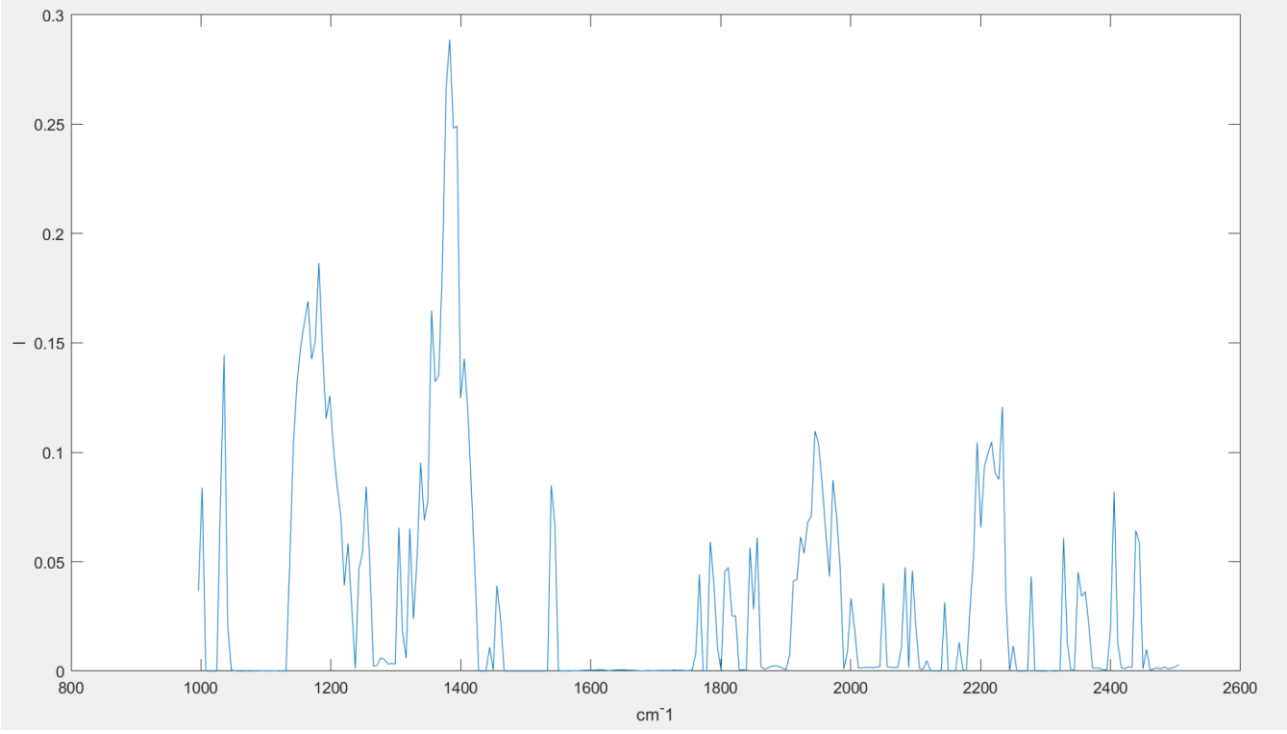
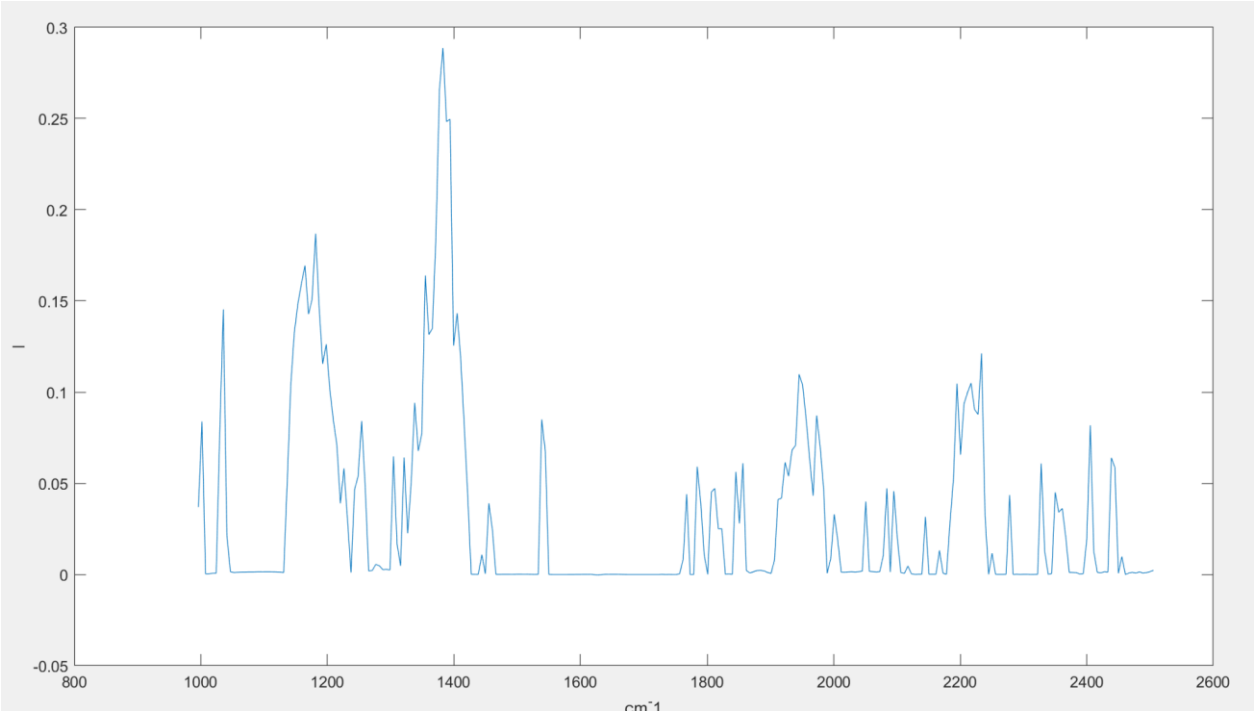
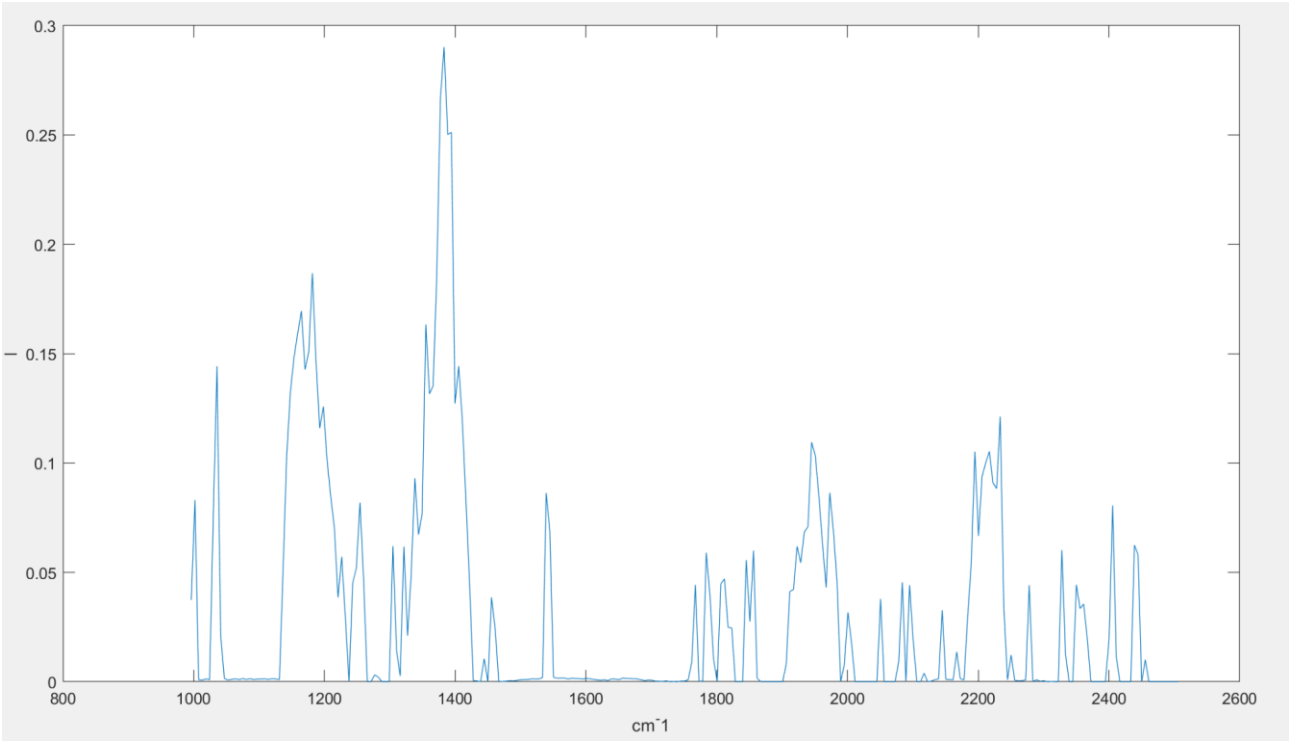



Figure 28. five resolve (redundant) concentration profiles.


0.3
0.25
0.2
I 0.15
0.1
0.05
0
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
cm^-1


0.3
0.25
0.2
I 0.15
0.1
0.05
0
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
cm^-1


0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
I
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
cm-1


0.3
0.25
0.2
0.15
0.1
0.05
0
I
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
cm-1

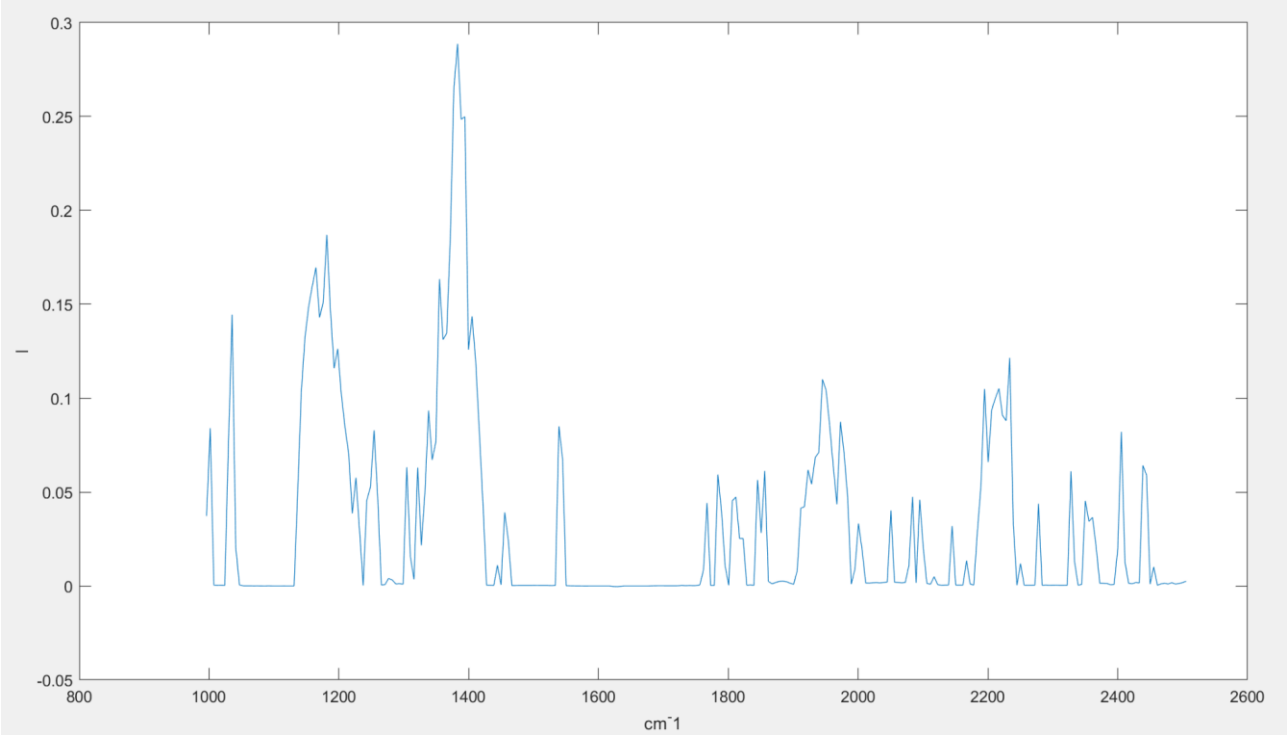



Figure 29. 5 resolved (redundant) spectra profiles.

