Environmental Science Advances

This article can be cited before page numbers have been issued, to do this please use: R. Fernandes, P. Miclea, M. Putzu, V. Morgado, J. Almeida, M. São Pedro, C. Drago, A. M. M. GIOVANNOZZI, C. Palma, R. M. P. A. Quendera, O. Pellegrino, R. J. N. Bettencourt da Silva and A. M. Rossi, *Environ. Sci.: Adv.*, 2026, DOI: 10.1039/D6VA00098C.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/esadvances

Environmental Significance Statement

View Article Online
DOI: 10.1039/D6VA00098C

This work delivers a transparent, low-computational-resource methodology for the inter-instrument identification of PET microplastics by μ-FTIR, combining statistical evaluation of spectral-matching algorithms with a robust signal-to-noise metric tailored to FTIR baseline variability. In contrast to existing approaches, the method was not only validated across multiple laboratories and instrumental conditions but was also subjected to an independent performance assessment using spectra not included in the model development, experimentally confirming the predicted true-positive and false-positive rates. This dual validation strategy establishes a practical and harmonised framework for reliable FTIR-based microplastic identification, offering accuracy comparable to complex chemometric tools while remaining accessible for routine environmental analysis.

View Article Online
DOI: 10.1039/D6VA00098C

Inter-equipment ongoing validation of microplastics identification by micro-FTIR using optimal resources

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Rafaela Fernandes[1], Paul-Tiberiu Miclea[2], Andrea M. Rossi [3], Andrea M. Giovannozzi[3], Mara Putzu[3,4], Vanessa Morgado[5,6], Carla Palma[6], José Almeida[6], Matilde São Pedro[6], Claudia Drago[7], Raquel Quendera[8], Olivier Pellegrino[8], Ricardo J.N. Bettencourt da Silva[1*]

FTIR is widely used for identifying microplastics due to its performance and affordable equipment. However, there is a need for a universal method for the automatic identification of microplastics by FTIR, robust to different spectra collection conditions, and capable of adapting to the diversity of particle spectra. This research presents a method for developing a reliable and simple algorithm for inter-instrumental microplastics identification using micro-FTIR. Spectra collected from different equipment and settings across various laboratories, from microparticles previously confirmed as PET (Positive cases) and non-PET (Negative cases), were compared with a PET reference using multiple algorithms (Match Methods) based on six weighted and unweighted correlation coefficients. After excluding low signal-to-noise spectra with a newly developed universal algorithm, the 5th percentile of the Match values for Positive cases, estimated using a robust bootstrap method, was used as a minimum Match (*P*5»P) to ensure a 95% True Positive rate (*TP*). Assuming a normal distribution of the Match values for Negative cases, the False Positive rate (*FP*) was calculated. The best method identified uses an unweighted correlation of differentiated signals, resulting in a *P*5»P=0.4140 and *FP*=0.0005% for spectra with a minimum signal-to-noise ratio of three. This identification performance, estimated from 117 spectra, is statistically equivalent to that observed from the identification of 405 additional particles using the validated method, at a 99% confidence level set by the Clopper–Pearson method, thereby extending the method's applicability to 522 particles. The developed method for PET identification by FTIR is robust across various acquisition conditions and instruments, was implemented in a user-friendly MS Excel spreadsheet, and can be easily tested with more particles and different instrumental settings. The validity of the identification algorithm and the defined criteria cannot be extrapolated to a different reality with respect to spectral collection conditions and the diversity of negative cases without a prior performance assessment.

1 *Centro de Química Estrutural, Institute of Molecular Sciences, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal*
2 *Fraunhofer Center for Silicon Photovoltaics (CSP), Germany,*
3 *Istituto Nazionale di Ricerca Metrologica (INRiM), 10135 Turin, Italy*
4 *University of Turin (UniTO), Department of Chemistry, 10125 Turin, Italy*
5 *Laboratório Nacional de Energia e Geologia, I.P., Lisboa, Portugal*
6 *Instituto Hidrográfico, Portugal,*
7 *Umweltbundesamt, German Environment Agency, Germany,*
8*Instituto Português da Qualidade, Portugal*
* *rjsilva@ciencias.ulisboa.pt*

Supplementary
Information available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

Introduction

Plastics are synthetic or semi-synthetic polymers employed in a range of applications, such as packaging, textiles, medical, electronic, and construction, primarily due to their malleability, chemical and thermal stability, lightweight nature, and low cost. [1] Due to the low price and societal weaknesses, plastics are often discarded improperly, contributing significantly to public waste. Plastic waste is extremely diverse in composition, colour, size, and shape. The range of sizes results from their initial dimensions and subsequent fragmentation through multiple processes. Plastic fragmentation mainly results from natural processes like mechanical action, biodegradation, and ultraviolet degradation, which mostly occur during the use or after the disposal of these materials. When the particles have dimensions between 1 µm and 1 mm, they are classified as microplastics. [2] Microplastics are widespread worldwide and have contaminated both soil, water, and air. [3,4] As studies demonstrate its harmful effects on human and animal health, the concern over microplastic pollution increases, and so does the need to recognise their contamination patterns. [4] It is not only essential to thoroughly understand the source of microplastics to the environment, but also their migration and distribution. To accomplish this goal, it is essential to collect and process information obtained from the efforts of researchers from all over the world, which demands reliable, and equivalent or interchangeable protocols for sampling, sample treatment, and microparticles identification and quantification.

