SoftwareX 16 (2021) 100857

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Original software publication

uFTIR: An R package to process hyperspectral images of environmental samples captured with μFTIR microscopes

Fabio Corradini [a,b,*], Nicolas Beriot [b,c], Esperanza Huerta-Lwanga [b,d], Violette Geissen [b]

[a] Instituto de Investigaciones Agropecuarias, INIA La Platina, Casilla 439, Correo 3, Santiago, Chile
[b] Soil Physics and Land Management Group, Wageningen University & Research, Droevendaalsesteeg 3, 6708PB Wageningen, The Netherlands
[c] Sustainable Use, Management and Reclamation of Soil and Water Research Group, Universidad Politécnica de Cartagena, Paseo Alfonso XIII, 48, 30203 Cartagena, Spain
[d] Agroecología, El Colegio de la Frontera Sur, Unidad Campeche, Av Polígono s/n, Cd. Industrial, Lerma, Campeche, Mexico

ARTICLE INFO

Article history:
Received 14 April 2020
Received in revised form 4 August 2021
Accepted 19 October 2021

Keywords:
Hyperspectral images
Environmental monitoring
Microplastics
IR microscopy
Plastic pollution

ABSTRACT

uFTIR is an R package that implements an automatic approach to analyze μFTIR hyperspectral images with a strong focus on microplastic recognition in environmental samples. The package performs image classification using a Spectral Angle Mapper algorithm in a library search approach. It interacts with other R packages used for spectral analysis. It exports its output as raster and vector files that can be post-processed in common Geographical Information Systems software. The package was designed around the principles of modular development, compatibility, and open-source software. We hope our contribution will serve researchers to size the occurrence of microplastics in ecosystems.

Code metadata

Current code version	v0.1.1
Permanent link to code/repository used for this code version	https://CRAN.R-project.org/package=uFTIR https://github.com/ElsevierSoftwareX/SOFTX_2020_170
Code Ocean compute capsule	https://doi.org/10.24433/CO.5579643.v1
Legal Code License	GPL-3
Code versioning system used	git
Software code languages, tools, and services used	R C++11
Compilation requirements, operating environments & dependencies	64-bit operating system R environment $\geq$ 3.5.2 *GDAL* $\geq$ 2.4.0 *PROJ* $\geq$ 5.2.0 *GEOS* $\geq$ 3.7.1 C++11 R packages: Rcpp, RcppArmadillo, raster, sp, rgdal.
If available Link to developer documentation/manual	https://cran.r-project.org/web/packages/uFTIR/uFTIR.pdf
Support email for questions	fabio.corradini@inia.cl

* Corresponding author at: Instituto de Investigaciones Agropecuarias, INIA La Platina, Casilla 439, Correo 3, Santiago, Chile.
E-mail address: fabio.corradini@inia.cl (Fabio Corradini).

https://doi.org/10.1016/j.softx.2021.100857

Software metadata

Current software version	v0.1.1
Permanent link to executables of this version	https://CRAN.R-project.org/package=uFTIR
Legal Software License	GPL-3
Computing platforms/Operating Systems	Linux, Microsoft Windows
Installation requirements & dependencies	64-bit operating system R environment $\geq$ 3.5.2 *GDAL* $\geq$ 2.4.0 *PROJ* $\geq$ 5.2.0 *GEOS* $\geq$ 3.7.1 C++11 R packages: Rcpp, RcppArmadillo, raster, sp, rgdal. On Microsoft Windows local versions of GDAL, GEOS, and PROJ are installed.
If available, link to user manual - if formally published include a reference to the publication in the reference list	https://cran.r-project.org/web/packages/uFTIR/uFTIR.pdf
Support email for questions	fabio.corradini@inia.cl

1. Motivation and significance

In the last decade, scientific concerns about environmental pollution by microplastic have scaled up, reached public opinion, and positioned within the political agenda [1,2]. With all the evidence that scientists have gathered, it is conceivable that policy-makers will promote routine environmental monitoring programs [3]. However, a problem might hamper the development of such initiatives. The same problem that has hindered research for years. Scientists have not agreed on standard methods to quantify or identify microplastics in environmental samples [4,5]. The problem hampers not only future monitoring initiatives, but it precludes study comparisons and metadata analyses [6].

To date, scientists have mainly used single purpose methods and low laboratory automation to address microplastic pollution [7]. The lack of standardized methods rises as a consequence of such approach [8]. Available analytical methods propose three step analysis that consider extraction, instrument detection, and particle identification and count. Although standards lack for each of the steps, scientists struggle the most to achieve both particle quantification and particle count using a single method or instrument [9]. Commonly, methods focus in one or the other, manual identification being the most cumbersome step [7]. To tackle the problem, scientists have proposed workflows that include laboratory automation.

Laboratory automation has two sides; hardware and software. The industry has tackled hardware requirements and scientists have at disposal equipment capable of identifying plastic polymers [7,8,10]. Literature reviews that summarize monitoring efforts identify FTIR spectroscopy as the most common method used to identify plastic polymers [11–13]. To provide a complete solution, manufacturers couple FTIR spectrometers with microscopes. The use of μFTIR instruments – as they are called – avoids unnecessary steps in sample handling. However, manufacturers do not provide tools to automate the analysis of the output image.

On the software side, companies do not provide built-in solutions to process the output images automatically. Equipment such as Agilent Cary 620 FTIR spectrometer come with a (proprietary) software that has only basic pixel classification features [14]. The situation is not uncommon for other equipment [7]. Researchers have taken the lead, proposing different approaches to fulfill software requirements which are based on machine learning [15–18] or bulk library search [8,19–23]. Bulk library search predominates [7,8,24]. Bulk library search presents the advantage that it can be adapted quickly through the implementation of new or extended reference libraries, but it can be computationally intensive [21]. Recently, researchers optimized the method performance by clustering the spectra before the search [22]. An approach that earth scientists follow when analyzing spectral data [25].

