Joint Research Programme
KWRW 2025.028 | March 2025

Microplastic Detection and EU Regulations: Infrared Spectroscopy and Machine Learning Approaches

Joint Research Programme

KWR

Bridging Science to Practice

Colophon

Microplastic Detection and EU Regulations: Infrared Spectroscopy and Machine Learning Approaches

KWRW 2025.028 | March 2025

This research is part of the Joint Research Programme of KWR, the water utilities and Vewin.

Project number
404300/023 and 402045/386

Project manager
Patrick Bäuerlein and Ina Vertommen

Client
BTO - Thematical research - Chemical Safety

Author(s)
Patrick S. Bäuerlein, Xin Tian, René van Doorn, Roan Streefland, Karel van Laarhoven

Quality Assurance
Peter van Thienen, Thomas ter Laak

Sent to
This report has been distributed to participants in the KWR Waterwijs programme.

Publicity
One year after publication of the report, this will become public.

Keywords
microplastics, machine learning, LDIR, infrared

Year of publishing
2025

More information
Dr. Patrick S. Bäuerlein
T 0306069702
E patrick.bauerlein@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

T +31 (0)30 60 69 511
E info@kwrwater.nl
I www.kwrwater.nl

Management summary

New sampling equipment and machine learning aid in identifying microplastics.

Authors Patrick S. Bäuerlein, Xin Tian, René van Doorn, Roan Streefland, Karel van Laarhoven.

The delegated act supplementing Directive (EU) 2020/2184 mandates that drinking water companies must monitor microplastics in drinking water once microplastics are added to the monitoring list. In this project, the necessary tools were developed to meet the requirements outlined in the document. A sampling device was built and tested that enables sampling in accordance with the delegated act. This device can sample microplastics ranging from 10/20 to 500 µm in a water matrix (drinking water or similar quality).
For the identification of microplastics, a software tool was developed with a machine learning algorithm that uses infrared spectra. This software replaces the supplier's software for the infrared spectrophotometer to improve accuracy. The software was tested on real samples from the Sijmons drinking water production site. The software was successfully utilized, and the results show that, on average, fewer than one microplastic particle per litre can be detected.

Sampling device compliant with the delegated act (EU 2020/2184) for the sampling of microplastics in drinking water. The left image shows the device connected to a faucet. The right figure displays the device along with the supporting meshes that hold the filters.

Importance: Practical Techniques for Monitoring Microplastics in Drinking Water Once the delegated act supplementing Directive (EU) 2020/2184 is transposed into national legislation and microplastics are added to the monitoring list, drinking water companies will be required to monitor microplastics in drinking water. They must count and identify microplastics using spectroscopic techniques. The directive specifies which microplastics must be monitored and how drinking water companies should collect and report this data.

Within this project, tools were developed and tested to meet these requirements. Additionally, a machine learning algorithm with a user interface was programmed to replace the commercial software provided by the infrared spectrophotometer supplier, aiming to improve the accuracy of the reports.

Approach: Development of Sampling Equipment and Software

The sampling equipment was designed and tested in accordance with the delegated decision. Additionally, a machine learning (ML) model was implemented in software with a user interface to easily identify the type of microplastics using infrared spectroscopy and report the data. The accuracy of the model was thoroughly tested.

Results: Software and Sampling Equipment Perform as Prescribed

Machine Learning Model

The developed ML model achieved high accuracy (>97%) and reliably identifies known types of microplastics based on their infrared spectra. The model can be easily trained with new data (e.g., additional polymers or more spectra of previously included polymers), enhancing accuracy and/or expanding the microplastic classes it can detect. Newly trained models can also be used for retrospective analysis of historical samples to search for plastics added to the model at a later time.

Sampling Equipment

The developed sampling equipment meets the requirements outlined in the delegated act. The device, constructed from stainless steel, includes chambers where filters are placed to capture microplastics. It is capable of sampling microplastics ranging from 10/20 to 500 µm in a water matrix (drinking water or similar quality). A key advantage of this setup is its flexibility: it can also collect other microplastics by simply replacing the filters. Readily available filters include 1, 50, and 100 µm, and other sizes can be ordered or custom-made as needed.

Validation

During the project, the software and sampling script were tested on real samples. Samples were collected at four locations within Vitens' distribution network (Sijmons drinking water production site), and microplastic concentrations were determined. Generally, low particle counts were found in drinking water (<1 particle/L), and particle counts were similar in outgoing water and tap water. Thus, there is no evidence that the distribution network releases microplastics.

Application: Developed Tools Can Be Utilized by Drinking Water Companies and Laboratories

The tools developed in this study can be used by water companies and other laboratories. For now, the other laboratories are limited to collecting samples, while the analysis is performed at KWR.