The MicroNIR datasets were analysed by applying two complementary constrained MCR-ALS strategies with the aim of assessing the capability of near-infrared spectroscopy to identify emerging contaminants embedded within a complex sorbent matrix. Unlike the Raman mapping experiments, where MCR-ALS was primarily employed to reconstruct the spatial distribution of the resolved chemical components, the MicroNIR analysis focused on the qualitative identification of the target analytes and on evaluating whether the introduction of prior chemical knowledge could improve component resolution. This aspect is particularly relevant in near-infrared spectroscopy, where the inherently broad and overlapping overtone and combination bands considerably reduce spectral selectivity compared with vibrational techniques based on fundamental transitions, such as Raman spectroscopy. To maximise the extraction of chemically meaningful information, two different constrained MCR-ALS approaches were investigated. Although both models relied on the same bilinear decomposition algorithm, they differed in the way prior information was incorporated into the optimisation procedure. The first approach exploited prior knowledge regarding the expected occurrence of the investigated compounds within the experimental datasets, whereas the second introduced selective constraints directly on spectral regions characteristic of the individual reference standards. The rationale behind the comparison of these two strategies was to evaluate whether different forms of prior information could enhance the capability of the algorithm to resolve weak spectral contributions masked by the dominant signal of the HLB sorbent. Overall, both approaches successfully resolved chemically meaningful spectral profiles, demonstrating that the introduction of appropriate constraints substantially improves the interpretability of complex NIR datasets. In particular, the multiset strategy enabled the identification of the spectrum representing the complete experimental mixture, the spectral

profile of the HLB sorbent and three of the five investigated contaminants, namely caffeine, hydrochlorothiazide and atenolol (Figures 30, 31, 32, 33, 34, 35). These results indicate that the model was capable of recovering the most spectrally distinctive constituents within the dataset despite the extensive band overlap that characterises near-infrared spectra. An improvement in spectral resolution was observed when the optimisation was performed using equality constraints (Figures 36, 37, 38, 39, 40, 41). In this case, the resolved Sopt profiles were representative of all constituents included in the experimental design, namely the complete mixture, the HLB sorbent and the five investigated standards (caffeine, hydrochlorothiazide, diclofenac, metoprolol and atenolol). Compared with the multiset approach, the incorporation of spectral constraints produced components that were more consistent with the known composition of the samples, suggesting that directing the optimisation towards characteristic wavelength regions improves the capability of MCR-ALS to resolve highly overlapping spectral features. These observations are consistent with previous studies demonstrating that the incorporation of chemically meaningful constraints significantly reduces solution ambiguity and enhances the interpretability of MCR-ALS models, particularly when analysing complex mixtures characterised by limited spectral selectivity. Despite the satisfactory resolution of the spectral profiles, both modelling strategies exhibited a common limitation concerning the resolved concentration profiles (Copt). The simulated SPE samples were prepared according to a Design of Experiments (DoE), in which the concentration of each pharmaceutical standard was systematically varied over predefined concentration levels while maintaining a controlled experimental design. Consequently, the resolved concentration profiles were expected to reproduce concentration trends coherent with the DoE, showing systematic variations associated with the nominal concentration of each analyte. Contrary to these expectations, neither constrained approach generated concentration profiles clearly reflecting the experimental design. Instead, the resolved Copt distributions appeared substantially random, without exhibiting concentration trends proportional to the known composition of the samples. This behaviour suggests that, although the algorithm successfully recovered chemically meaningful spectral signatures, it was unable to accurately reconstruct the quantitative contribution of the individual analytes within the investigated concentration range. Several factors may account for this behaviour. First, the analysed samples consisted predominantly of the HLB sorbent, whereas the pharmaceutical standards represented only a minor fraction of the total sample mass. Consequently, the spectral variance of the dataset was largely dominated by the polymeric matrix, while the variability associated with the concentration differences introduced by the experimental design represented only a limited contribution to the overall variance. Secondly, the intrinsic characteristics of near-infrared spectroscopy further complicate quantitative resolution. Since NIR spectra originate from overtone and combination vibrations, the resulting absorption bands are broad, highly overlapped and considerably less chemically selective than Raman bands. Under these conditions, relatively small concentration variations may produce spectral differences that are comparable to, or even smaller than, the natural variability associated with the sorbent itself. Moreover, although the simulated SPE samples were prepared following a rigorously defined Design of

Experiments, slight local heterogeneities cannot be completely excluded. During sample preparation, the pharmaceutical standards were dissolved together with the HLB sorbent in methanol, allowed to dry for approximately 24 hours and subsequently transferred into plastic sachets prior to spectral acquisition. Minor differences in the local distribution of the analytes within the sorbent after solvent evaporation may therefore have introduced additional variability not directly related to the nominal concentration levels defined by the DoE. Such effects may have further reduced the capability of MCR-ALS to reconstruct concentration profiles consistent with the experimental design. It should also be emphasised that MCR-ALS is fundamentally designed to minimise the reconstruction error of the experimental data matrix rather than to reproduce external quantitative information. Consequently, the successful recovery of chemically meaningful spectral profiles does not necessarily imply an equally accurate estimation of the corresponding concentration profiles, particularly when the available chemical variance is limited and dominated by a single matrix component. The results obtained in the present study therefore suggest that the constrained models were particularly effective for qualitative chemical interpretation, whereas their quantitative capability remained limited under the investigated experimental conditions. From a methodological perspective, the obtained results highlight several opportunities for future developments. A promising strategy would consist of combining the complementary advantages of the two constrained approaches investigated in this work within a single hybrid MCR-ALS framework. Specifically, the multiset strategy could provide information regarding the expected occurrence of each analyte within the different experimental datasets, while equality constraints could simultaneously guide the optimisation towards the spectral regions characteristic of the individual reference compounds. The simultaneous incorporation of prior information at both the dataset and spectral levels may substantially reduce rotational ambiguity and improve both the robustness and the chemical interpretability of the resolved solutions. A further improvement could involve the incorporation of constraints directly related to the concentration profiles. Since the simulated SPE samples were prepared according to a predefined Design of Experiments, the nominal concentration levels of each pharmaceutical compound are known *a priori*. This information could be exploited during the optimisation by introducing concentration-based constraints capable of guiding the algorithm towards solutions consistent with the experimental design. Integrating the DoE within the MCR-ALS optimisation may therefore improve the reconstruction of the resolved concentration profiles while preserving the quality of the spectral resolution. More generally, future developments should aim at integrating multiple sources of prior knowledge within a unified chemometric framework. Rather than relying exclusively on spectral information, MCR-ALS could simultaneously exploit information related to sample composition, spectral selectivity, experimental design and expected concentration gradients. Such a multi-constrained strategy would likely improve both the qualitative identification of the investigated analytes and the quantitative reliability of the resolved concentration profiles.

Finally, the combination of constrained MCR-ALS with supervised chemometric methods, such as Partial Least Squares (PLS) regression, or with machine-learning algorithms

represents another promising research direction. While MCR-ALS is particularly effective in resolving chemically meaningful spectral profiles, supervised predictive models may exploit the known concentration values defined by the experimental design to improve quantitative predictions. The integration of these complementary methodologies may therefore contribute to the development of robust analytical workflows capable of supporting rapid, sustainable and reliable environmental screening using portable near-infrared instrumentation.

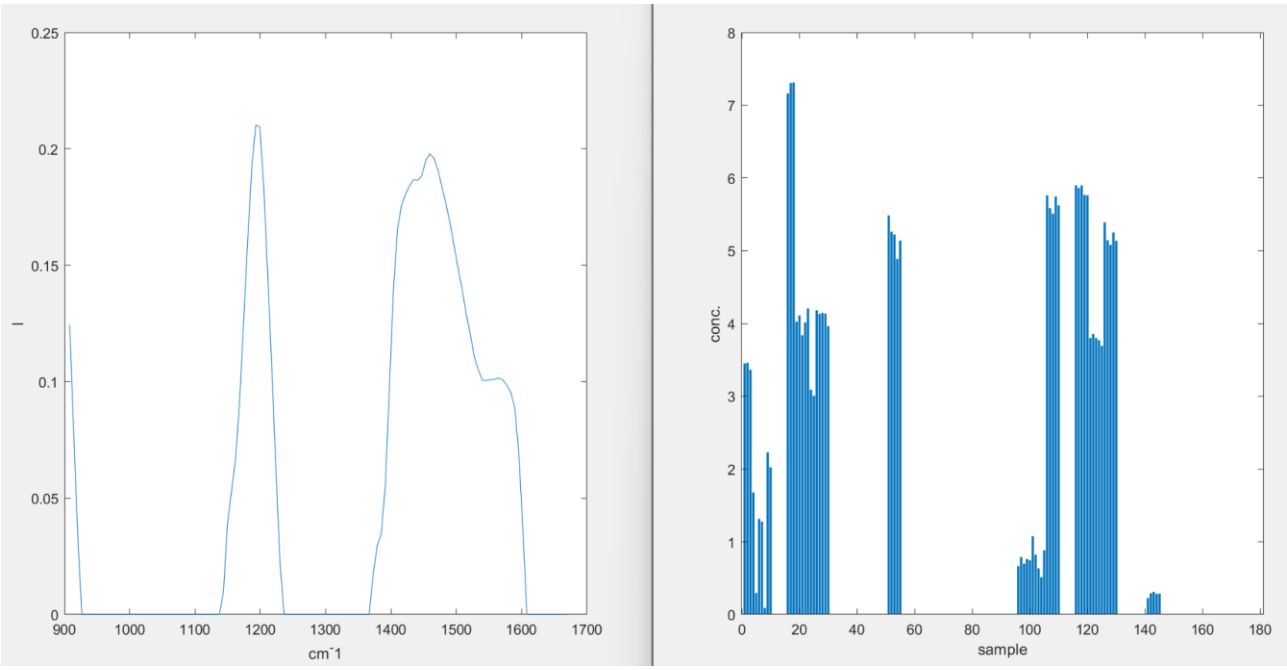



Figure 30. First resolved component by multiset.

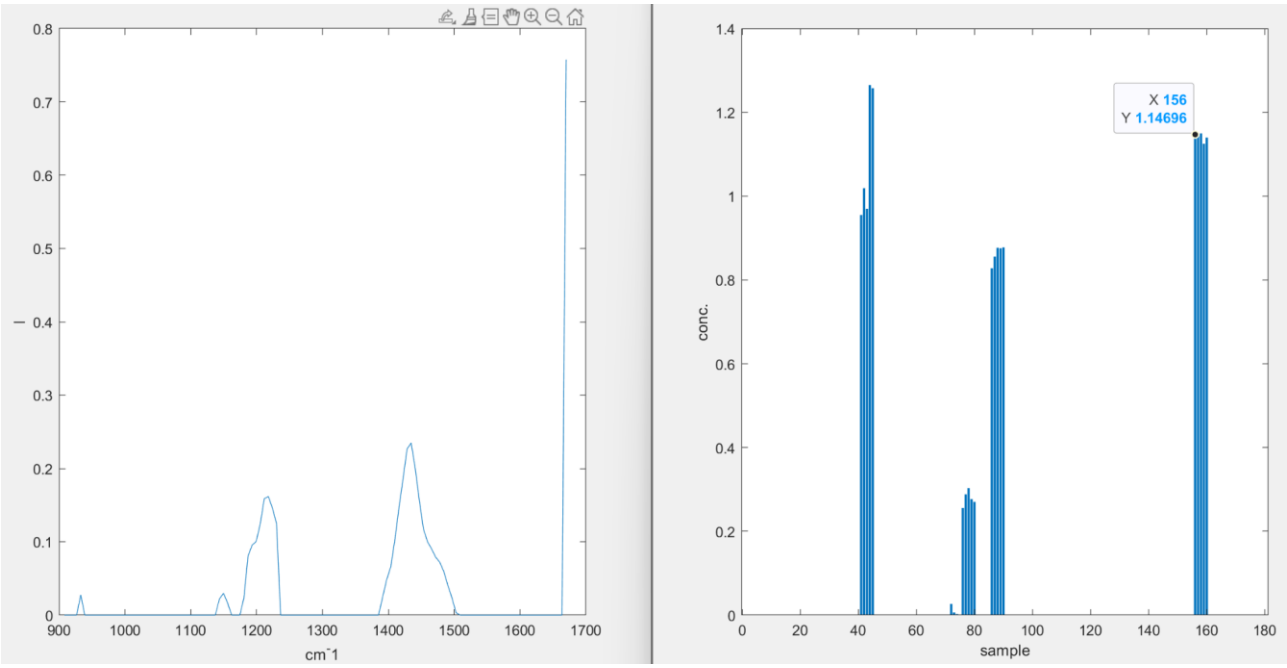



Figure 31. Second resolved component by multiset (Caffeine).