Typically, polymer identification is achieved by comparing the particle spectra with a reference spectrum of the target material, using vibrational spectroscopies such as Raman or Fourier-transform infrared (FTIR). Both are non-destructive analysis methods, allowing re-analysis and analysis through other techniques, and can be used to characterise particles isolated from different matrices, such as those collected from various environmental compartments. These techniques, however, are not competitive but complementary: Raman spectroscopy measures variations in molecular polarizability, whereas FTIR measures variations in electric dipole moment. [5] In FTIR, information about molecules' functional groups is retrieved from the spectra region between 3600 cm^{-1} and 1250 cm^{-1}. The fingerprint region, between 1200 cm^{-1} and 600 cm^{-1}, where small structural and constitutional changes lead to significant differences in both appearance and band distribution, has bands challenging to attribute to a specific molecular movement. [5,6] Overall, the bond environment affects vibrational modes, leading to vibrational coupling and alterations in their characteristic wavenumber. Consequently, there may be uncertainty in identifying the functional groups responsible for certain bands. Nevertheless, this also creates a unique spectrum which allows for a selective compound identification. [5]

Raman is superior to FTIR spectroscopy when analysing particles smaller than 10 µm. However, FTIR is less affected by fluorescence and particle pigmentation than Raman. For laboratories unable to afford both FTIR and Raman spectroscopies due to equipment and operational costs, FTIR offers the additional benefit of more cost-effective instruments. [7]

The main challenge of FTIR and Raman spectroscopic methods is the necessity for skilled analysts to interpret spectra manually, which is time-consuming. Automatization of polymer identification of microplastics involves quantifying the similarity between particle and reference spectra through a simple algorithm, such as a correlation coefficient, or a complex chemometric tool, such as a Machine learning (ML) algorithm like k-nearest neighbours (k-NN) or support vector machines (SVM). [8] These ML algorithms are highly complex, require appropriate computational resources, and depend heavily on large libraries that can be private and/or limited to a particular spectral range. [8]

Primpke et al. (2020) [9] developed *siMPle*, an open, independent software with over 90 % of correct assignment rates for microplastic identification using FTIR spectroscopy. The software offers harmonised spectral matching across multiple instruments from four major manufacturers and various detection modes, covering common polymers found in environmental studies such as polyethylene (PE), polypropylene (PP), polyethylene terephthalate (PET), and polystyrene (PS). The match is calculated by combining the weighted Pearson's correlation coefficient for untreated data and their first and second derivatives. The user-selected weights for each coefficient must be specified for inter-study comparison, requiring subsequent thorough validation. Willans et al. (2023) [10] applied partial least squares–discriminant analysis (PLS-DA) to FTIR reflectance imaging of marine salt samples, where particles were typically 25–75 µm in size. Their method achieved a very high true negative rate (TN) (100 % for PE and PP) but a somewhat lower true positive rate (TP) (around 78–82 %), highlighting both the potential and limitations of advanced statistical models when dealing with real-world noisy spectra.

Clough et al. (2024) [11] introduced conformal prediction as a novel framework for microplastic identification. Contrasting with previous literature, this method promises to associate an uncertainty with each individual match, based on the user-chosen confidence level. To achieve robustness to real samples, an open spectra library, including both pristine polymer samples and aged environment-sourced particles, was utilised.

Regardless of the computational tool used, all automated processes require setting a threshold for the quantified spectral similarity above which there is a sufficiently high chance of correct classification. [1,12] Defining this threshold requires the statistical analysis of similarity values for particles known to be, or not, from the same polymer as the reference spectrum. A challenge in setting the threshold is the unique combination of instrumental conditions and preprocessing steps from each spectrometer software, as well as their unknown effects on the match value, which complicates comparisons between instruments.

Fernandes et al. (2025) [13] have recently developed and validated an algorithm for the inter-instrumental identification of PET microparticles by Raman spectroscopy, where reference spectra acquired from a single instrument are applied across multiple instruments and varying instrumental conditions. Identification with a minimum Person's correlation coefficient, *P5*»P, of 0.6244, a *TP* not lower than 95%, and a false positive rate, *FP*, of no greater than 4.9×10^{-7}% was achieved for spectra with a signal-to-noise ratio (S/N) greater than 10. These authors have developed an algorithm to quantify the S/N of Raman spectra regardless of the values of the scattered light, reported in arbitrary units, that depends on particle size and spectra collection parameters. This research did not discuss how to assess the need to update the identification algorithm as additional particles are studied.