Report

This research is reported in report *Microplastic Detection and EU Regulations: Infrared Spectroscopy and Machine Learning Approaches* (KWRW 2025.028).

More information
Dr. Patrick S. Bäuerlein
T 0306069702
E patrick.bauerlein@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

Managementsamenvatting

Nieuwe bemonsteringsapparatuur en machine learning helpen microplastics identificeren

Auteurs Patrick S. Bäuerlein, Xin Tian, René van Doorn, Roan Streefland, Karel van Laarhoven.

Het gedelegeerde besluit ter aanvulling van Richtlijn (EU) 2020/2184 schrijft voor dat drinkwaterbedrijven microplastics in drinkwater moeten monitoren zodra microplastics zijn toegevoegd aan de toezichtlijst. In dit project zijn de benodigde instrumenten ontwikkeld om te voldoen aan de eisen die in het document worden genoemd. Er is een monsternameapparaat gebouwd en getest waarmee monsters kunnen worden genomen volgens de gedelegeerde handeling. Dit apparaat is in staat om microplastics tussen 10/20 en 500 µm te bemonsteren in een watermatrix (drinkwater of een vergelijkbare kwaliteit). Voor de identificatie van microplastics werd een softwaretool geschreven met een algoritme voor machinaal leren dat infraroodspectra gebruikt. Deze software vervangt de software van de leverancier van de infraroodspectrofotometer om de nauwkeurigheid te verhogen. De software werd getest op echte monsters van de drinkwaterproductielocatie Sijmons. De software werd met succes gebruikt. De resultaten laten zien dat er gemiddeld minder dan één microplasticdeeltje per liter kan worden gevonden.

Bemonsteringsapparaat volgens de gedelegeerde besluit (Eu 2020/2184) voor de bemonstering van microplastics in drinkwater. De linker afbeelding toont het apparaat aangesloten op een kraan. De rechterfiguur toont het apparaat en de ondersteunende netten die de filters vasthouden.

More information
Dr. Patrick S. Bäuerlein
T 0306069702
E patrick.bauerlein@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

Belang: Bruikbare technieken om microplastics in drinkwater te monitoren

Zodra de gedelegeerde besluit ter aanvulling van Richtlijn (EU) 2020/2184 wordt omgezet in nationale wetgeving en microplastics zijn toegevoegd aan de toezichtlijst worden drinkwaterbedrijven verplicht om microplastics in drinkwater te monitoren. Ze moeten microplastics gaan tellen en identificeren met spectroscopische technieken. De richtlijn beschrijft welke microplastics moeten worden gemonitord en hoe de drinkwaterbedrijven deze gegevens moeten verzamelen en rapporteren. Binnen dit project zijn tools ontwikkeld en getest om aan deze verplichting te voldoen. Daarnaast werd een machine learning-algoritme met gebruikersinterface geprogrammeerd ter vervanging van commerciële software van de leverancier van de infraroodspectrofotometer om de nauwkeurigheid van de rapportages te vergroten.

Aanpak: Ontwikkeling van bemonsteringsapparatuur en software

De bemonsteringsapparatuur werd gebouwd en getest volgens het gedelegeerde besluit. Daarnaast werd een machine learning (ML) model geïmplementeerd in software met een gebruikersinterface om eenvoudig het type microplastics middels infraroodspectroscopie te kunnen identificeren en de gegevens te rapporteren. De nauwkeurigheid van het model werd getest.

Resultaten: Software en bemonsteringsapparatuur werken zoals voorgeschreven

Machine learning model

Het ontwikkelde ML-model heeft een hoge nauwkeurigheid (>97%) en identificeert op betrouwbare wijze bekende typen microplastics op basis van hun infraroodspectra. Het model kan eenvoudig worden getraind met nieuwe gegevens (zoals extra polymeren of meer spectra van reeds opgenomen polymeren), waardoor de nauwkeurigheid wordt verbeterd en/of de microplastics-klassen worden uitgebreid. De nieuw getrainde modellen kunnen dan ook worden gebruikt voor retrospectieve analyse van historische monsters om te zoeken naar kunststoffen die later aan het model zijn toegevoegd.

Bemonsteringsapparatuur

De ontwikkelde bemonsteringsapparatuur voldoet aan de eisen in de gedelegeerde handeling. Het apparaat bestaat uit roestvrij staal en omvat kamers waarin filters worden geplaatst om de microplastics vast te houden. Het kan worden gebruikt voor het bemonsteren van microplastics tussen 10/20 en 500 µm in een watermatrix (drinkwater of een vergelijkbare kwaliteit). Het voordeel van de gebouwde opstelling is dat ook andere microplastics kunnen worden verzameld. Hiervoor hoeven alleen de filters te worden vervangen. Andere filters die direct beschikbaar zijn, zijn 1, 50 en 100 µm. Andere maten kunnen worden besteld of speciaal worden gemaakt.