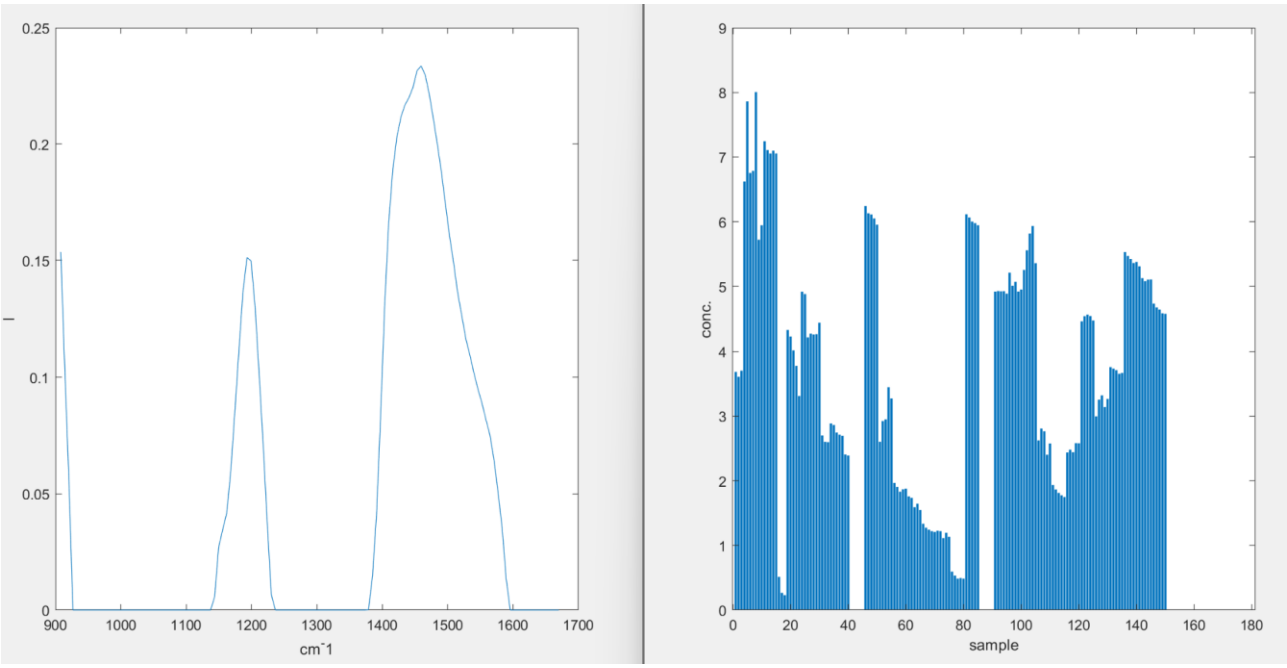



Figure 32. Third resolved component by multiset.

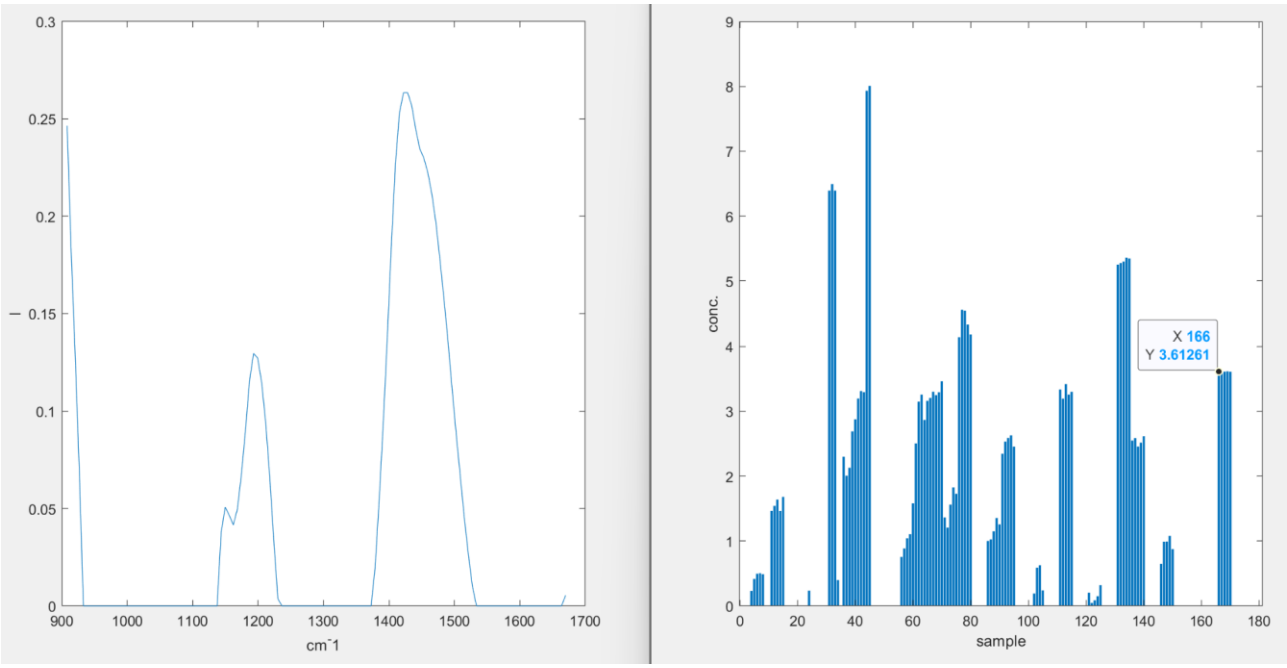



Figure 33. Fourth resolved component by multiset (Hydrochlorothiazide).

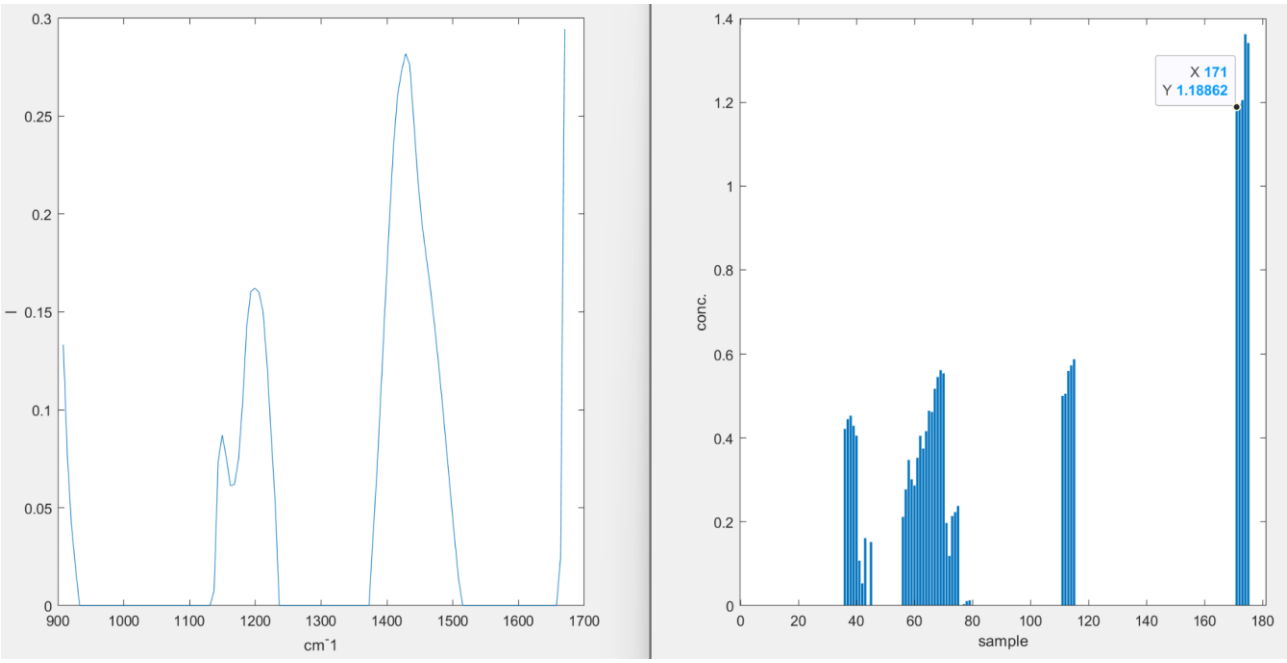



Figure 34. Fifth resolved component by multiset (Atenolol).

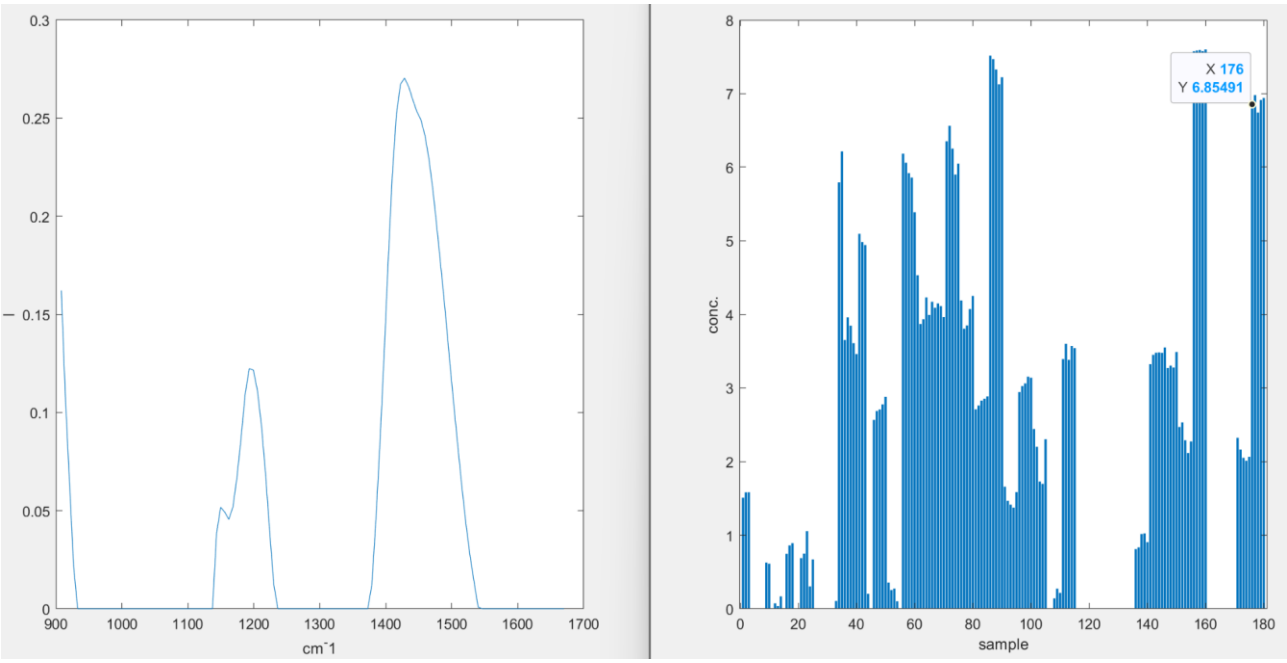



Figure 35. Sixth resolved component by multiset (HLB resine).

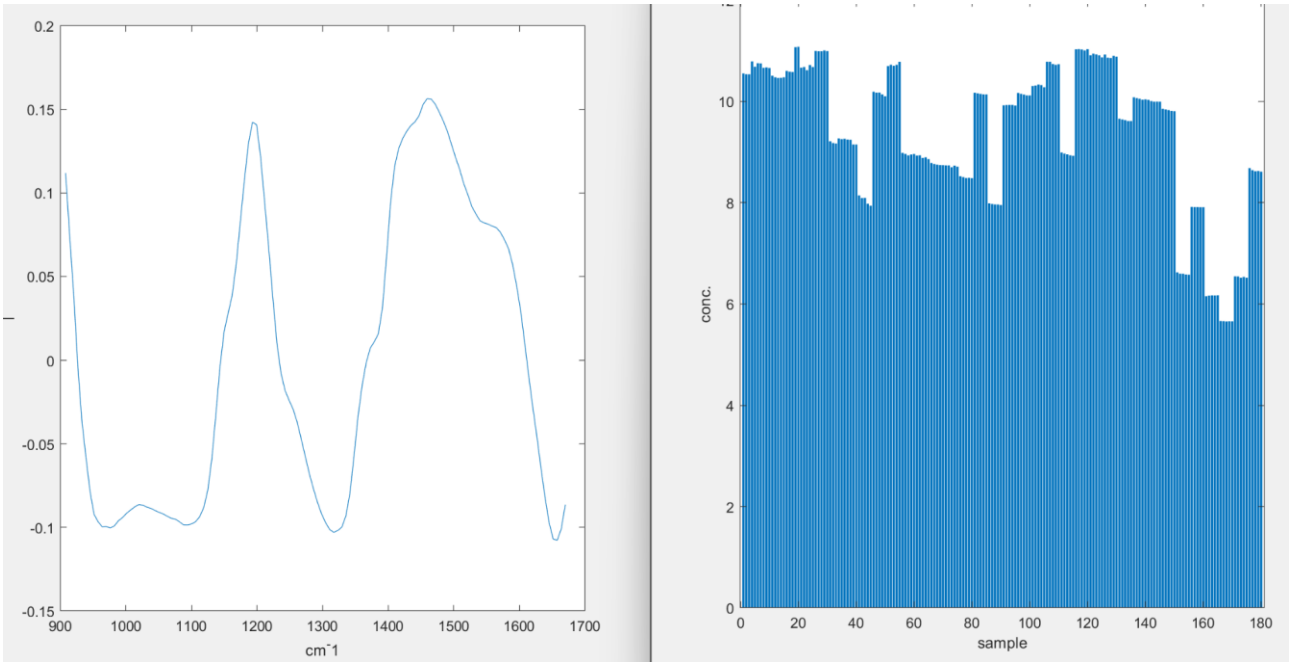



Figure 36. First resolved component by equality constraint.