Scientists have implemented a few alternatives to overcome the absence of software officially supported by μFTIR instrument providers. The Systematic Identification of MicroPLastics in the Environment (siMPle) software is a good representative of the alternatives to date in one package suit [20]. The software has some limitations. First, it has shortcomings when dealing with large files —a single sample file size starts from 12Gb. Second, the developers restrict the access to the source code, reducing the possibilities of software modification and adaptation to particular research needs [19,26]. Third, the code's obscurantism veils the analytical workflow and forces the user to choose between a finite set of pre-processing steps. The software limits the analytical possibilities to two algorithms to pre-process the data (calculate spectra's first or second derivative) and one algorithm to perform the library search (correlation). These options fall short when compared with typical spectral analysis steps [27].

Given the software limitations, we set out to develop a program able to automate the analysis of μFTIR images built on trustworthy and reproducible research principles(see Chambers's [26] explanation on the subject). Our main goal was to implement a set of front-end tools to analyze the output of μFTIR spectrometers. Our main focus was its application in environmental research, especially for microplastics analysis. We addressed our goal by writing an R package that structures a library search workflow around the principles of modular development, compatibility, and open-source software. In this article we introduce the uFTIR R package architecture, describe its functionality, present a step by step processing of a soil sample, and contrast the results with alternative software (siMPle).

2. Conceptualization and requirements

2.1. Analytical steps

The analysis of hyperspectral images comprises five sequential steps that we tackled independently; load, pre-process, process, post-process, and summarize (see Fig. 1). The current version (v0.1.1) of the uFTIR package implements all the analytical steps to process Agilent Resolutions Pro Software outputs files [14]. The Agilent Resolutions Pro Software comes with the Agilent μFTIR Microscope and Bench (Agilent Technologies, Inc., USA) and together constitute Agilent's suit for FTIR microscope analysis.

Agilent's suit for FTIR microscope analysis allows a spectra recording between 3600 and 700 cm^{-1} with a collection resolution between 0.5 and 16 cm^{-1}. The manufacturer offers three

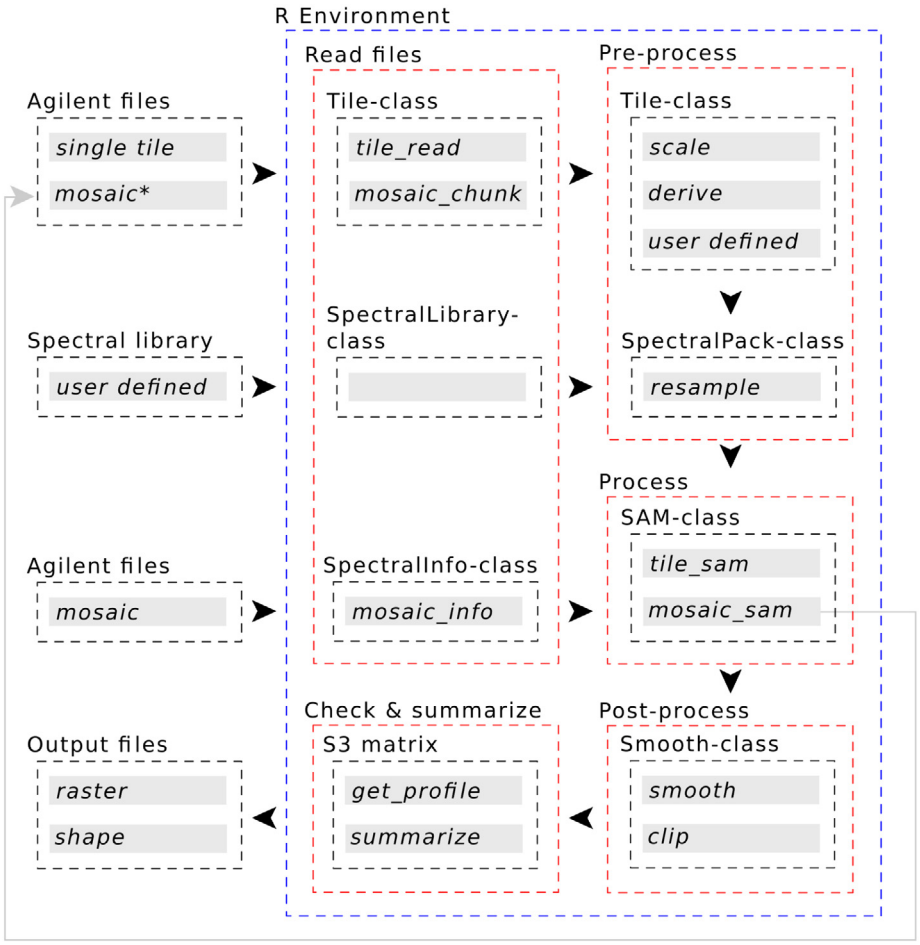

Fig. 1. uFTIR R Package architecture. The blue box aggregates the processes that the package performs, all of them inside the R environment. The red boxes individuate each analytical step. The gray arrow shows how mosaic files are read. By calling the mosaic_sam function the user process each mosaic* (sub)file(s) in one call. The summary method returns a three column table with information about the number of particles, their area, and the cluster or substance to which they correspond. The method vectorizes the image that can be saved as ESRI shapefile format. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

objectives to equip the microscope; 4x, 15x, and 25x, which yield images with a pixel size of 20.6, 5.5, and 1.1 μm respectively. Agilent Resolutions Pro Software can do only pixel wise library search to compare its output with a known reference. The software comes with a (privative) spectral library for plastic polymers identification called poly_8. Agilent's software stores the images in a file format with special characteristics that we use as input to start the microplastic recognition analysis in uFTIR. In the next section we describe the image characteristics.