This paper describes the development of a strategy for the ongoing validation of an inter-equipment method for the identification of microplastics by FTIR spectroscopy. The method takes the specificity and characteristics of micro-FTIR spectra into account and allows an easy assessment of the need for algorithm updating as more particles are studied.

The identification algorithm adjusts to the sensitivity, band resolution, baseline variations, spectral patterns, and spectral variability function of particle characteristics and spectra collection conditions of micro-FTIR. Due to the wider spectral bands and relevant baseline variations, a new universal algorithm for quantifying the S/N ratio for FTIR spectra was developed.

The algorithm, validated with a limited number of spectra used to define a minimum S/N and set identification criteria, was tested on a set of particles different from those used in the validation, extending the method's scope or determining the need to update the algorithm. This strategy supported an ongoing validation process, meaning that as more particles are studied, the scope of the method can be expanded with or without adjustment.

The predicted performance of the method from their validation and the observed performance from the identification of an independent set of spectra are compared using the Clopper–Pearson method at a 99% confidence level. This approach considers the number of spectra involved in the assessment.

The developed approach merits comprehensibility, statistical soundness, and low computational resource requirements without compromising performance, making it suitable for general use.

Experimental

Material and equipment

Spectra were collected by three spectrometers from three laboratories: Lab.1 – Fraunhofer CSP, Lab. 2 - Istituto Nazionale di Ricerca Metrologica, and Lab. 3 - Instituto Hidrográfico. Only non-PET spectra are available from Lab. 1.Table 1 lists the spectrometers and the most relevant instrumental conditions.

In this work, the identification method was developed and validated using 117 particles, of which 67 belonged to non-PET particles and 50 to PET particles. Non-PET particles are from organic matter (19 particles), other polymers such as PE, PP and PS (8 particles), or were not identified (40 particles). PET particles were supplied by the German Federal Institute for Materials Research and Testing (BAM) as part of the EU-funded PlasticTrace project. These reference materials consisted of soluble tablets with secondary PET microparticles, which had irregular shapes and sizes ranging from 10 µm to 100 µm. Non-PET plastic polymers were analysed in various physical forms (e.g., pellets, powders, films) with a size range from 0.05 mm to 5 mm. While most PET tablets analysed were dissolved in ultrapure water, some were dissolved in powdered milk (infant formula) after reconstitution in ultrapure water. Out of these 50 PET spectra, 14 were within the milk matrix.

Confirmation of method performance was achieved by analysing 405 particles from the same source of particles used for method validation, 57 of which were classified as PET and 348 as non-PET, to verify the determined *TP* and *FP* during method validation.

In micro-FTIR, the spectral resolution is mainly influenced by the beam splitter, detector type, and selected wavenumber resolution setting. The beam splitter determines the spectral range and efficiency of light transmission, which in turn affects how distinctly spectral bands are separated. The detector type affects sensitivity and response speed; for instance, mercury cadmium telluride (MCT) detectors provide higher sensitivity and faster response times than deuterated triglycine sulphate (DTGS) detectors, allowing for higher-resolution data collection. The microscope magnification also influences the results: higher magnification decreases the sampling area, which can limit the number of photons reaching the detector and thus reduces the effective resolution of fine spectral details if insufficient signal is collected. The selected wavenumber resolution setting (e.g., 2 cm^{-1} vs. 8 cm^{-1}) directly controls how sharply adjacent peaks are distinguished — the lower the spectral resolution, the higher the resolving power, but the longer the acquisition time. [14]

The S/N is heavily affected by time per scan, the number of co-added scans, and detector sensitivity. Longer scan times and more co-added scans both enhance S/N by averaging out random noise, with S/N improving approximately as the square root of the number of scans. However, these settings increase acquisition time. Detector type also plays a crucial role: MCT detectors typically yield a higher S/N than DTGS detectors due to their superior sensitivity, although they may require cooling. Microscope magnification reduces the sampling area, lowering photon throughput and thus decreasing S/N, if the sample is small or weakly absorbing. [14]

Table 1. FTIR spectrometers and instrumental conditions used to collect IR spectra.

Lab.	Spectrometer	BS	Detector/ Mode	MM	[WI]/WR (cm^{-1})	TS (s)/ NS	Substrate
1	Bruker Vertex 70 coupled with Hyperion 2000 IR Microscope	KBr	MCT/ PC	15x	[600 to 4000]/2	1/32	Silicon filter [5 µm pore size, 9 mm round]
2	Thermo Scientific™ Nicolet iN10	KBr	MCT/ PC	15x	[650 to 4000]/8	3/16	Silicon filter [5 µm pore size, (10 x 10) mm]
3	PerkinElmer FT-IR Microscope Spotlight 200i	OptKBr	MCT/ PC	1 x	[600 to 4000]/4	1/4	Glass fibre membranes [1.2 µm pore size, Ø 47 mm]

BS/S - Beam splitter; MCT - mercury cadmium telluride detector; PC – Point collection; MM - Microscope magnification in times; WI - Wavenumber interval in cm^{-1}; WR - Wavenumber resolution in cm^{-1}; TS – Time per scan in s; NS – Number of co-added scans.