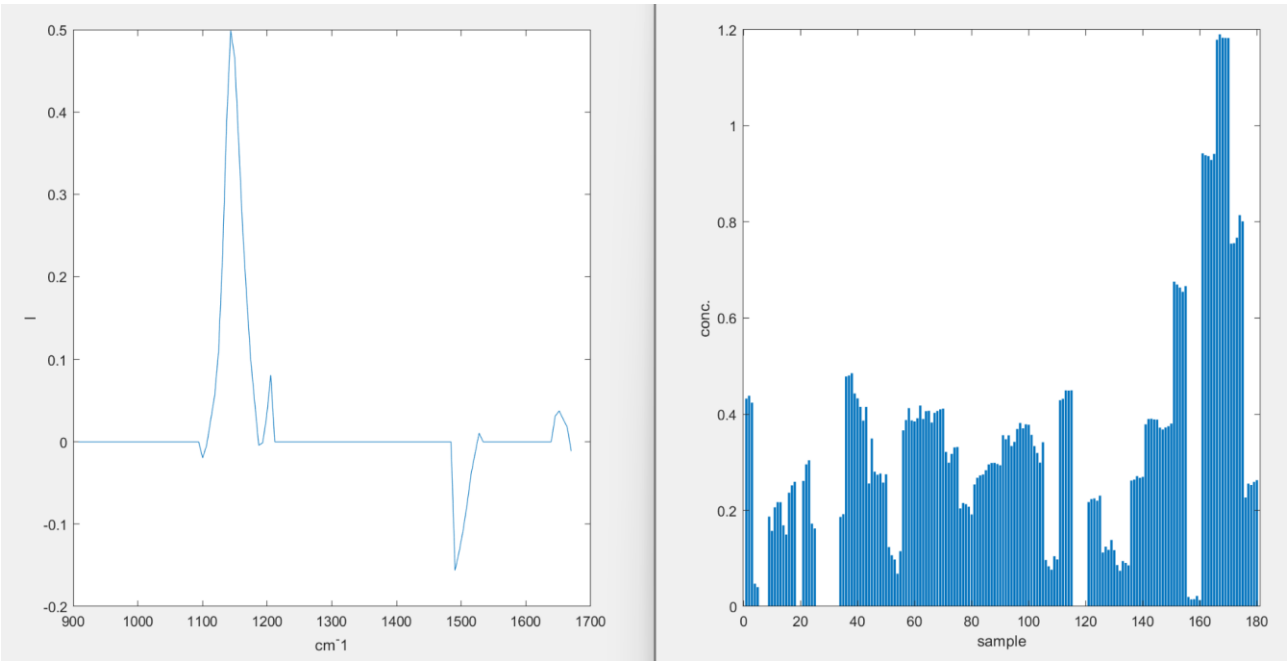



Figure 37. Second resolved component by equality constraint.

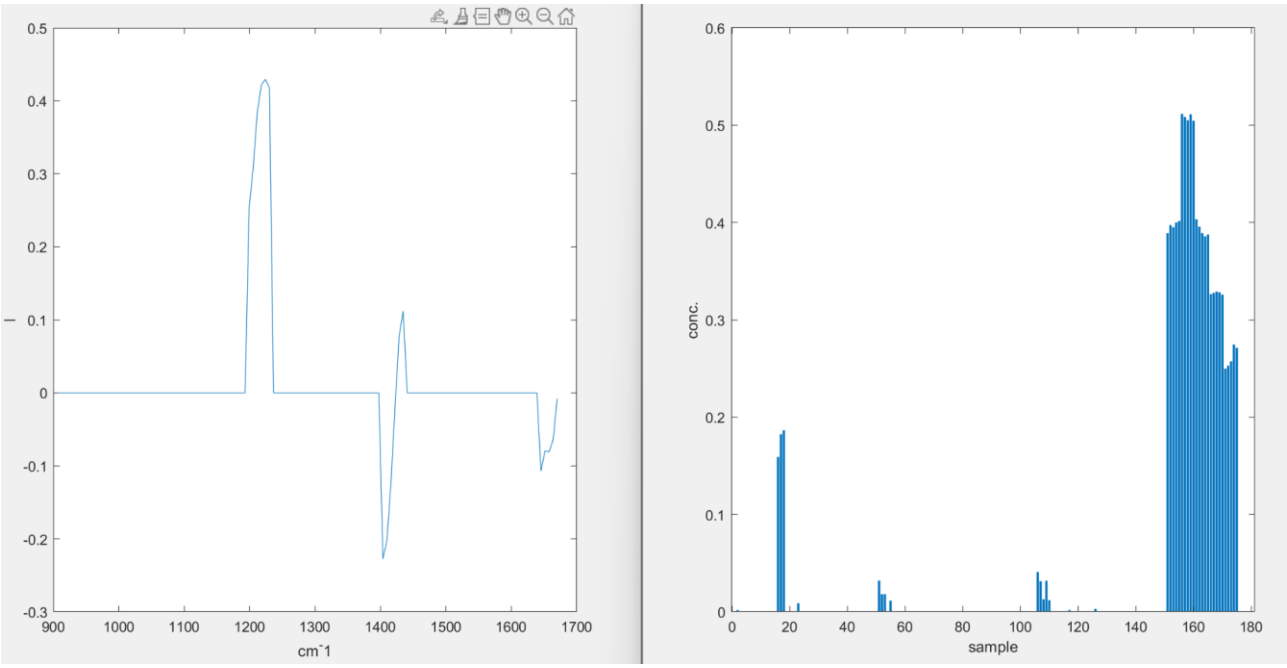



Figure 38. Third resolved component by equality constraint.

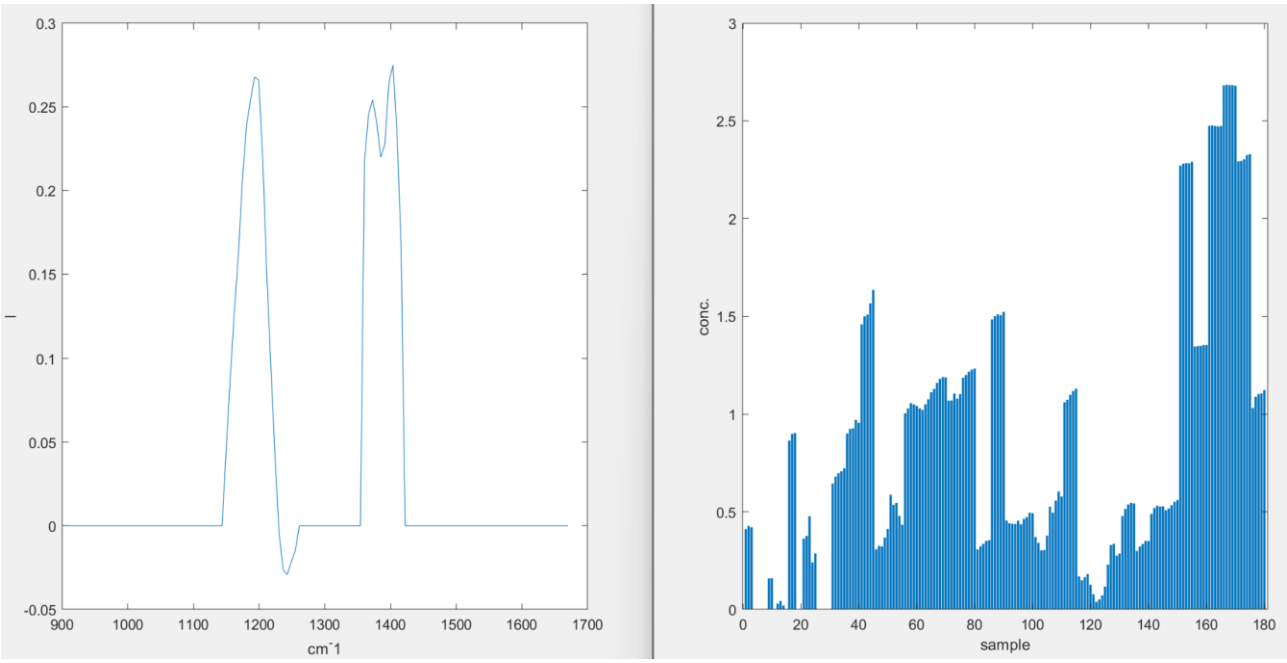



Figure 39. Fourth resolved component by equality constraint.

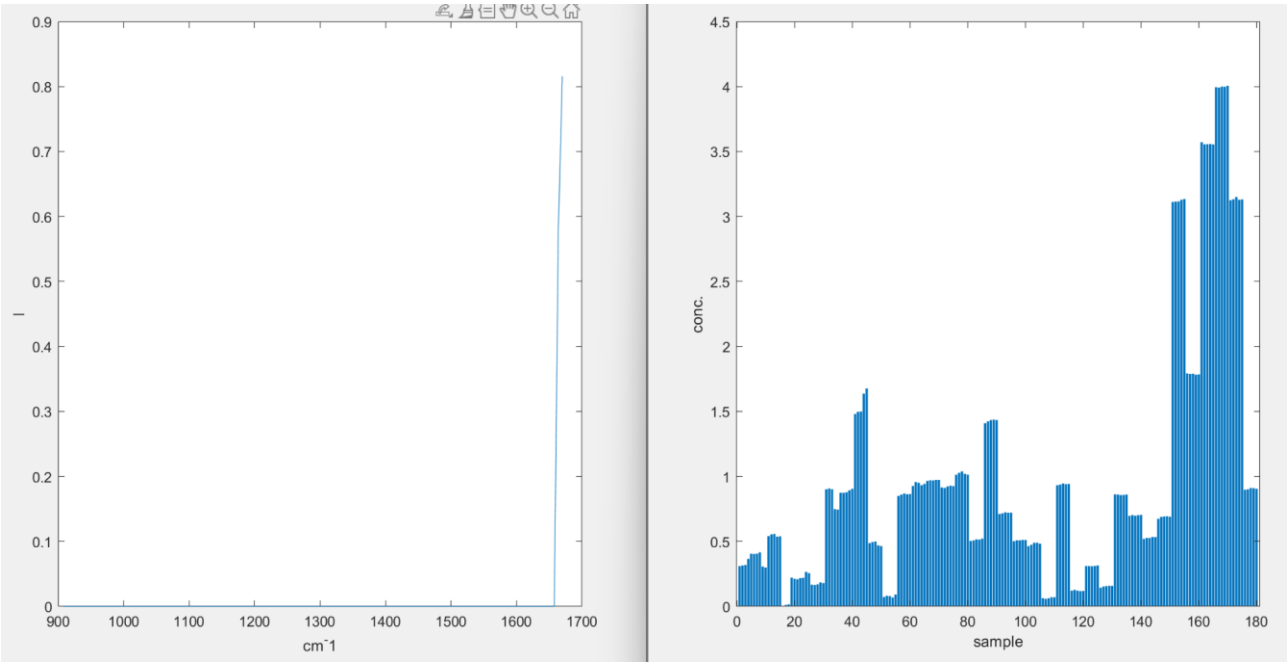



Figure 40. Fifth resolved component by equality constraint.