It is worth noticing that, despite we focus the analysis and comparison on Agilent's FTIR microscope images, the package works with any hyperspectral image. We use here the case of Agilent images as: (a) it exemplifies the difficulties that close source formats pose to science; (b) it represents the most difficult analytical scenario.

2.2. Input files

In this subsection we present the challenges of extending R reading functionalities to load Agilent's FTIR microscope images. Agilent's FTIR microscope had two main output file formats. In its most simple usage, the microscope takes the spectra of a *single tile*; a single hyperspectral image taken at a fixed position. *Mosaics* extends the single tile format to multiple images. Mosaics constitute the working horse of all automation efforts. They allow the user to take hyperspectral images of an area larger than the microscope field of view. When the mosaic approach is used, the user defines *a priori* an area to record. Then, the microscope takes the images and moves its tray until it covers the whole area. As a result, mosaic images are a record of multiple single images (chunks) with a header that identifies them.

Agilent's output formats pose challenges for post-processing. Agilent's software stores its output in a proprietary file format. It does provide a translation feature to convert the files to ENVI, another proprietary software commonly used to analyze spatial imagery. Currently, the Comprehensive R Archive Network (CRAN) does not register any packages to read Agilent file formats. Although the R package caTools can load ENVI files into memory [28], the problem persists as mosaic files are typically too large to be loaded without processing them first.

2.3. Reference library

Library search methods rely on the availability of comprehensive reference libraries. Unfortunately, researchers lack free access to such resources. Primpke et al. [21] published the first freely available library tailor-made for microplastic identification. The library includes 270 substances manually aggregated in 32 clusters that stand for different plastic polymers. It includes other polymers commonly found in environmental samples which might cause misclassification, such as chitin, cellulose, and animal fur. Since Primpke et al. library is the only spectral

library freely distributed among scientists, we included it in the uFTIR package as accompanying data.

3. Software description

3.1. Software architecture

The scientific context defined in Section 1 and the characteristics of the input files served as the cornerstone to design the package principles:

- Researchers are the program end-users.
- The program must be modular and accept user modifications.
- The program must be compatible with processing algorithms implemented already for spectral analysis.
- The program must support stepwise checking of module success and user exploration.
- The program must not overload the host memory. Mosaics should be processed in chunks, since they are usually large files for personal computers.
- Memory intensive processes should be parallelized, taking advantage of the chunk-processing approach.

We implemented the application as an R package and defined its output in a format common for geographical information systems (GIS) analysis. This approach has three positive consequences. First, the R environment [29] has a variety of tools implemented already for hyperspectral image analysis. The program can integrate with those, if the user wants to extend the package built-in features. Second, researchers use the R environment frequently to explore, process, and analyze data. The familiarity that they have with the R environment should soften the learning curve of our software. Third, environmental researchers have at least a common knowledge of GIS and GIS software. GIS allows the user to visualize, manipulate, and process spatial data. Open source libraries and software to work with GIS formats are free and well maintained (see GDAL [30], GRASS [31], and QGIS [32]). Researchers can use these suits to summarize and check the data.

Fig. 1 shows the package general workflow. We subdivide the workflow in five sequential steps (see Section 2.1). We describe each feature and its characteristics in Section 3.2. Fig. 3 shows how each step works when processing an environmental sample.

3.2. Software functionalities

The R package uFTIR presents the following major functionalities:

Read files. The uFTIR package defines two classes to manipulate Agilent Resolution Pro FTIR files. One reads a single tile directly into memory and the other creates a virtual class that holds the location of all mosaic subfiles. We based the code that implements the reading process on Henderson's MATLAB solution [33]. We translated the MATLAB code to R (single tiles) and C++ (mosaics) to import the reading functionality to R.

Pre-process. The program implements three methods to pre-process the spectra: scale, calculate first and second derivatives, and resample. The user might use any of these methods to pre-process the spectra. The program includes one additional method to allow user defined pre-process functions. The user can pass either a lambda function or functions defined in other packages. By these means, the user can perform other common pre-processing steps such as applying a Savitzky–Golay filter [34] (see the R package *signal,signal*).