Method validation strategy

To validate the method for PET microparticles classification using FTIR, spectra from PET and non-PET particles were collected using the spectrometers under the spectral conditions listed in Section “Material and equipment”. Spectra were converted into 95 points of “Transmittance vs Wavenumber” at equidistant values between 670 cm^{-1} and 3960 cm^{-1}. All spectra, provided as Supplementary Material 01, were manually identified by considering characteristic bands to act as a reference for method validation.

PET is a polymer whose monomer contains multiple and distinct functional groups, such as carboxylic acid and ester, both presenting a shared benzene substituent, as it is shown in Fig 1. In the literature [6,15], it is expected to observe a C=O stretch for aryl esters in wavenumbers between 1730 cm^{-1} and 1715 cm^{-1}, while for carboxyl acids the wavenumber should be between 1725 cm^{-1} and 1700 cm^{-1}. Aryl esters also exhibit two strong C-O stretches, one between 1275 cm^{-1} and 1200 cm^{-1}, and another between 1075 cm^{-1} and 1020 cm^{-1} [6]. It is also expected to observe a peak between 3000 cm^{-1}

View Article Online
DOI: 10.1039/D6VA00098C

and 2900 cm^{-1}, due to the C-H stretch [5]. Additionally, PET spectra may present a characteristic broad band around 2500 cm^{-1} and 2700 cm^{-1} as a result of hydrogen-bonded carboxylic acids [5].

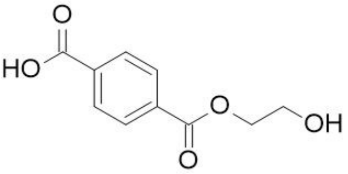

Fig. 1 Monomeric structure of PET.

Consequently, spectra were manually identified as PET if they exhibit, at least, the simultaneous presence of peaks at approximately 1730 cm^{-1} – 1715 cm^{-1}, 1275 cm^{-1} – 1200 cm^{-1}, 1175 cm^{-1} – 1075 cm^{-1}, 840 cm^{-1} – 874 cm^{-1}, and 700 cm^{-1} – 740 cm^{-1}. [6,16] Two different collected PET spectra, shown in Fig. 2, present all five characteristic peaks considered, although they differ in the intensity of the broad band at wavenumbers 3100 cm^{-1} to 2850 cm^{-1}. This band can represent both characteristic C-H stretching and/or O-H stretching.

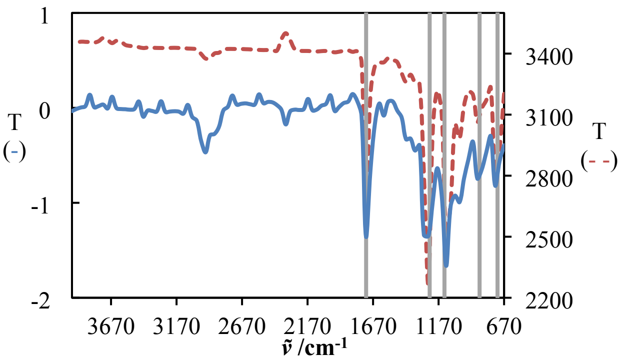

Fig 2. Harmonised PET spectra collected with different equipment and under various spectra collection conditions. The vertical lines indicate wavenumbers at 1723 cm^{-1}, 1238 cm^{-1}, 1125 cm^{-1}, 857 cm^{-1} and 720 cm^{-1}, representing PET characteristic peaks.

The match between reference, a PET spectrum obtained by Attenuated Total Reflectance (ATR), and PET or non-PET particle spectra was determined for each algorithm listed in Table 2. The studied correlation coefficients (*CC*) are Pearson's, *r*, Spearman's, *ρ*, and an alternative coefficient, C_{PE}, as well as their respective weighted versions. Overall, 81 matches were analysed, comprising 27 with unweighted coefficients and 54 with weighted coefficients.

In this work, PET particles were classified as positive cases, while non-PET particles were classified as negative cases. For the spectra of positive cases, the 5th percentile (*P5*»P) was determined using the bootstrap resampling method, which simulates its distribution without assuming normality [1,17]. In fact, *P*5»P is the 5th percentile of the 5th percentile distribution of match values for positive cases simulated by the bootstrap method; 95 % of match values for positive cases are above *P*5.

FP is then calculated as the complementary of the cumulative t-distribution of (*P5*»P – *M*)/*s* for *n*-1 degrees of freedom, where *M*, *s*, and *n* represent the mean, standard deviation, and total number of match values between reference and negative cases, respectively. The *FP* estimates the probability of a non-PET particle being misclassified as a PET particle, meaning it assesses the chances of the algorithm establishing a correlation between a non-PET particle spectrum and a PET reference spectrum above *P5*»P.