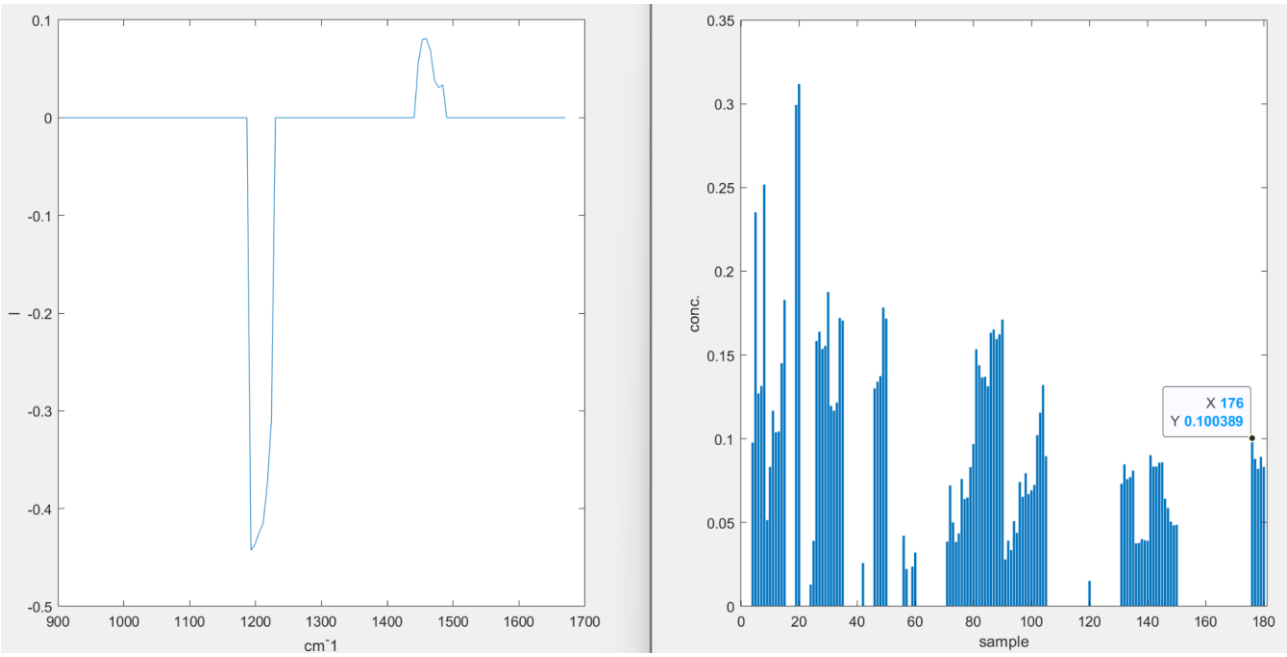



Figure 41. Sixth resolved component by equality constraint.

The datasets acquired using the portable Raman spectrometer were analysed following the same chemometric strategy adopted for the MicroNIR measurements in order to directly compare the analytical performance of the two portable spectroscopic platforms under identical experimental conditions. Two complementary constrained MCR-ALS approaches were investigated. The first consisted of a multiset strategy, in which prior information concerning the occurrence of the investigated compounds within the different datasets was incorporated into the optimisation procedure, whereas the second exploited equality constraints by introducing spectral information related to the characteristic Raman bands of

the investigated pharmaceutical standards. The comparison between the two approaches revealed a clear difference in their capability to resolve the spectral contributions of the analysed samples. Although the multiset strategy converged correctly and generated stable solutions, the resolved spectral profiles (Figures 42, 43, 44, 45, 46) remained largely dominated by the contribution of the HLB sorbent, preventing the unambiguous identification of the individual pharmaceutical compounds. Conversely, the application of equality constraints substantially improved the quality of the resolved spectral profiles. By guiding the optimisation towards characteristic Raman regions of the reference standards, the algorithm successfully identified spectral components attributable to the investigated pharmaceuticals, demonstrating that the incorporation of selective spectral information represents an effective strategy for resolving weak spectral contributions embedded within a dominant matrix. This behaviour is consistent with the intrinsic characteristics of Raman spectroscopy. Unlike near-infrared spectroscopy, where spectral information originates from broad overtone and combination bands, Raman spectra are characterised by sharp and chemically selective vibrational bands. Consequently, even when the analytes are present at relatively low concentrations, their characteristic Raman features can still provide sufficient chemical specificity to facilitate the resolution of the corresponding spectral profiles, provided that suitable constraints are incorporated into the optimisation procedure. Despite the satisfactory reconstruction of the resolved spectral profiles, both constrained approaches exhibited the same limitation observed for the MicroNIR datasets with respect to the concentration profiles. The simulated SPE samples were prepared according to a predefined Design of Experiments, and therefore the resolved Copt profiles were expected to reproduce the concentration gradients associated with each pharmaceutical standard. However, the obtained concentration profiles (Figures 47, 48, 49, 50, 51, 52) did not follow the expected experimental trends and appeared substantially unrelated to the nominal concentration levels of the investigated analytes. The discrepancy between the resolved concentration profiles and the experimental design suggests that, although Raman spectroscopy provides greater molecular specificity than near-infrared spectroscopy, the quantitative resolution of low-concentration analytes dispersed within a highly predominant polymeric matrix remains particularly challenging. Under these conditions, the spectral variance associated with the pharmaceutical compounds represents only a limited fraction of the total variance of the dataset, while the contribution of the HLB sorbent dominates the optimisation process. Consequently, MCR-ALS successfully retrieves chemically meaningful spectral signatures but encounters greater difficulty in reconstructing concentration profiles that accurately reflect the known composition of the samples. From a methodological perspective, the obtained results suggest several possible directions for future research. Since the Raman spectrum of the Oasis HLB sorbent is known through the blank measurements acquired during the experimental campaign, a promising strategy would consist of removing or attenuating the spectral contribution of the sorbent before applying the chemometric analysis. A blank-based spectral correction, performed either through spectral subtraction or by incorporating the blank spectrum as a fixed contribution during the optimisation procedure, could reduce the

dominance of the polymeric matrix and increase the relative contribution of the pharmaceutical standards. Such an approach is particularly relevant because several characteristic Raman bands of the investigated compounds partially overlap with the most intense vibrational bands of the HLB sorbent, thereby limiting the capability of the algorithm to resolve the individual analytes. An additional and particularly promising perspective involves exploiting the complementarity between laboratory and portable Raman instrumentation. The high-quality spectra acquired using the Horiba Raman microscope could be employed as reference spectral profiles or as initial estimates during the optimisation of the portable Raman datasets. In this way, the superior spectral resolution achievable with the benchtop instrument would provide chemically reliable reference information capable of guiding the decomposition of spectra acquired under field conditions. Such a knowledge-transfer strategy could substantially improve the robustness of the resolved spectral profiles while preserving the operational advantages of portable Raman spectroscopy, including rapid analysis, minimal sample preparation and potential in situ applications. The integration of benchtop and portable Raman platforms within a common chemometric framework therefore represents a promising research direction for the development of sustainable analytical methodologies capable of combining laboratory-grade chemical specificity with the flexibility required for routine environmental monitoring. Overall, the results obtained demonstrate that portable Raman spectroscopy, when combined with appropriately constrained MCR-ALS models, represents a valuable qualitative screening technique for the investigation of emerging contaminants retained on polymeric sorbents. Although further methodological developments are required to improve the quantitative reliability of the resolved concentration profiles, the high chemical specificity of Raman spectroscopy and the possibility of integrating prior spectral information acquired from laboratory instrumentation highlight the considerable potential of this approach for future environmental applications.

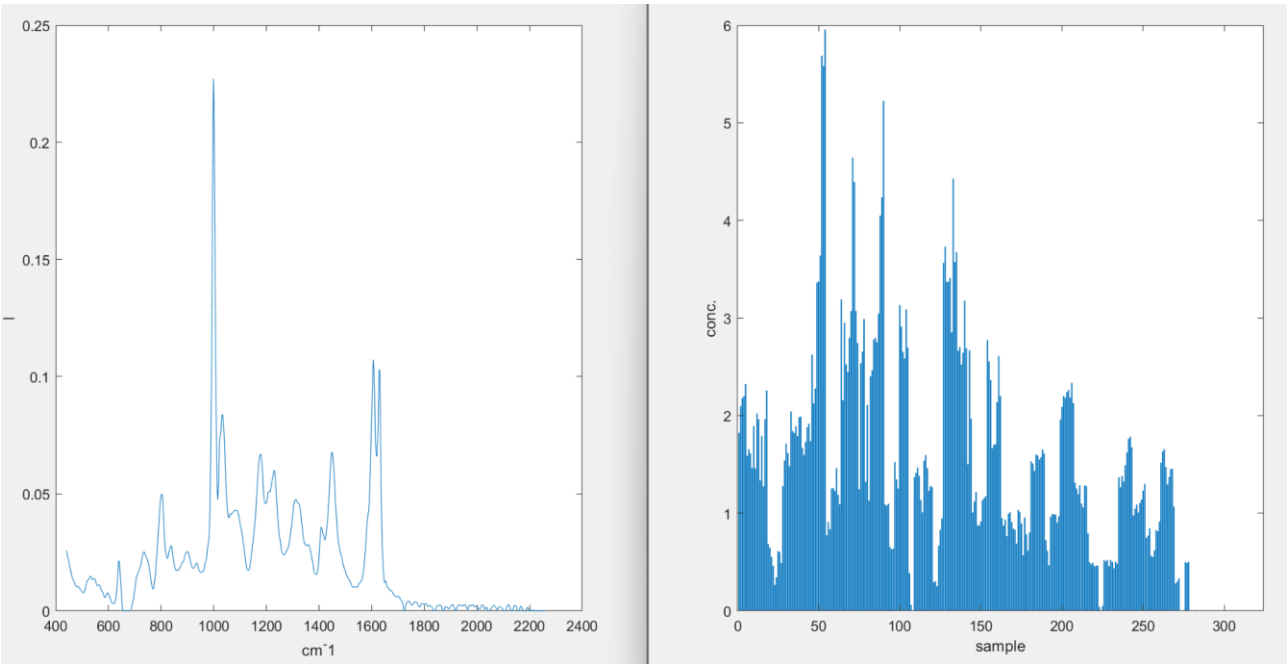


Figure 42. First resolved component by multiset (HLB resine).

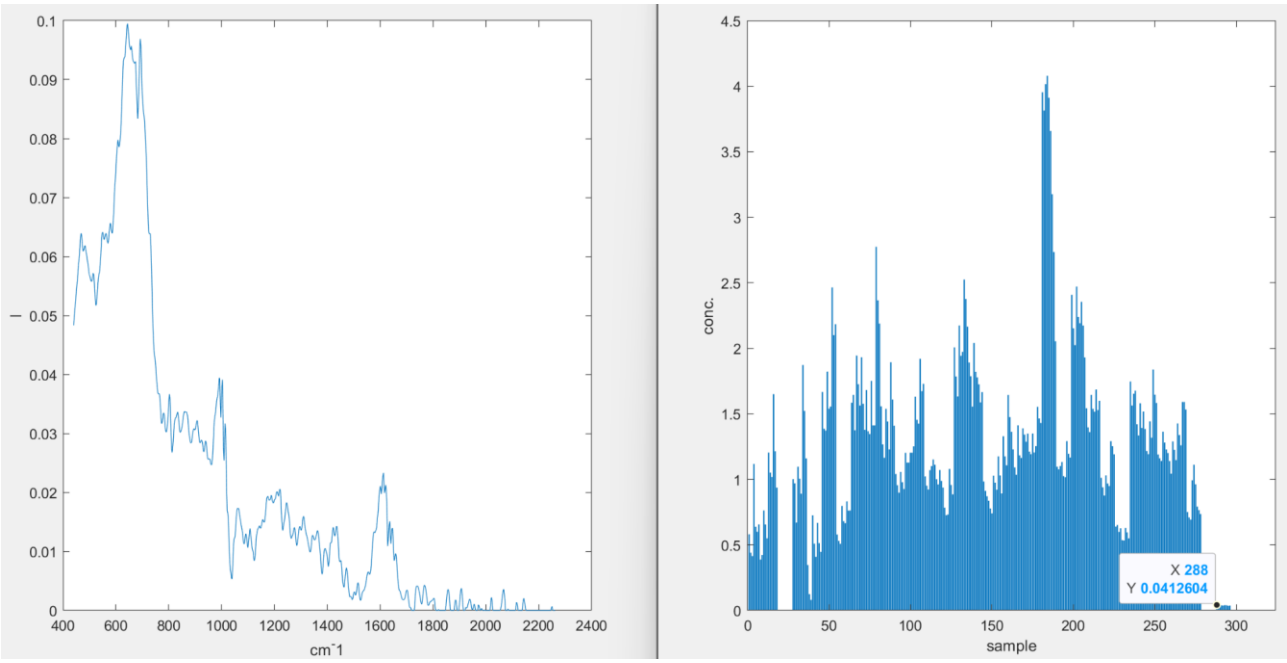



Figure 43. Second resolved component by multiset (Caffeine).