Validatie

Tijdens dit project zijn de software en het script getest op echte monsters. Op vier locaties in het distributienetwerk van Vitens (drinkwaterproductielocatie Sijmons) werden monsters genomen en microplastic-concentraties bepaald. Over het algemeen werden lage deeltjesaantallen gevonden in drinkwater (< 1 deeltje/L) en waren de deeltjesaantallen in uitgaand water en water aan de kraan vergelijkbaar. Daarom is er geen bewijs dat het distributienet microplastics vrijgeeft.

Toepassing: De ontwikkelde tools kunnen worden ingezet bij drinkwaterbedrijven en -laboratoria

De instrumenten die in dit onderzoek zijn ontwikkeld, kunnen door de waterbedrijven en de andere laboratoria worden gebruikt. Voorlopig kunnen de andere laboratoria alleen monsters nemen. De analyse wordt uitgevoerd bij KWR.

Rapport

Dit onderzoek is beschreven in het rapport *Microplastic Detection and EU Regulations: Infrared Spectroscopy and Machine Learning Approaches* (KWRW 2025.028).

Meer informatie
dr. ir. Nikki van Bel
T 030-6069516
E Nikki.van.Bel@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

Contents

e

Meer informatie
dr. ir. Nikki van Bel
T 030-6069516
E Nikki.van.Bel@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

1 Introduction

1.1 General introduction

Polymers, which include microplastics, have been detected in oceans, rivers, lakes, bottled water and at considerably lower concentrations also in drinking water (Bäuerlein et al., 2022a; Bäuerlein et al., 2022b; Shruti et al., 2020). Research about the occurrence of polymers in water, especially drinking water, is highly relevant given that polymers are explicitly mentioned in regulations such as the new European Drinking Water Directive (DWD), which was adopted in December 2020 and came into force on January 12, 2021 [1]. The EU mandates that Member States had to implement the provisions of the revised DWD into national laws and regulations by January 12th, 2024. The directive aims to address growing public and scientific concern about the effects of polymers on human health, in particular if exposure occurs through ingestion of water intended for human consumption and to address the risks involved. This means that polymers will have to be monitored and risk assessments made. For this reason the number of particles in drinking water needs to be determined. Fast and reliable analysis methods and knowledge about the occurrence of polymers are hence crucial. Currently, monitoring is on hold until at least 2025 as microplastics need to be added to the watch list before the monitoring will become mandatory.

Several optical techniques exist to detect and characterize polymers, such as Fourier Transform Infrared (FTIR), laser direct infrared spectroscopy (LDIR) and Raman spectroscopy (Carruthers et al., 2022; Mintenig et al., 2020; Schymanski et al., 2021). These techniques rely on databases containing reference spectra to which acquired particle spectra are compared. Reference spectra are generally collected from pristine polymers studied in matrices under controlled conditions. However, particles detected in environmental samples are seldom pristine and have often undergone more or less extensive weathering due to environmental conditions (e.g., UV-light, oxidation) or might be covered by microorganisms (biofouling). Both weathering and biofouling will affect their physicochemical properties, and hence will influence their interaction with infrared light. Ultimately, this will translate in modified spectral profiles, which can potentially lead to misidentifications (Phan et al., 2022). Compiling a database with spectra of polymers with different degrees of weathering, biofouling and combinations thereof would be extremely time consuming and is not a viable option. Therefore, alternative approaches combining more rapid analytical techniques, as well as more advanced data analysis approaches, which take advantage of existing data from both pristine and weathered samples, are necessary.

The application of LDIR is gaining ground as an alternative for the analysis of polymers in environmental samples. Compared to approaches such as FITR, it has the advantage of scanning the surface area, that contains the sample, with a single frequency prior to initiate a full frequency scan. As a full frequency sweep is time-consuming, this will safe time as now only area's with actual particles will undergo a full frequency scan. This has obvious advantages in terms of analysis time, as spectra will only be acquired when particles have been detected during the preliminary scan. Samples that are almost void of particles are therefore measured considerably faster with LDIR. However, the infrared band recorded with LDIR instruments is narrower (ca. 1800 to 900 cm^{-1}) than FITR (4000 to 500 cm^{-1}). As less information is collected of a particular particle, LDIR is likely more prone to misidentification, especially when analyzing weathered particles compared to FTIR. Therefore, more effective classification methods are required to minimize the risk of misidentification. Machine learning (ML) can be a way to obtain more reliable classification approaches; yet few examples of their implementation in the field of polymer identification exist. In earlier studies, we developed an effective approach to identify environmentally exposed polymers using a combination of LDIR and

[1] https://ec.europa.eu/transparency/documents-register/detail?ref=C(2024)1459&lang=en&lang=en

ML (Tian et al., 2022; Tian et al., 2023). This model is further refined by increasing the training data set. Additionally, the user can now decide that based on a match score (a score that says how certain the model is with its prediction) if the prediction is accepted. This refined model has been used for the identification of polymers according to the delegated act.