Table 2. Compilation of match algorithm identification codes. The codes are listed in the order (*S*|*d*|*CC*|*SW*|*RW*), where *"S"* indicates signal type (*I*, 1-*I* and 1-I_N , for original, inverse and complementary intensity after signal normalization – numbers 1, 2 and 3); *"d"* denotes original signal, first derivate or second derivate (numbers 1, 2 and 3); *"CC"* represents the studied correlation coefficient; *"SW"* signifies signal weighting, applicable only to weighted coefficients and always set to 1; *"RW"* indicates direct ($\tilde{\nu}$) or inverse ($1/\tilde{\nu}$) wavenumber weighting (numbers 1 or 2).

Match	*S*	*d*	*RW*	Match	*S*	*d*	*RW*
1|1|*CC*|1|1	*I*	*S*	$\tilde{\nu}$	1|1|*CC*|1|2	*I*	*S*	$1/\tilde{\nu}$
2|1|*CC*|1|1	(1/*I*)	*S*	$\tilde{\nu}$	2|1|*CC*|1|2	(1/*I*)	*S*	$1/\tilde{\nu}$
3|1|*CC*|1|1	(1 - I_N)	*S*	$\tilde{\nu}$	3|1|*CC*|1|2	(1- I_N)	*S*	$1/\tilde{\nu}$
1|2|*CC*|1|1	*I*	*f'*(*S*)	$\tilde{\nu}$	1|2|*CC*|1|2	*I*	*f'*(*S*)	$1/\tilde{\nu}$
2|2|*CC*|1|1	(1/*I*)	*f'*(*S*)	$\tilde{\nu}$	2|2|*CC*|1|2	(1/*I*)	*f'*(*S*)	$1/\tilde{\nu}$
3|2|*CC*|1|1	(1 - I_N)	*f'*(*S*)	$\tilde{\nu}$	3|2|*CC*|1|2	(1 - I_N)	*f'*(*S*)	$1/\tilde{\nu}$
1|3|*CC*|1|1	*I*	*f''*(*S*)	$\tilde{\nu}$	1|3|*CC*|1|2	*I*	*f''*(*S*)	$1/\tilde{\nu}$
2|3|*CC*|1|1	(1/*I*)	*f''*(*S*)	$\tilde{\nu}$	2|3|*CC*|1|2	(1/*I*)	*f''*(*S*)	$1/\tilde{\nu}$
3|3|*CC*|1|1	(1 - I_N)	*f''*(*S*)	$\tilde{\nu}$	3|3|*CC*|1|2	(1 - I_N)	*f''*(*S*)	$1/\tilde{\nu}$

Among the tested algorithms, only those with a *TP* of 95 % and *FP* no greater than 5 % were regarded as adequate. Since an *FP* of 5 % is adequate for assessing the contamination with microplastics of environmental samples, a likelihood ratio (*LR*) or accuracy (*A*) should not be lower than 19 (*LR* = *TP/FP)* or 95 % (*A* = *(TP+TN)/(TP+TN+FP+FN)*), respectively. [1,17,18]

Determination of signal-to-noise ratio

In FTIR, baseline inconsistencies are occasionally observed throughout the spectra. Often, the baseline appears sloping down or raised, either in a proportional manner throughout the entire wavenumber range or in specific spectral zones. The baseline variations are typically caused by light scattering resulting from instrumental problems, background measurement, sample characteristics, or biofilm. Biofilm tends to present a lowered baseline in the fingerprint region (between 1280 cm^{-1} and 860 cm^{-1}).[1] Therefore, an algorithm to determine the signal-to-noise ratios suitable for such baseline movements was developed. The spectra region examined was divided into 20 uniformly spaced wavenumber intervals. Subsequently, each wavenumber interval was split into two equal segments, and the baseline for each was established using a two-point regression considering the maximum signal intensity and the corresponding wavenumber of both equal segments. Using this regression, the baseline signal at the wavenumber of the lowest signal peak within the interval was obtained. After subtracting the lowest intensity signal from the baseline signal, the highest absolute difference obtained from all 20 intervals was considered the signal, S_H, of the more intense peak of the spectrum. Noise value, N, was determined by subtracting the maximum signal intensity from the average signal for each interval, then averaging these values across all 20 intervals. The ratio S_H/N, designated as the maximum signal-to-noise ratio, was used to quantify the intensity of the spectral bands. Figure 3 illustrates the determination of the S_H/N of a spectrum.