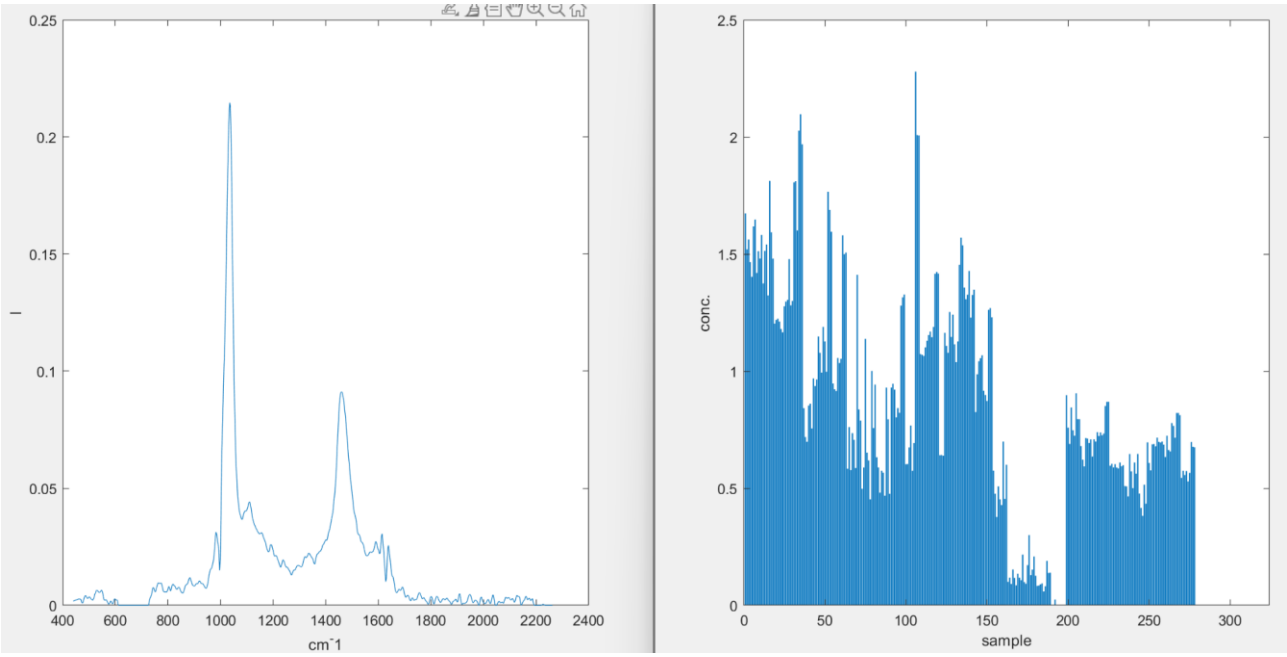



Figure 44. Third resolved component by multiset.

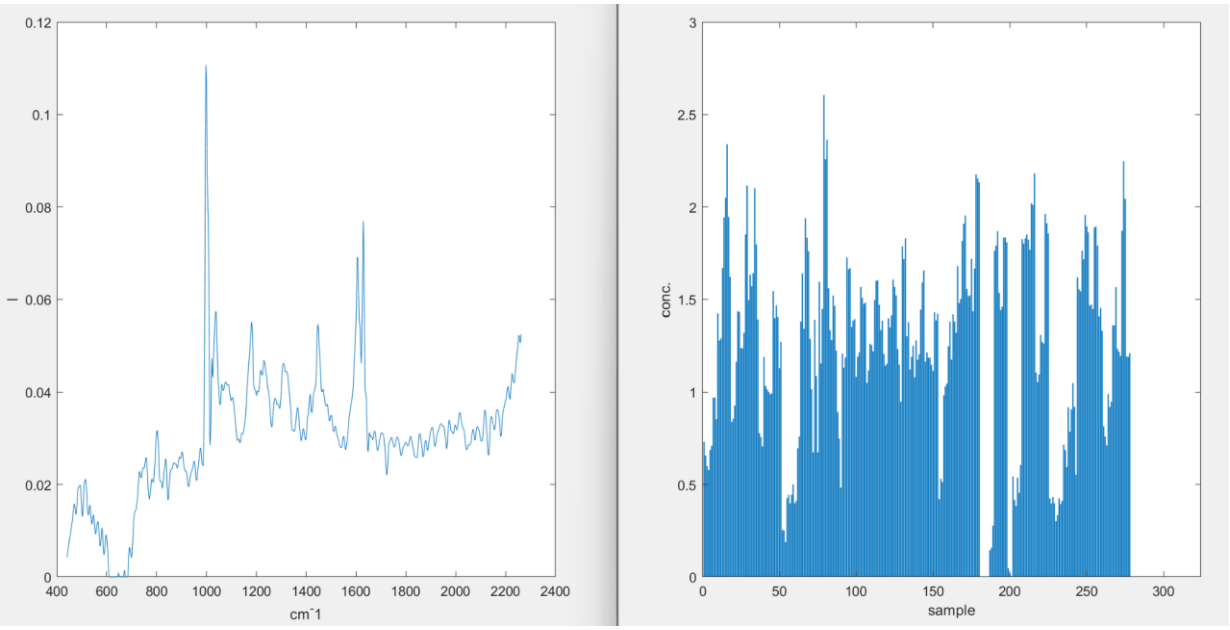



Figure 45. Fourth resolved component by multiset.

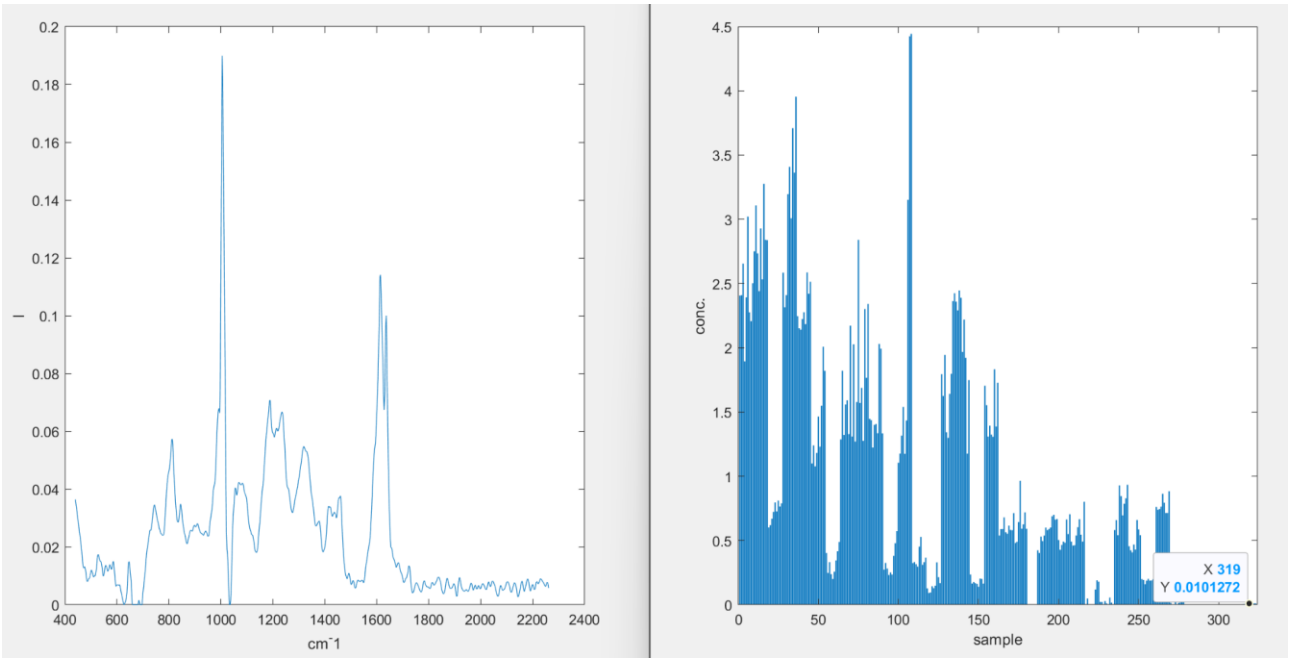



Figure 46. Fifth resolved component by multiset (Atenolol).

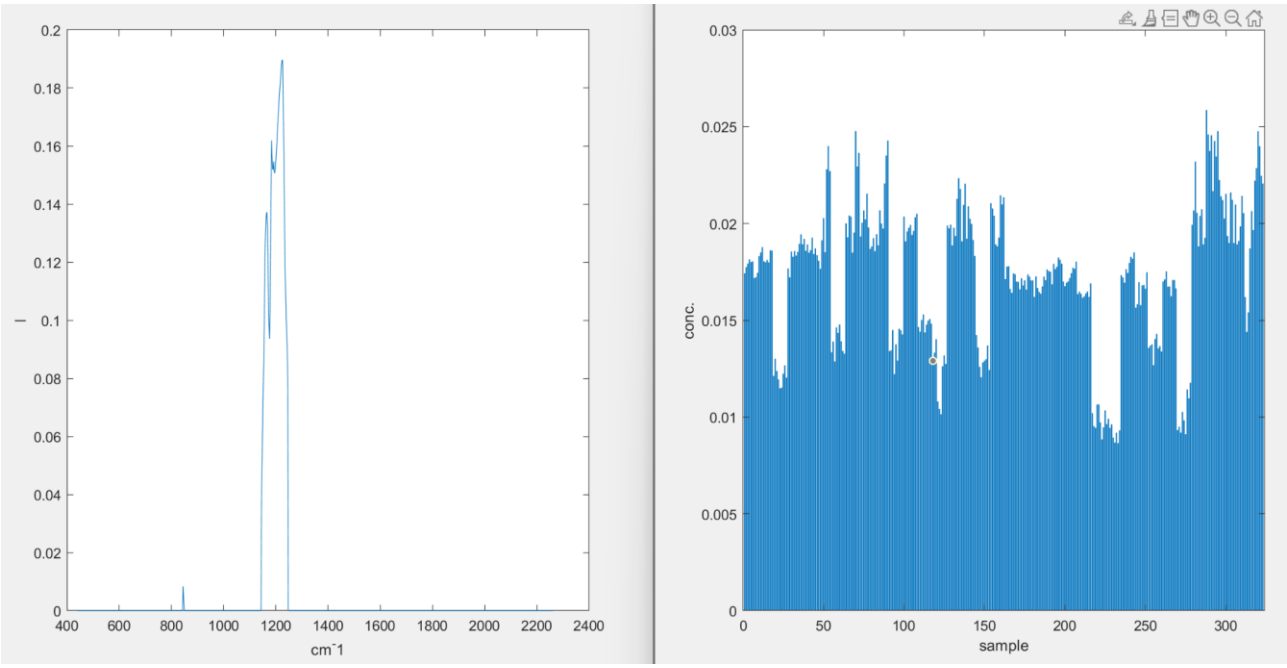



Figure 47. First resolved component by equality constraint.

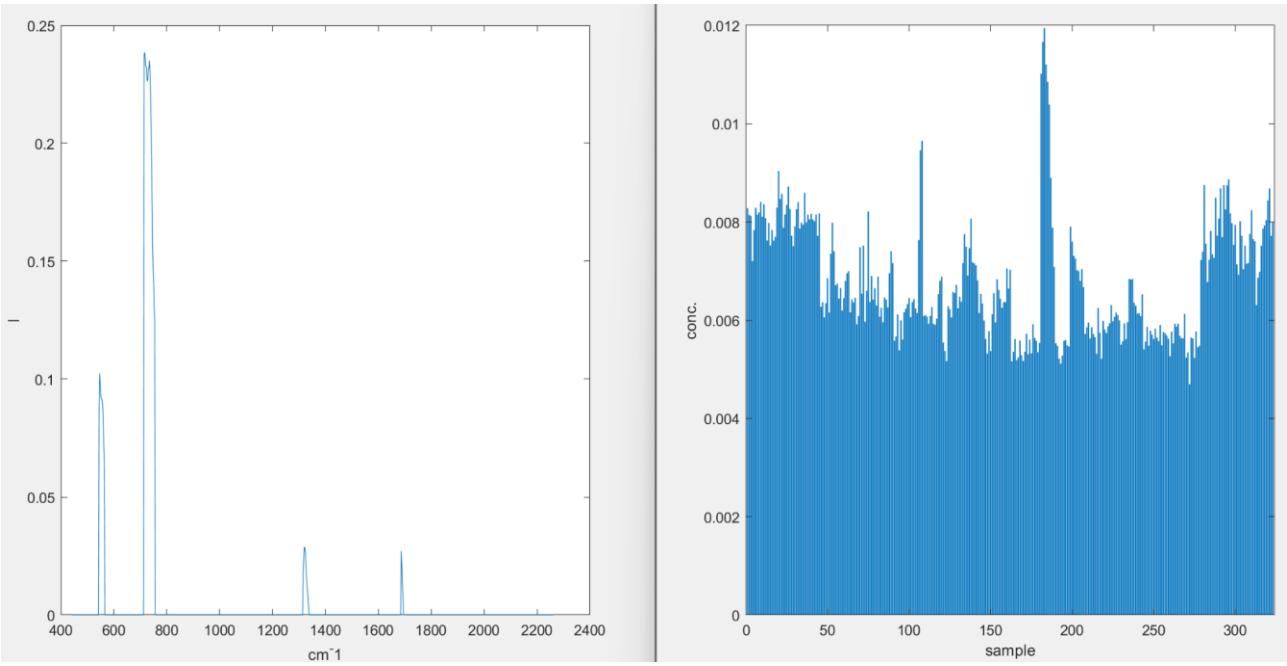



Figure 48. Second resolved component by equality constraint.