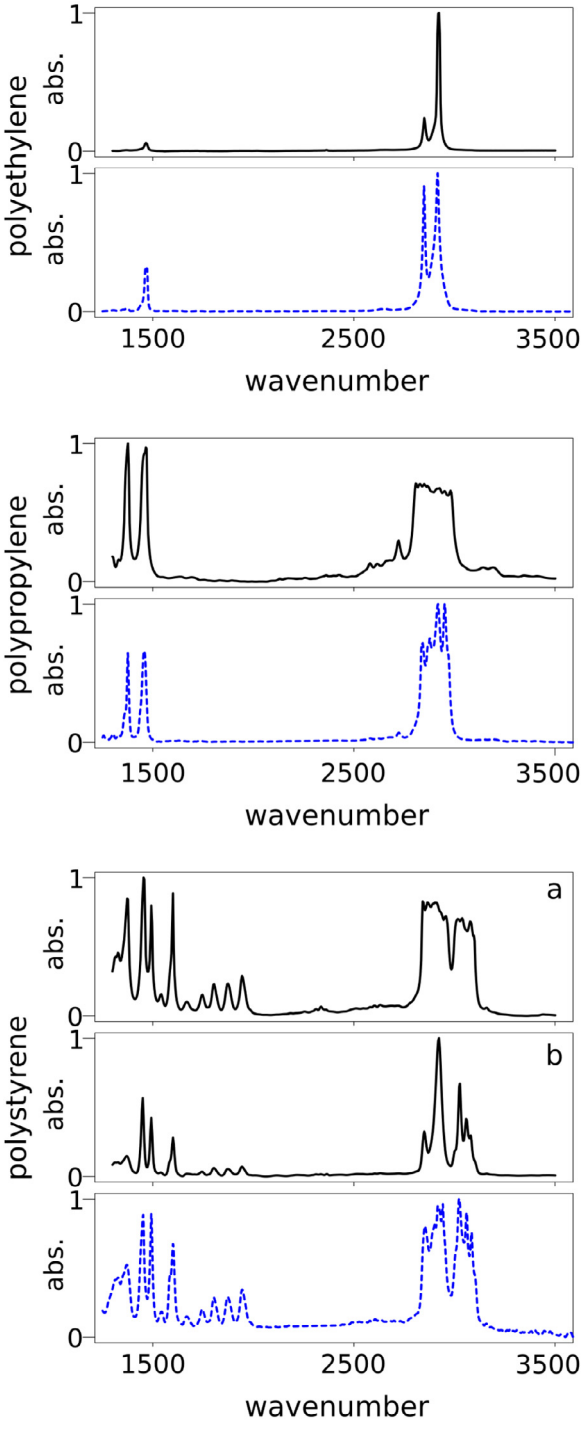

Fig. 2. Spectra recorded for each plastic polymer used to validate the matching algorithm. Black lines present the average spectra of all pixels that matched the target polymer while the blue-dotted lines show the spectral library's spectra for the target polymers. Polystyrene (a) corresponds to the plastic cup sample, while polystyrene (b) corresponds to the polystyrene standard.

Process. Currently the package implements only one algorithm to match pixel spectra with known references. It uses Spectral Angle Mapper (SAM) as implemented in the R package RStoolbox [35]. The algorithm recognizes different polymers successfully [36], and to this end the waste recycling industry has used it for over 15 years [37]. The SAM algorithm is, however, just one of the classical methods used

Table 1
Polymers scanned and analyzed in the validation test: number of particles detected, total area ($pixel^2$), proportion of the total area, other polymers identified in the same image, and the area of those other polymers ($pixel^2$).

Polymer	Part.	Area		Other	Area
	(n)	($pixel^2$)	(prop.)	–	($pixel^2$)
polyethylene	2	15,705	0.96	ethylene-vinyl-acetate	679
polypropylene	1	16,296	0.99	polyethylene	88
polystyrene	1	16,351	>0.99	polypropylene	33
polystyrene standard	1	16,384	1	–	0

for hyperspectral image classification. Researchers have proposed both algorithm optimizations [38,39] and alternative approaches [40]. The package modularity allows the user to add new processing algorithms by calling other R packages.

Post-process. The package implements optional post-processing methods. The method to smooth the output of the SAM algorithm has the highest relevance. The program uses a moving window to remove single-point particles [41, 42].

Check and summarize. To check the accuracy of the library search algorithm, it is possible to retrieve the spectra of pixels that matched a particular substance or cluster. The program defines methods to plot at every step, to allow stepwise inspection of the process performance.

The summary method returns a three column table with information about the number of particles, their area, and the cluster or substance to which they correspond. The user can save the output as common vector or raster formats.

4. Algorithm validation

Although the Spectral Mapper Algorithm discriminates well between polymers [36] we tested whether it was correctly implemented in the uFTIR package. To do so, we recorded the spectra of one polyethylene bag, two plastic cups –one made of polypropylene and the other made of polystyrene–, and a polystyrene standard film (VARIAN P/N 883-9120). A single tile was recorded for each polymer, in transmission mode with a spectral resolution of 8 cm^{-1} through a spectral range of 3500–1300 cm^{-1} and 8 co-added scans. Data was recorded in absorbance (%). The microscope magnification was x4 with a pixel size resolution of 20.6 μm. The analysis used the spectra's first derivatives. The images post-process included smoothing them using a 3 × 3 moving window. We used a freely available spectral library for the library search [21].

Results showed that the algorithm matches the expected polymer in all cases (Table 1). uFTIR classified correctly all pixels of the standard polystyrene film, and almost all pixels of the polystyrene cup. The algorithm was confused in 1% of the cases when it classified the polypropylene cup, attributing wrongly 88 pixels to polyethylene. The analysis of the polyethylene bag had the lowest success rate, misclassifying 4% of the pixels. However, the algorithm attributed those pixels to ethylene-vinyl-acetate, which is a polymer composed by polyethylene and vinyl-acetate in a ratio from 10:1 to 10:4. Fig. 2 shows the average spectra recorded for each of the polymers used in the validation test and contrasts them with their reference spectra.

5. Illustrative example

To illustrate the workflow of uFTIR and compare its output with its alternative (siMPle), we prepared a soil sample and captured its spectral signal (Section 5.1). We processed the image using both uFTIR and siMPle software with similar settings (Section 5.2). siMPle was developed to automate a similar analytical procedure. To produce comparable results, we used Primpke et al. [21] library for both analyses. Sections 5.3 and 5.4 show the results of uFTIR and siMPle, respectively.

5.1. Sample preparation and image acquisition

We selected a soil sample from our archive [43] that had 1.4 plastics particles per gram of soil by Zhang et al. method [44]. The soil sample was suspended in $ZnCl_2$, stirred, centrifuged, and vacuum-filtered three times. At the end of the preparation process, a filter (Whatman(R) Anodisc Inorganic Membranes) that collected all buoyant particles was ready for μFTIR analysis. The μFTIR analysis was performed in transmission mode with a spectral resolution of 8 cm^{-1} through a spectral range of 3500–1300 cm^{-1} and 8 co-added scans. Data was recorded in absorbance (%). The microscope magnification was x4 with a pixel size resolution of 20.6 μm. The final mosaic comprised 64 tiles and 12Gb.

The collected image showed a large plastic particle placed on the filter's lower half. We opened the image in Agilent's Resolution Pro software and performed a library search in 10 random pixels within the particle. We used the correlation algorithm and the *poly_8* built-in library. The particle matched polystyrene in all the 10 runs.