View Article Online
DOI: 10.1039/D6VA00098C

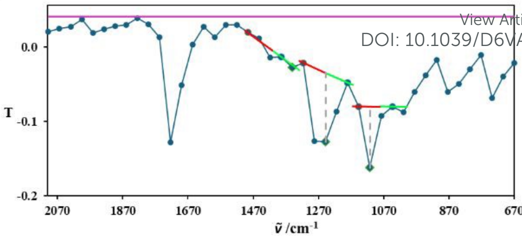

Fig 3. FTIR spectrum demonstrating the robustness of S/N measurement despite baseline variations for three different peaks. Red and green lines indicate the first and second baseline segments for three wavenumber intervals. The dashed lines identify transmittance peaks.

Results and discussion

Table 3 lists the valid algorithms for identifying PET particles, considering the first 117 spectra (50 n_P and 67 n_N spectra for positive and negative cases, respectively), and only those spectra with an S/N ratio greater than 3 (n_P = 50 and n_N = 55). The table lists the minimum match value for a *TP* = 95 %, *P5*»P, the *FP,* and the *LR*. For an identification criterion set for a *TP* of 95 %, identifications are considered valid if *FP* ≤ 5 %. Fig. 4 displays a graphical representation of the correlations between reference and particle spectra for both positive and negative cases, as well as threshold, *P5*»P, for each valid match, including a) all spectra, and b) only spectra with S/N > 3. Overall, all matches with *FP* ≤ 5 % demonstrate a clear ability to differentiate between positive and negative cases, which strongly indicates the method's effectiveness. The identification performance improves slightly when only signals with an S/N ratio above 3 are considered.

Table 3. The best-performing match algorithms for all spectra and those with a signal-to-noise ratio (S/N) greater than 3. n_P and n_N denote the total number of PET and non-PET spectra, respectively. The match notation listed in Table 1 is utilised.

All spectra (n_P = 50 and n_N = 67)				Spectra with S/N > 3 (n_P = 50 and n_N = 55)			
Match	*P5*»P	*FP* / %	*LR*	Match	*P5*»P	*FP* / %	*LR*
1|2|C_{PE}	0.4140	7.4×10^{-4}	13×10^{4}	1|2|C_{PE}	0.4140	5.0×10^{-4}	19×10^{4}
1|3|*r*	0.3240	9.0×10^{-4}	11×10^{4}	1|2|*r*	0.4140	6.6×10^{-4}	14×10^{4}
1|3|C_{PE}	0.3242	9.1×10^{-4}	10×10^{4}	1|3|*r*	0.3021	1.4×10^{-3}	6.6×10^{4}
1|2|*r*	0.4082	1.2×10^{-3}	7.8×10^{4}	1|3|C_{PE}	0.2910	2.2×10^{-3}	4.3×10^{4}
1|2|*p*	0.2895	1.6	58	1|2|*p*	0.3212	0.5	17×10^{1}

The match with the lowest false positive rate, specifically 1|2|C_{PE}, was subjected to a performance assessment. This involved applying the validated PET identification tool to 405 particles, of which 57 had been manually identified as PET and 348 as non-PET. In this test, one false negative (a PET particle misclassified as non-PET) and zero false positives (a non-PET particle misclassified as PET) were observed, resulting in an experimentally determined *FN* of 1.8 % (1/57) and *FP* of 0 % (0/348). Assuming binomially distributed classification errors with nominal *FN* = 5 % and *FP* = 5.0×10^{-4} %, the 99 % Clopper–Pearson confidence intervals are 1.1 % to 13.9 % for the *FN* and 0 % to 1.3 % for the *FP*.[19] Therefore, the independently estimated *FN* and *FP* values, 1.8 % and 0 %, respectively, are fully consistent with those obtained during method validation.

Figure 5 shows the PET spectra with an S/N of 7.1 that was incorrectly identified as non-PET, overlapped with the reference PET spectra. The 1|2|C_{PE} Match value is 0.363, which is below the *P5*»P of 0.4140, due to more significant spectral mismatches at the lower wavenumbers of the fingerprint region.

It should be noted that the match algorithm selected as the best performing, and the determined *FP*, depend on the equipment studied, their instrumental conditions, and the diversity of positive and negative cases. For instance, only spectra from negative cases are available from Laboratory 1. More importantly, the described Clopper–Pearson method can be used to identify cases where new data represent a reality more diverse than the previously studied one, requiring the revision of the identification method.