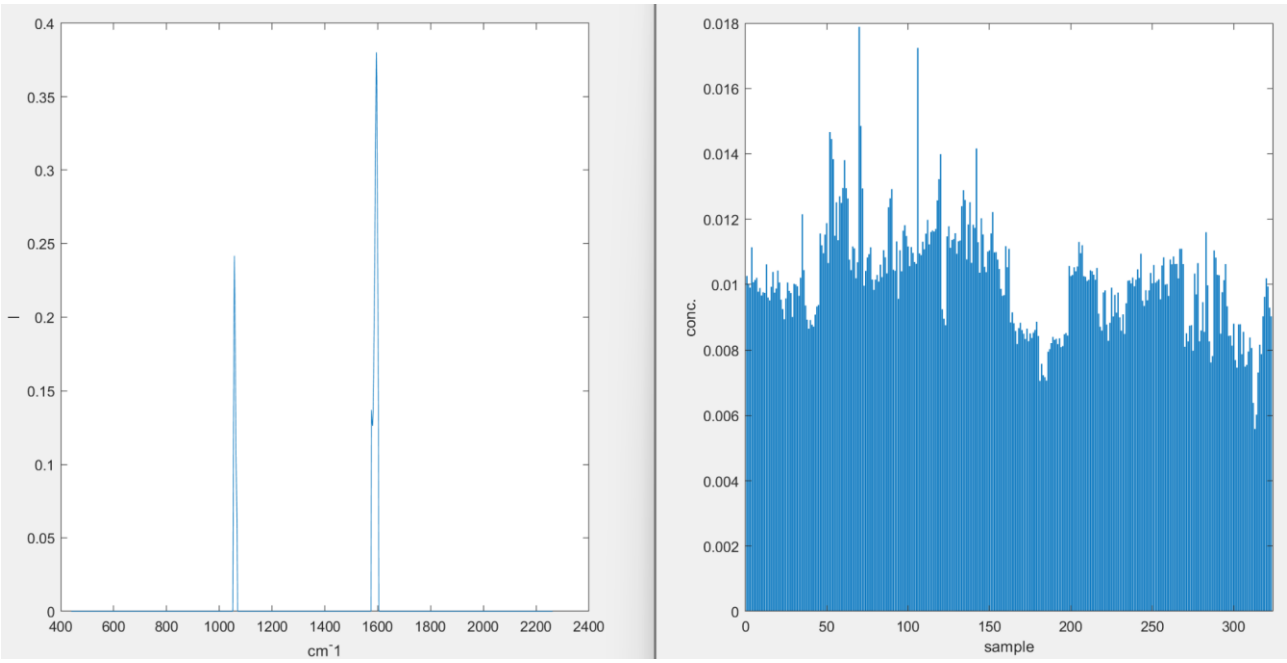



Figure 49. Third resolved component by equality constraint.

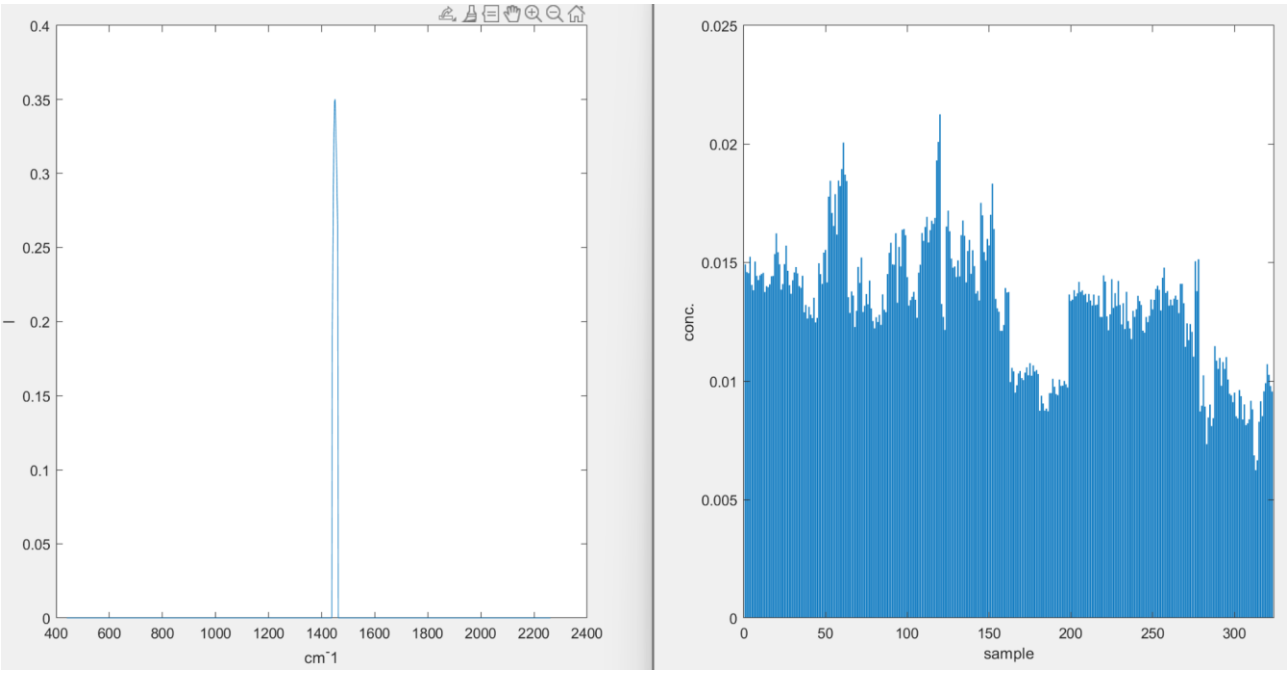



Figure 50. Fourth resolved component by equality constraint.

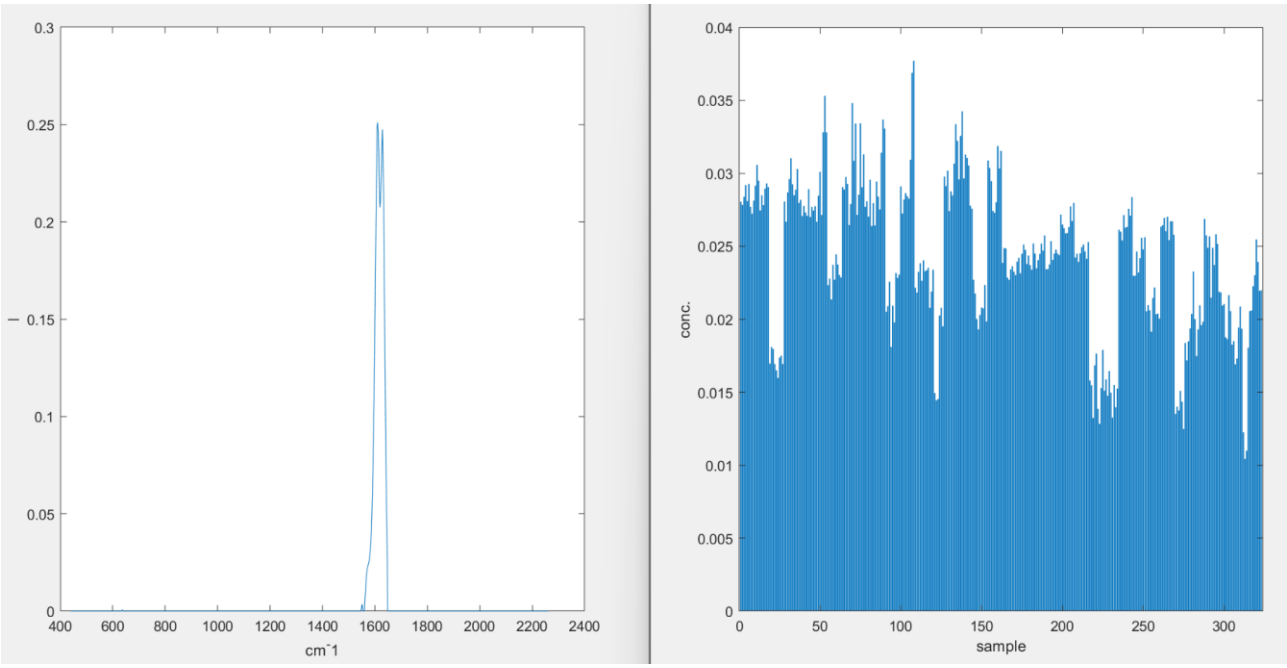



Figure 51. Fifth resolved component by equality constraint.

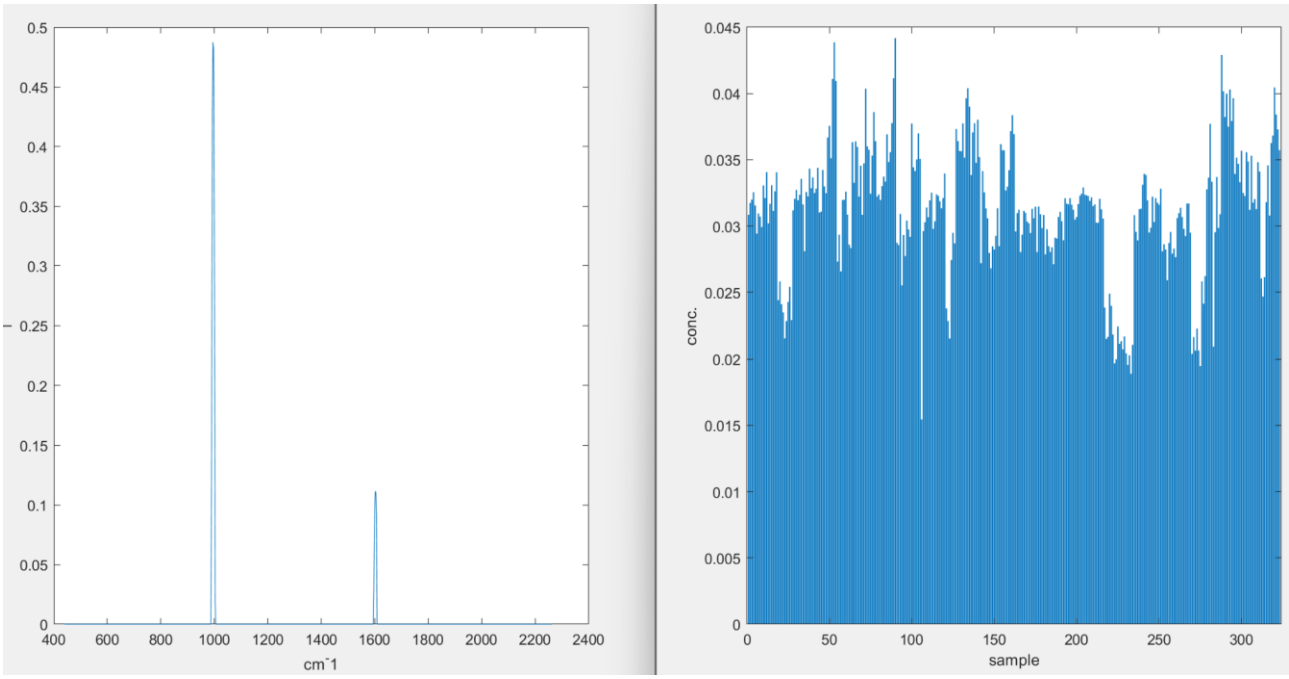



Figure 52. Sixth resolved component by equality constraint.