5.2. Hardware information

The image analysis with uFTIR r-package and siMPle software (see Section 5.4) was done in a HP EliteBook 840-g3, Intel(R) Core(TM) i7-6600U CPU @ 2.60 GHz, with 2 cores and 4 threads, and 8GiB of memory. The testing environment was Windows 10 enterprise, with R version 3.5.2, and siMPle Version 1.0.0.

5.3. uFTIR pre-processing and results

The image was processed as mosaic using the package parallel features. The pre-process included scaling and taking the spectra first derivatives. The post-process included smoothing the image with a moving window of 3 × 3 pixels and clipping it to the extent of the filter to leave out the filter's polypropylene support ring. The clipping mask was a circle with a radius equal to 490px and its center placed at (512,512)px. Fig. 3 shows the output of each analytical step.

The analysis revealed the presence of two different polymer clusters on the filter. Polystyrene dominated with 3 particles that accounted for more than 5000 $pixels^2$. Polypropylene was the other, having 48 particles and a total area of ∼800 $pixels^2$. Table 2 reports the summarized output. Fig. 3(f) shows the correspondence between the library spectra (blue-dotted) and the average spectrum of the 3 particles that matched polystyrene (red-solid). The polystyrene particle was fragmented into three particles. However, two of them had an area of $2 pixels^2$ and one encompassed >99% of the particle area.

The analysis took 9 min 52 s to complete (elapsed time).

5.4. siMPle pre-processing and results

The pre-process included cutting out the CO_2 signal and taking the spectra first derivatives. We exported siMPle results as comma delimited to summarize them in R. Fig. 4 shows the image output and a close-up to the large polystyrene particle.

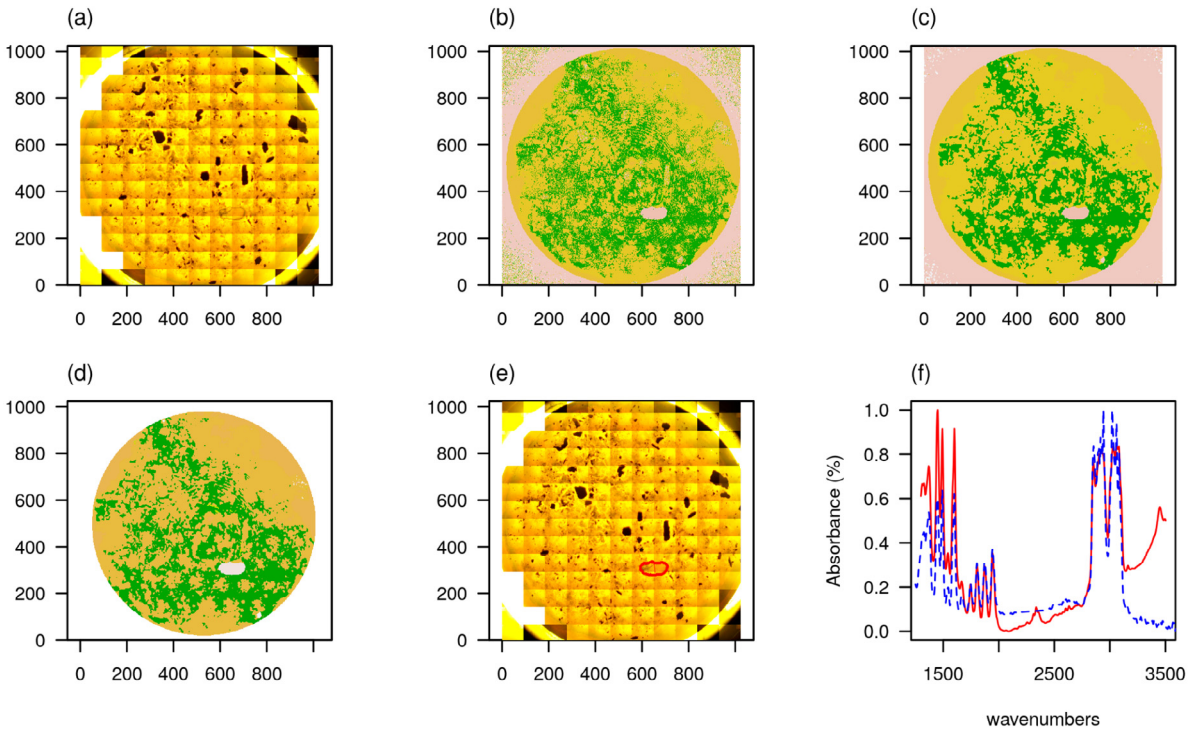

Fig. 3. uFTIR workflow and analysis at different steps for an environmental sample: (a) load the sample (visual) image into memory, (b) pre-process the spectra and run the library search algorithm (spectral angle mapper), (c) post-process the image with a smoother algorithm (moving window), (d) post-process the image removing unnecessary information (clip), (e) check the accuracy by tracing a polygon over all particles matching a given polymer (polystyrene), (f) check the accuracy by comparing the spectral signal (mean) of all particles that matched a given polymer (polystyrene - red-solid line) and the reference spectra of the polymer (polystyrene in Primpke et al. library [21] - blue-dotted line).

Table 2
Summary of uFTIR's analysis.

Cluster name	Number of particles (n)	Total area ($pixel^2$)
animal fur	155	60,877
chitin	41	1,109
coal	482	239,974
plant fibers	433	415,434
polypropylene	48	791
polystyrene	3	5,224

Table 3
Abbreviated summary of siMPle's analysis.