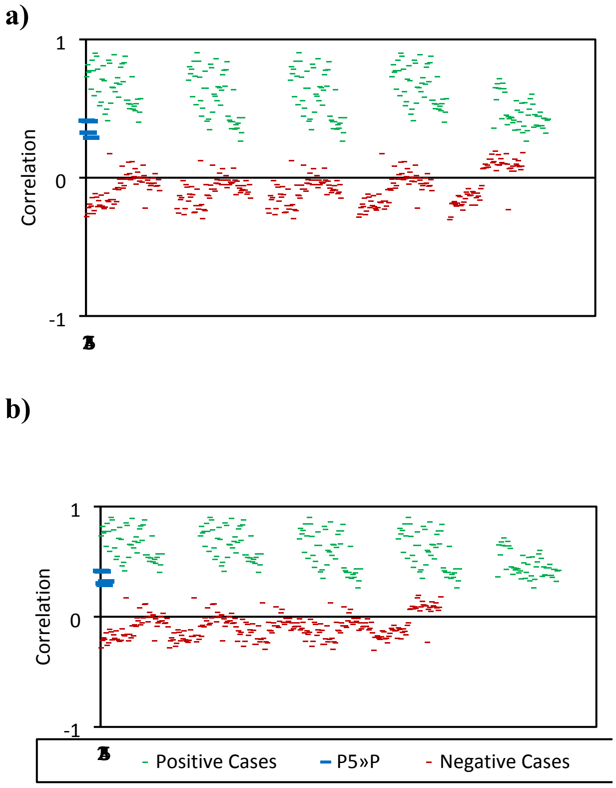

Fig. 4 Representation of match values for positive and negative cases by the valid match methods for a) all spectra, and b) spectra with S/N > 3.

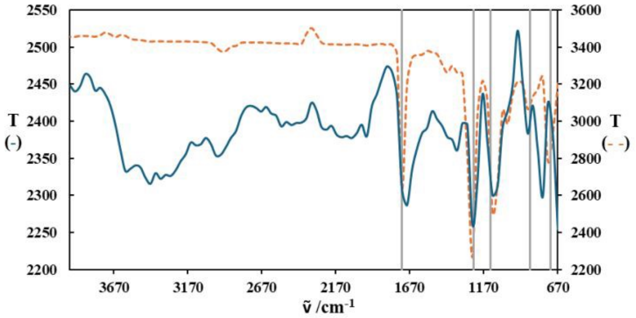

Fig. 5 Overlapped FTIR spectra of the PET reference (dashed line) and a manually identified PET particle (continuous line) misclassified by the validated automatic method 1|2|C_{PE}. The vertical lines correspond to wavenumbers 1723 cm^{-1}, 1238 cm^{-1}, 1125 cm^{-1}, 857 cm^{-1} and 720 cm^{-1} of PET characteristic peaks.

Conclusions

The methodology described for developing valid methods to identify microplastics using micro-FTIR, applicable across various instruments and conditions, was successfully applied to the identification of PET microplastics with three spectrometers and different instrumental settings—these settings influenced spectral resolution and S/N, thereby affecting comparison with reference spectra. PET spectra from two spectrometers were considered. Identifications were considered valid if they were associated with a true positive rate of 95 % and a false positive rate of no more than 5 %.

Various Match methods, including spectral preprocessing such as differentiation and unweighted or weighted correlation coefficients, were tested. It was concluded that when spectra have an S/N greater than three, the unweighted correlation coefficient C_{PE}, an alternative to Pearson’s and Spearman’s, applied to the original spectra after differentiation, yields identification with a Match threshold of 0.44 (*P5*»P), a true positive rate of 95%, and a false positive rate of 0.0005%. The performance estimated during method validation was assessed by experimentally determined false result rates from the identification of 405 particles not considered in the method validation. The estimates *FN* of 1.8 % and *FP* of 0 % are compatible with the *TP* of 95 % and *FP* of 0.0005 %, given the number of particles identified as proven by the Clopper–Pearson method for a 99 % confidence level. This assessment expanded the method scope to a total of 522 particles.

The performance of the developed identification method is equivalent to that reported in the bibliography, which involves a more computationally intense processing using complex chemometric tools.

The developed methodology should be applied to additional equipment and used for identifying other polymers and aged particles to assess their robustness or any need for an algorithm update.

Author contributions

Rafaela Fernandes: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Paul-Tiberiu Miclea:** Writing – review & editing, Resources, Investigation, Funding acquisition, Data curation. **Andrea M. Rossi:** Resources. **Andrea M. Giovannozzi:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **Mara Putzu:** Writing – review & editing, Resources, Investigation, Data curation. **Vanessa Morgado** Conceptualization, Investigation, Writing – reviewing and editing. **Carla Palma:** Writing – review & editing, Funding acquisition, Conceptualization. **José Almeida:** Data curation. **Matilde São Pedro:** Resources. **Claudia Drago:** Project administration, Conceptualization. **Raquel Quendera:** Funding acquisition. **Olivier Pellegrino:** Writing – review & editing, Funding acquisition, Conceptuali zation. **Ricardo J.N. Bettencourt da Silva:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Conflicts of interest

There are no conflicts of interest to declare.

Data availability

Data for this article, including spectra used, for training and validation, are available at Zenodo at https://doi.org/10.5281/zenodo.17791782.

Acknowledgements

The authors thank the European Commission for funding the PlasticTrace project (21GRD07) from the European Partnership on Metrology (Funder ID: 10.13039/100019599), co-financed from the

European Union's Horizon Europe Research and Innovation Program and by the Participating States. This work was supported by Fundação para a Ciência e a Tecnologia (FCT) (Projects UIDB/00100/2020 and UIDP/00100/2020), and Institute of Molecular Sciences (Project LA/P/0056/2020).