### *4.2. Results and Discussion of Bilbao*

The analysis of the samples collected from the wastewater treatment plants in the Basque Country allowed an evaluation of both the occurrence and the behaviour of microplastics (MPs) along the different stages of the treatment process. At the same time, it provided useful feedback on the effectiveness and limitations of the analytical workflow adopted in this study. In an initial phase, the samples from the Galindo wastewater treatment plant were mainly used for method development. These samples were particularly useful to optimize the sample

preparation protocol and to better understand how the matrix could affect the spectroscopic response. Raman analyses carried out on selected particles showed a clear predominance of polyethylene terephthalate (PET), mostly in the form of thin fibres. However, the presence of residual organic matter on the filters often interfered with the spectral quality, sometimes making the identification less straightforward. This aspect highlighted the need to improve the digestion step, leading to the implementation of a more effective pre-treatment procedure. The optimised protocol was then applied to the samples collected from the Crispijana wastewater treatment plant, which represent the main dataset of this work. The combined use of optical microscopy and Raman spectroscopy confirmed that most of the detected microplastics were fibrous. PET was by far the most frequently identified polymer, while polyethylene (PE) and, less commonly, polystyrene (PS) were also detected. In several cases, fibres with different colours (transparent, blue, red or green) were observed. This variability is likely related to the presence of dyes or additives, which can influence the Raman signal and, in some cases, complicate the spectral interpretation. Looking at the distribution of MPs along the treatment line, a clear trend can be observed. The highest number of particles was generally found at the inlet of the plant (ERAZ), where 19 MPs were identified. A comparable or slightly lower number was observed in the primary treatment stage (DEKA), with 14 particles. Moving towards the biological treatment (BIOL_sar and BIOL_irt), the number of detected MPs decreased to 5 and 1, respectively. Finally, the lowest values were observed in the tertiary treatment effluent (TERT), although a residual presence of MPs was still detected (3 particles). Even if the dataset is not intended for a strict quantitative balance, this progressive decrease suggests that the treatment processes are able to remove a significant fraction of microplastics. At the same time, the fact that MPs are still present in the final effluent indicates that conventional wastewater treatments are not fully effective in eliminating these contaminants. In particular, the persistence of fibrous particles could be related to their morphology, which may favour their passage through the different treatment steps. From a methodological point of view, the approach based on optical pre-selection followed by point-by-point Raman analysis proved to be reliable for polymer identification, but also quite time-consuming. For these reasons, complementary approaches such as Raman mapping and near-infrared hyperspectral imaging (HSI-NIR) were explored in order to analyse larger areas of the filters. While these techniques showed potential, especially for screening purposes, some limitations were observed. In particular, the lower spatial resolution of HSI made it less effective for the analyses of the smallest particles compared to Raman spectroscopy.

Overall, the results obtained in this study confirm that Raman spectroscopy remains one of the most suitable techniques for the identification of microplastics in complex environmental samples. At the same time, they highlight the importance of proper sample preparation and the advantage of combining different analytical approaches to obtain a more complete picture of the system under investigation.

The experimental work carried out during the internship at the Universidad del País Vasco / Euskal Herriko Unibertsitatea (UPV/EHU) was part of a broader research project focused on the identification and characterisation of microplastics in wastewater treatment plants in the Basque Country. For this reason, part of the results obtained during this activity, together with their detailed discussion, were included in a scientific article developed in collaboration with the research group and a PhD candidate involved in the same project. Consequently, the present thesis reports and discusses the most relevant aspects of the experimental work, while some analytical details and extended datasets are addressed within the associated publication[65].

[65] FGonzález e IBeA, *CARACTERIZACIÓN QUÍMICA Y MORFOLÓGICA DE MICROPLÁSTICOS EN LA EDAR DE CRISPIJANA, VITORIA-GASTEIZ (KRISPI-PLASTIC)*.

## 5. Conclusion and Perspectives

The present thesis focused on the development and evaluation of sustainable spectroscopic and chemometric strategies for the investigation of emerging contaminants retained on passive sampling devices and simulated solid-phase extraction matrices. The increasing occurrence of contaminants of emerging concern (CECs) in aquatic environments represents a critical analytical challenge, as conventional chromatographic techniques, although providing excellent sensitivity and selectivity, often require complex sample preparation procedures, extensive solvent consumption and laboratory-based instrumentation. In this context, the development of complementary analytical approaches capable of providing rapid, cost-effective and environmentally sustainable screening methodologies represents an important research direction. The experimental work carried out in this thesis demonstrated the potential of combining Raman spectroscopy, hyperspectral imaging and near-infrared spectroscopy with advanced chemometric approaches for the characterisation of complex environmental matrices. Particular emphasis was placed on the application of Multivariate Curve Resolution–Alternating Least Squares (MCR-ALS) as a powerful tool for extracting chemically meaningful information from spectroscopic datasets characterised by spectral overlap, matrix effects and low analyte concentrations. The Raman mapping analyses performed using the Horiba Raman microscope highlighted the capability of high-resolution Raman spectroscopy combined with chemometric modelling to investigate the spatial distribution of spectral contributions within POCIS samples. The implementation of a PCA-based data geometry approach allowed the identification of representative spectral profiles and the subsequent reconstruction of chemical distribution maps through MCR-ALS. The results demonstrated that the dominant spectral contribution corresponded to the sorbent matrix, while additional components were associated with regions characterised by different spectral signatures, potentially related to retained environmental contaminants. However, the analysis also highlighted the intrinsic complexity of untargeted identification within environmental matrices, where multiple compounds and matrix contributions can generate highly overlapping spectral profiles. The application of hyperspectral imaging to POCIS samples provided further insight into the advantages and limitations of spatially resolved spectroscopic approaches. Although HSI enabled the rapid acquisition of a very large number of spatially distributed spectra and successfully highlighted regions characterised by different spectral responses, the limited spectral selectivity of the technique represented a critical factor for the untargeted identification of individual contaminants. The high spatial resolution provided by the system was not sufficient to overcome the limited chemical resolution of the acquired spectra, resulting in redundant spectral contributions and preventing a clear separation between the sorbent matrix and the retained contaminants. These observations underline the importance of balancing spatial and spectral resolution when selecting hyperspectral platforms for complex environmental applications. The analysis of simulated SPE samples using portable Raman and MicroNIR spectroscopy demonstrated the potential of compact spectroscopic devices as rapid screening tools for contaminant investigation. The

application of constrained MCR-ALS approaches allowed the recovery of chemically meaningful spectral profiles, particularly when prior information from reference standards was incorporated into the model. The equality-constrained strategy proved particularly effective in improving the resolution of individual contaminant signatures by guiding the optimisation towards characteristic spectral regions. However, the reconstruction of concentration profiles remained challenging due to the dominance of the sorbent matrix and the limited contribution of low-concentration analytes to the total spectral variance. Overall, the results obtained throughout this thesis highlight that the major challenge associated with spectroscopic monitoring of emerging contaminants is not only related to detection capability, but also to the development of robust chemometric strategies able to discriminate against weak analyte signals within complex environmental backgrounds. The combination of spectroscopic information, prior chemical knowledge and advanced mathematical constraints therefore represents a key aspect for improving the reliability and applicability of these methodologies.
Several future developments can be proposed based on the outcomes obtained in this work. Regarding Raman-based analyses, a promising direction would involve exploiting the information provided by the laboratory Raman microscope to improve the interpretation of portable Raman measurements. The high spectral quality obtained with the Horiba system could be used as a reference source for portable Raman datasets through knowledge-transfer approaches, allowing the spectral resolution of laboratory instrumentation to guide the decomposition of lower-resolution field measurements. Furthermore, the systematic subtraction or modelling of the sorbent contribution obtained from blank samples could reduce matrix-related interference and enhance the detectability of retained contaminants.
For hyperspectral imaging applications, future improvements should focus on optimising the compromise between spatial and spectral resolution. The use of alternative optical configurations, including higher magnification objectives or reduced field-of-view acquisitions, could provide increased spectral information per pixel while maintaining the advantages of rapid imaging. Moreover, the integration of hyperspectral data with complementary spectroscopic techniques through data-fusion strategies could overcome the limitations observed for single-platform approaches. From a chemometric perspective, future research should focus on the development of more advanced constrained models capable of integrating multiple sources of prior information. The combination of multiset strategies, equality constraints and concentration-based constraints derived from experimental designs could represent an effective approach to improve both qualitative identification and quantitative prediction. In addition, the integration of MCR-ALS with supervised machine-learning approaches and predictive modelling could further enhance the capability of portable spectroscopic systems for environmental monitoring. The perspectives identified in this thesis are directly aligned with the research activities proposed within the PhD project entitled “Development of sustainable sample pretreatment and detection methods for emerging contaminants and microplastics in water systems”, under the supervision of Prof. Costanza Scopetani and Prof. Marina Cocchi at the University of Modena and Reggio Emilia, in collaboration with Prof. Jose Manuel Amigo Rubio at the University of the Basque Country

(UPV/EHU). The doctoral research will represent a natural continuation of the present work by extending the developed spectroscopic and chemometric strategies towards broader environmental applications, including the simultaneous investigation of emerging contaminants and microplastic-related pollutants. Future activities will therefore aim at integrating sustainable sample preparation procedures, portable spectroscopic platforms, hyperspectral approaches and advanced chemometric modelling into comprehensive analytical workflows capable of supporting faster, greener and more accessible environmental monitoring. The combination of analytical innovation, data-driven methodologies and sustainable approaches represents a promising pathway towards the development of next-generation strategies for the assessment and management of aquatic contamination.

## 6. References

**APPENDIX A**

*Routine MatLab:*

```
%% essential information prior to MCR

%% prepare your files

dataset=spectra_norm;%(substitute with the name of your data, data should be in matlab workspace)

if isdataset (dataset)

  matrix=dataset.data;

else

  matrix=dataset;

end

% if you have original map dimension in pixel put them here

sizet=[8 33 1147];

%if you want to exclude some spectra create pixin index

flag1=input('do you want to exclude some spectra?','s') % type 'Y' (yes) or 'N' (not)

if flag1=='Y'

  pixin=input('give vector with index of spectra to include');

  % esempio se digiti: [1:4830 4832:8281]   il 4831 spettro è stato escluso

  % ovviamente il numero di spettri va aggiornato con quello nel tuo dataset

else

  pixin=[1:size(matrix,1)];

end

%if you want to exclude some variables create includevar index

flag2=input('do you want to exclude some wavelengths?','s'); % type 'Y' (yes) or 'N' (not)
```

```
if flag2=='Y'
  includevar=input('give vector with index of variables to include');
  % put the reange of wavelengths you want to consider
else
  includevar=[1:size(matrix,2)];
end
matrix=matrix(pixin,includevar);
%% run Data Geometry
n=14; % number of total PCs to consider , can be changed
% run PCA and normalize
scores_n=[];
ormatrix=matrix;
[u,s,v]=svds(matrix,n);
t=u*s;
for i=1:n
  scores_n(:,i)=t(:,i)./t(:,1);
end
figure;
pp=input('Give components to plot as [nx xy]'); % se vuoi PC2 vs PC3 scrivi [2 3]
plot(scores_n(:,pp(1)),scores_n(:,pp(2)),'.k','MarkerSize',5);
xlabel(strcat('t',int2str(pp(1))));
ylabel(strcat('t',int2str(pp(2))));
set(gca,'fontsize', 18), grid on,
axis tight

% get the convexhull points
```

```
chulPC=[2 3] % indicate the number of components you want to consider in the convexhull
% default is the first 2 , which are 2 and 3 because of normalization
index_convh = convhulln(scores_n(:,chulPC));
selected_spectra.data=matrix(index_convh(:,1),:);
selected_spectra.PCs=chulPC;
selected_spectra.index=index_convh;
selected_spectra.normscores=scores_n;
save selected_spectra selected_spectra %vengono salvati i risultati
% plot on the normalized scores plot the convexhul point in red
hold on;
plot(scores_n(index_convh,pp(1)),scores_n(index_convh,pp(2)),'.r','MarkerSize',5);
hgsave('normscore_plot')
%% MCR dal modello salvato
S=sopt_sel;
spettri_pret= spectra_norm;  % ci metti il nome del file iniziale (matrice degli spettri preprocessati senza mean centering)
C=spettri_pret*S'; % non so se S' o S , dipende dalle dimensioni di S , se  è lunghezze d'onda x numero componenti il ' non serve, se è il contrario ci vuole.
%%
Xcoord=coords.raw(:,1);
Ycoord=coords.raw(:,2);
xvals=coords.x ;
yvals=coords.y ;
[~, ix] = ismember(Xcoord, xvals);
[~, iy] = ismember(Ycoord, yvals);
%% allocate cube and fill
```

```
Ccube = zeros(33, 8, 6);  % nMCR è il numero di componenti MCR che hai selezionato (numero colonne di copt)

for p = 1:264

  Ccube(iy(p), ix(p), :) = C(p, :);

end

figure;

for i=1:6

subplot(1,6,i);imagesc(Cube(:,:,i));

end
```