Cluster name	Number of particles (n)	Total area ($pixel^2$)
polypropylene	26,633	1,214,800
not identified	14,516	608,093
cellulose chemical modified	2,793	27,192
acrylates/PUR/varnish	905	3,249
polyimide	16	3,046
polyethylene	878	2,864
polyester	600	1,940
polysulfone	61	1,142
polycaprolactone	286	974
plant fibers	108	447
⋮	⋮	⋮
animal fur	3	10
polystyrene	3	9

The analysis revealed the presence of 18 synthetic polymers. Table 3 shows a synthesis of the output. The polystyrene particle matched both polyimide and polysulfone (and not polystyrene). siMPle identified only 3 polystyrene particles, with a total area of 9 pixels. The large number of particles for each cluster revealed a problem of particle fractionation. The large amount of polypropylene corresponds to the filter's support ring (see Fig. 4(c)). The program has no features to crop, or smooth the output.

The software took 50 s to convert the image to siMPle's format, 18 s to load the reference library, 46 min to analyze the image for spectra fit, and 3 h 48 min 9s to run the *MP detection* algorithm to find the particles. The total time was 275 min.

6. Impact

The uFTIR package provides a general-purpose software to automatize hyperspectral images acquired in μFTIR spectrometers. Its primary orientation is towards microplastic detection. It constitutes a step forward for environmental research as it provides a tool for researchers to increase the accuracy of state-of-the-art analytical methods. The software implements a scalable methodology – in a language familiar to scientists – that quantifies and identifies microplastics in environmental samples. It uses spectral angle mapper, and algorithm that had not been implemented before in any similar software (such as siMPle [20] or MPhunter [8]).

The idea of using the spectral angle mapper algorithm came from earth sciences and Geographical Information Systems. Harris [25] proposed a two step library search that first runs a feature recognition algorithm and calculates spectral end-members and then runs a library search algorithm [25]. The method, known as spectral angle mapper, is widely used in earth sciences as it outperforms other common classification procedures [36–39]. However, the method loses information when calculating the end-members, thus for some applications it should be used with care [45]. Despite its disadvantage, the method reduces the processing time compared to bulk library search [22].

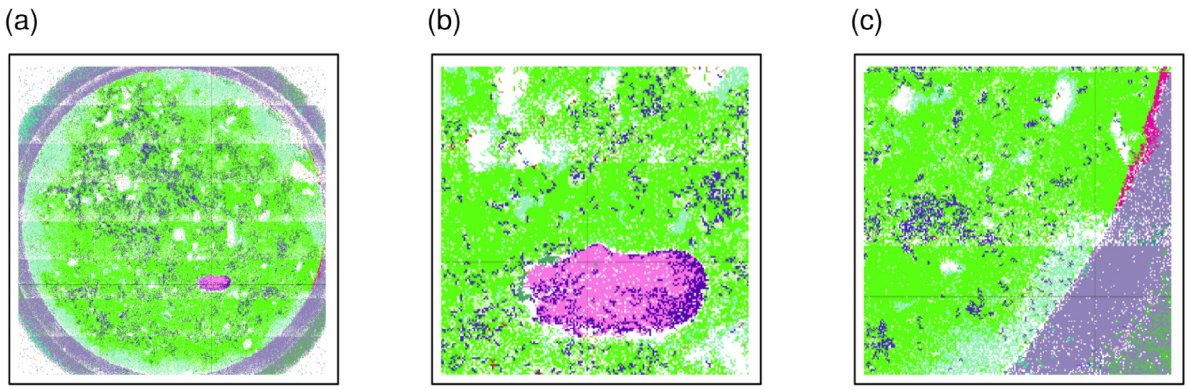

Fig. 4. Output of siMPle: (a) image map after the *MP detection*, (b) close-up to the polystyrene particle, revealing the particle fractionation problem, (c) close-up to the filter support ring that matched polypropylene.

The algorithm and the analytical workflow implemented in uFTIR allows the package to work several orders of magnitude faster than its alternative (siMPle, see Section 5). The increase in speed will allow researchers to increase the number of samples in assessment efforts. This will contribute to size the problem of plastic pollution in ecosystems without the current limitations imposed by time consuming and tedious laboratory routines [46].

The package improves the reproducibility of the results, since procedural scripts can be shared and published together with scientific articles. The software open-source nature allows trustworthy analysis and scientific communication [26]. Moreover, R – a functional programming language – is strongly modular, facilitating the addition of new functions and analytical techniques. To date, researchers have recognized the need for a Open Source Community devoted to contribute to microplastics spectral classification [19]. Our software walks one step into that direction.

In its first release, the package implements only one matching algorithm. However, the R environment is full of packages that can interact in any of the analytical steps described. As a proof of concept, we show an example that uses the R *signal* package [47] to include a Savitzky–Golay filter to pre-process a sample in the CodeOcean capsule that accompanies the article (). In future releases, we expect to include a support vector machine algorithm, another hyperspectral image classification method with good reputation among scientists [40].

7. Conclusions

We presented uFTIR, an R-based software that implements an automatic approach to analyze μFTIR images. The package is mainly oriented towards the analysis of environmental samples and microplastic identification. It supports parallel computations, and interaction with other R packages and procedures. It is fast, compared with other library search alternatives, and it promotes collaborative science through an open-source approach. uFTIR is an ongoing project. We intend to implement additional matching algorithms in future releases, and a pre-processing feature for *a priori* feature recognition. As presented, we hope that our contribution will serve researchers to size the occurrence of microplastics in ecosystems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by Comisión Nacional de Investigación Científica *y* Tecnológica, CONICYT, Chile [grant 72170044].