Notes and references

1. Morgado V, Gomes L, Bettencourt da Silva RJN, Palma C. Validated spreadsheet for the identification of PE, PET, PP and PS microplastics by micro-ATR-FTIR spectra with known uncertainty. Talanta. 2021;234(April):122624.
2. ISO 24187:2023(en), Principles for the analysis of microplastics present in the environment [Internet]. [cited 2025 Nov 14]. Available from: https://www.iso.org/obp/ui/en/#iso:std:iso:24187:ed-1:v1:en
3. Redko V, Wolska L, Potrykus M, Olkowska E, Cieszyńska-Semenowicz M, Tankiewicz M. Environmental impacts of 5-year plastic waste deposition on municipal waste landfills: A follow-up study. Science of The Total Environment. 2024 Jan 1;906:167710.
4. Yang Y, Jalalah M, Alsareii SA, Harraz FA, Thakur N, Zheng Y, et al. Plastic wastes (PWs) and microplastics (MPs) formation: Management, migration, and environmental impact. J Environ Chem Eng. 2024 Jun 1;12(3).
5. Skoog DA, Holler FJ, Crouch SR. Instrumental Analysis Principles. 7th ed. Veronika R. Meyer. Cengage Learning; 2018. 429 p.
6. Fleming I, H. Williams D. Métodos Espectroscópicos en Química Orgánica. 1st ed. Carrasco Guimeno F (translator), editor. Bilbao: Ediciones Urmo; 1968.
7. Araujo CF, Nolasco MM, Ribeiro AMP, Ribeiro-Claro PJA. Identification of microplastics using Raman spectroscopy: Latest developments and future prospects. Water Res. 2018;142: 426-440.
8. Villegas-Camacho O, Francisco-Valencia I, Alejo-Eleuterio R, Granda-Gutiérrez EE, Martínez-Gallegos S, Villanueva-Vásquez D. FTIR-Based Microplastic Classification: A Comprehensive Study on Normalization and ML Techniques. Recycling. 2025;10(2):1–20.
9. Primpke S, Cross RK, Mintenig SM, Simon M, Vianello A, Gerdts G, et al. Toward the Systematic Identification of Microplastics in the Environment: Evaluation of a New Independent Software Tool (siMPle) for Spectroscopic Analysis. Appl Spectrosc. 2020;74(9):1127–38.
10. Willans M, Szczecinski E, Roocke C, Williams S, Timalsina S, Vongsvivut J, et al. Development of a rapid detection protocol for microplastics using reflectance-FTIR spectroscopic imaging and multivariate classification. Environmental Science: Advances. 2023 Apr 3;2(4):663–74.
11. Clough ME, Rivera EO, Parham RL, Ault AP, Zimmerman PM, Mcneil AJ, et al. Enhancing Confidence in Microplastic Spectral Identification via Conformal Prediction. Environ Sci Technol. 2024;58:21740–9.
12. Kozloski R, Cowger W, Arienzo MM. Moving toward automated µFTIR spectra matching for microplastic identification: addressing false identifications and improving accuracy. Microplastics and Nanoplastics. 2024;4(1).
13. Fernandes R, Miclea PT, Fadda M, Putzu M, Sacco A, Rossi AM, et al. Inter-instrument definition of valid criteria for the automatic identification of microplastics by micro-Raman spectroscopy. Talanta. 2025 Sep 9;128834.
14. Griffiths PR, De Haseth JA. Fourier Transform Infrared Spectrometry: Second Edition. Fourier Transform Infrared Spectrometry: Second Edition. 2006 Jun 30;1–529.
15. M. Hardwood L, D.W. Claridge T. Introduction to Organic Spectroscopy. 1st ed. New York: Oxford University Press; 1997.
16. Chen Z, Hay JN, Jenkins MJ. FTIR spectroscopic analysis of poly(ethylene terephthalate) on crystallization. Eur Polym J. 2012;48(9):1586–610.
17. Morgado V, Palma C, Bettencourt da Silva RJN. Microplastics identification by infrared spectroscopy – Evaluation of identification criteria and uncertainty by the Bootstrap method. Talanta. 2021;224(October 2020):121814.
18. R Bettencourt da Silva, SLR Elison. EURACHEM / CITAC Guide Assessment of performance and uncertainty in qualitative chemical analysis. 2021.
19. Casella G, Berger RL. Statistical Inference, Second Edition. Statistical Inference, Second Edition. 2024 Jan 1;1–535.

View Article Online
DOI: 10.1039/D6VA00098C

View Article Online
DOI: 10.1039/D6VA00098C

Data availability

Data for this article, including spectra used, for training and validation, are available at Zenodo at https://doi.org/10.5281/zenodo.17791782.