References

[1] Henderson L, Green C. Making sense of microplastics? Public understandings of plastic pollution. Mar Pollut Bull 2020;152. http://dx.doi.org/10.1016/j.marpolbul.2020.110908.
[2] Foteinis S. How small daily choices play a huge role in climate change: The disposable paper cup environmental bane. J Cleaner Prod 2020;255. http://dx.doi.org/10.1016/j.jclepro.2020.120294.
[3] Fossi MC, Vlachogianni T, Galgani F, Innocenti FD, Zampetti G, Leone G. Assessing and mitigating the harmful effects of plastic pollution: the collective multi-stakeholder driven euro-mediterranean response. Ocean & Coastal Management 2020;184:105005. http://dx.doi.org/10.1016/j.ocecoaman.2019.105005.
[4] Alexy P, Anklam E, Emans T, Furfari A, Galgani F, Hanke G, et al. Managing the analytical challenges related to micro- and nanoplastics in the environment and food: filling the knowledge gaps. Food Addit Contam A Chem Anal Control Expo Risk Assess 2020;37(1):1–10. http://dx.doi.org/10.1080/19440049.2019.1673905.
[5] Van Raamsdonk L, Van Der Zande M, Koelmans A, Hoogenboom R, Peters R, Groot M, et al. Current insights into monitoring, bioaccumulation, and potential health effects of microplastics present in the food chain. Foods 2020;9(1). http://dx.doi.org/10.3390/foods9010072.
[6] Li J, Song Y, Cai Y. Focus topics on microplastics in soil: Analytical methods, occurrence, transport, and ecological risks. Environ Pollut 2020;257:113570. http://dx.doi.org/10.1016/j.envpol.2019.113570.
[7] Renner G, Nellessen A, AlexanderSchwiers, Wenzel M, Schmidt TC, Schrama J. Data preprocessing and evaluation used in the microplastics identification process: A critical review and practical guide. TRAC Trends Anal Chem 2019;111:229–38. http://dx.doi.org/10.1016/j.trac.2018.12.004.
[8] Primpke S, Lorenz C, Rascher-Friesenhausen R, Gerdts G. An automated approach for microplastics analysis using focal plane array (FPA) FTIR microscopy and image analysis. Anal Methods 2017;9:1499–511. http://dx.doi.org/10.1039/C6AY02476A.
[9] Wang W, Ge J, Yu X, Li H. Environmental fate and impacts of microplastics in soil ecosystems: Progress and perspective. Sci Total Environ 2020;708:134841. http://dx.doi.org/10.1016/j.scitotenv.2019.134841.
[10] Dümichen E, Eisentraut P, Bannick CG, Barthel A-K, Senz R, Braun U. Fast identification of microplastics in complex environmental samples by a thermal degradation method. Chemosphere 2017;174:572–84. http://dx.doi.org/10.1016/j.chemosphere.2017.02.010.
[11] Möller J, Löder M, Laforsch C. Finding microplastics in soils: A review of analytical methods. Environ Sci Technol 2020;54(4):2078–90. http://dx.doi.org/10.1021/acs.est.9b04618.
[12] Huang Y, Qing X, Wang W, Han G, Wang J. Mini-review on current studies of airborne microplastics: Analytical methods, occurrence, sources, fate and potential risk to human beings. TRAC Trends Anal Chem 2020;125. http://dx.doi.org/10.1016/j.trac.2020.115821.
[13] Wong J, Lee K, Tang K, Yap P-S. Microplastics in the freshwater and terrestrial environments: Prevalence, fates, impacts and sustainable solutions. Sci Total Environ 2020;719. http://dx.doi.org/10.1016/j.scitotenv.2020.137512.
[14] Agilent Technologies, Inc. Resolutions pro FTIR software. 2011, URL https://www.agilent.com/en/products/software-informatics/molecular-spectroscopy-data-systems/ftir-software/resolutions-pro-ftir-software.

[15] Kedzierski M, Falcou-Préfol M, Kerros ME, Henry M, Pedrotti ML, Bruzaud S. A machine learning algorithm for high throughput identification of FTIR spectra: Application on microplastics collected in the Mediterranean Sea. Chemosphere 2019;234:242–51. http://dx.doi.org/10.1016/j.chemosphere.2019.05.113.
[16] Shan J, Zhao J, Zhang Y, Liu L, Wu F, Wang X. Simple and rapid detection of microplastics in seawater using hyperspectral imaging technology. Anal Chim Acta 2019;1050:161–8. http://dx.doi.org/10.1016/j.aca.2018.11.008.
[17] Serranti S, Palmieri R, Bonifazi G, Cózar A. Characterization of microplastic litter from oceans by an innovative approach based on hyperspectral imaging. Waste Manage 2018;76:117–25. http://dx.doi.org/10.1016/j.wasman.2018.03.003.
[18] Paul A, Wander L, Becker R, Goedecke C, Braun U. High-throughput NIR spectroscopic (NIRS) detection of microplastics in soil. Environ Sci Pollut Res 2018;26:7364–74. http://dx.doi.org/10.1007/s11356-018-2180-2.
[19] Cowger* W, Steinmetz Z, Gray A, Munno K, Lynch J, Hapich H, et al. Microplastic spectral classification needs an open source community: Open specy to the rescue!. Anal Chem 2021;93:7543–8. http://dx.doi.org/10.1021/acs.analchem.1c00123.
[20] Primpke S, A. Dias P, Gerdts G. Automated identification and quantification of microfibres and microplastics. Anal Methods 2019;11:2138–47. http://dx.doi.org/10.1039/C9AY00126C.
[21] Primpke S, Wirth M, Lorenz C, Gerdts G. Reference database design for the automated analysis of microplastic samples based on Fourier transform infrared (FTIR) spectroscopy. Anal Bioanal Chem 2018;410:5131–41. http://dx.doi.org/10.1007/s00216-018-1156-x.
[22] Wander L, Vianello A, Vollertsen J, Westad F, Braun U, Paul A. Exploratory analysis of hyperspectral FTIR data obtained from environmental microplastics samples. Anal Methods 2020;12(6):781–91. http://dx.doi.org/10.1039/c9ay02483b.
[23] Renner G, Sauerbier P, Schmidt TC, Schram J. Robust automatic identification of microplastics in environmental samples using FTIR microscopy. Anal Chem 2019;91:9656–64. http://dx.doi.org/10.1021/acs.analchem.9b01095.
[24] Liu F, Olesen K, Borregaard A, Vollertsen J. Microplastics in urban and highway stormwater retention ponds. Sci Total Environ 2019;671:992–1000. http://dx.doi.org/10.1016/j.scitotenv.2019.03.416, cited By 20.
[25] Harris A. Spectral mapping tools from the earth sciences applied to spectral microscopy data. Cytometry A 2006;69(8):872–9. http://dx.doi.org/10.1002/cyto.a.20309.
[26] Chambers J. Software for data analysis: programming with R. first ed.. New York: Springer-Verlag New York; 2008, http://dx.doi.org/10.1007/978-0-387-75936-4.
[27] Raczkowska M, Koziol P, Urbaniak-Wasik S, Paluszkiewicz C, Kwiatek W, Wrobel T. Influence of denoising on classification results in the context of hyperspectral data: High definition FT-IR imaging. Anal Chim Acta 2019;1085:39–47. http://dx.doi.org/10.1016/j.aca.2019.07.045.
[28] Tuszynski J. caTools: Tools: Moving window statistics, GIF, base64, ROC AUC, etc. 2020, R package v1.18.0. URL https://CRAN.R-project.org/package=caTools.
[29] R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2019, URL https://www.R-project.org/.
[30] GDAL/OGR contributors. GDAL/OGR geospatial data abstraction software library. 2020, Open Source Geospatial Foundation. URL https://gdal.org.
[31] GRASS Development Team. Geographic resources analysis support system (GRASS GIS) software, version 7.2. 2017, Open Source Geospatial Foundation. URL http://grass.osgeo.org.
[32] QGIS Development Team. QGIS geographic information system. 2009, Open Source Geospatial Foundation. URL http://qgis.osgeo.org.
[33] Henderson A. Agilent FTIR mosaic image file reader. 2017, http://dx.doi.org/10.5281/zenodo.399238, Work relates to EPSRC Grant number EP/L012952/1 Clinical Infrared and Raman Spectroscopy Network (CLIRSPEC) http://clirspec.org/.
[34] Jardim R, Morgado-Dias F. Savitzky–Golay filtering as image noise reduction with sharp color reset. Microprocess Microsyst 2020;74. http://dx.doi.org/10.1016/j.micpro.2020.103006.
[35] Leutner B, Horning N, Schwalb-Willmann J, Hijmans R. Tools for remote sensing data analysis. 2019, R package v0.2.6. URL https://CRAN.R-project.org/package=RStoolbox.
[36] Wu X, Li J, Yao L, Xu Z. Auto-sorting commonly recovered plastics from waste household appliances and electronics using near-infrared spectroscopy. J Cleaner Prod 2020;246:118732. http://dx.doi.org/10.1016/j.jclepro.2019.118732.
[37] Kulcke A, Gurschler C, Spöck G, Leitner R, Kraft M. On-line classification of synthetic polymers using near infrared spectral imaging. J Near Infrared Spectrosc 2003;11(1):71–81. http://dx.doi.org/10.1255/jnirs.355.
[38] Tang B, Liu Z, Xiao X, Nie M, Chang J, Jiang W, et al. Spectral–spatial hyperspectral classification based on multi-center SAM and MRF. Opt Rev 2015;22(6):911–8. http://dx.doi.org/10.1007/s10043-015-0139-9.
[39] Galal A, Hassan H, Imam I. A novel approach for measuring hyperspectral similarity. Appl Soft Comput 2012;12(10):3115–23. http://dx.doi.org/10.1016/j.asoc.2012.06.018.
[40] Kakhani N, Mokhtarzade M. A new neuro-fuzzy-based classification approach for hyperspectral remote sensing images. J Earth Syst Sci 2019;128(2). http://dx.doi.org/10.1007/s12040-018-1054-9.
[41] Galletti C, Myint S. Land-use mapping in a mixed urban-agricultural arid landscape using object-based image analysis: A case study from Maricopa, Arizona. Remote Sens 2014;6(7):6089–110. http://dx.doi.org/10.3390/rs6076089.
[42] Zheng H, Zhou X, He J, Yao X, Cheng T, Zhu Y, et al. Early season detection of rice plants using RGB, NIR-G-B and multispectral images from unmanned aerial vehicle (UAV). Comput Electron Agric 2020;169. http://dx.doi.org/10.1016/j.compag.2020.105223.
[43] Corradini F, González N, Casado F, Rojas V, van der Ploeg M. Usefulness of an opportunistic data analysis approach to evaluate if environmental regulations aim at relevant applications. Geoderma 2019;351:261–9. http://dx.doi.org/10.1016/j.geoderma.2019.05.007.
[44] Zhang S, Yang X, Gertsen H, Peters P, Salánki T, Geissen V. A simple method for the extraction and identification of light density microplastics from soil. Sci Total Environ 2018;616–617:1056–65. http://dx.doi.org/10.1016/j.scitotenv.2017.10.213.
[45] Keranen P, Kaarna A, Toivanen PJ. Spectral similarity measures for classification in lossy compression of hyperspectral images. Proc SPIE - Int Soc Opt Eng 2002;4885:285–96. http://dx.doi.org/10.1117/12.463160.
[46] Jany B, Janas A, Krok F. Automatic microscopic image analysis by moving window local Fourier Transform and Machine Learning. Micron 2020;130. http://dx.doi.org/10.1016/j.micron.2019.102800.
[47] Signal developers. signal: Signal processing. 2014, R package v0.7-6. URL http://r-forge.r-project.org/projects/signal/.