The objective of this project is to improve the previously developed machine learning model (subspace-knn) and implement these in a user interface, to make it more accessible for users less familiar with machine learning. The new models accuracy has been determined and the user interface can be run on any windows computer. Furthermore, the requirements for monitoring microplastics mentioned in an delegated act supplementing Directive (EU) No 2020/2184 are mentioned. Based on these requirements a sampling device has been built and tested as well as a script has been written that allows for reporting the microplastic data as mentioned in the delegated act. The machine learning model and the script were applied to real samples from the drinking water production site Sijmons (Vitens).

1.2 Delegated Act – Important parameters

Article 13(6) of Directive (EU) No 2020/2184 of the European Parliament and of the Council of 16 December 2020 deals with the quality of water intended for human consumption. The accompanying delegated act adds monitoring of microplastics to the list of compounds of interest. In this document the technical requirement for monitoring of microplastics in drinking water are written down (see Table 1). The most important parameters regarding microplastic monitoring are mentioned in this section. The full delegated act (Dutch version) can be found in the supporting information and online[2]. The main purpose of the delegated act is to ensure that data is gathered and reported across all laboratories in the same way. This will enable comparability, which up to this day is often not given as no generally accepted ways of sampling, measuring and reporting have been agreed on.

Table 1: Table containing the parameters of polymers that need to be measured and how they need to be reported according to Article 13(6) of Directive (EU) No 2020/2184 of the European Parliament and of the Council (+ reference).

Parameter	Details
Polymers	Polyethylene (PE), Polypropylene (PP), Polyethylene terephthalate (PET), Polystyrene (PS), Polyvinyl chloride (PVC), Polyamide (PA), Polyurethane (PU), Polymethylmethacrylate (PMMA), Polytetrafluorethylene (PTFE), Polycarbonate (PC)
Size Range	20 µm to 5 mm for particles; 20 µm to 15 mm for fibres
Fibres	A particle is a fibre if the ratio of length and width is 3 or larger. The size of a fibre is defined by the average value of the projected width of the fibre.
Particles	A particle is considered to be not a fibre if the ratio between length and width is < 3. The size of a particle is defined by the surface-equivalent diameter. d = sqrt(4A/π), where A is the area of the two-dimensional projection of the particle.

[2] https://ec.europa.eu/transparency/documents-register/detail?ref=C(2024)1459&lang=en&lang=en

Size Brackets	20 ≤ surface-equivalent diameter < 50 µm 50 ≤ surface-equivalent diameter < 100 µm 100 ≤ surface-equivalent diameter < 300 µm 300 ≤ surface-equivalent diameter < 1,000 µm 1,000 ≤ surface-equivalent diameter < 5,000 µm
Sampling Device	Cascade of four filters: two primary filters (a: 100 µm, b: < 20 µm) for collecting suspended particles; two secondary filters (c: 100 µm, d: < 20 µm) for procedure blanks. (see Figure 1).
Filter Size	Primary Filters: 100 µm and < 20 µm Procedure Blanks: 100 µm and < 20 µm
Detection Techniques	Vibrational spectroscopy (µ-FTIR, µ-Raman, QCL-IR), Optical microscopy for size and shape determination
Sample Volume	Minimum 1,000 litres of water
Data Reporting Categories	Shape: Particles or fibres Size: Specific size classes as defined above Composition: Priority polymers, non-priority synthetic or chemically modified natural polymers, others or unidentified
Additional Information	Total volume of sampled water Location and time of sampling and analysis Sample handling details Spectroscopic method and instrument used Subsampling details Any deviation from the methodology

Important Considerations:

- Minimise Contamination: All reasonable precautions must be taken to prevent contamination from external plastic particles, personal protective equipment, or laboratory equipment. All liquids must be filtered through 0.45 µm prior to use.
- Procedure Blanks: The water for the procedure blanks is the same as for the real sample. Use filters c and d to obtain procedure blanks for assessing contamination from laboratory equipment, reagents, and ambient air during sampling, handling, and analysis. These procedure blanks are subjected to the same processing and analysis steps as the corresponding collection filters a and b. At least ten procedural blanks should be taken. Regular collection and analysis of further procedure blanks are required to monitor variations in background contamination levels. If the contamination level of a regularly collected blank exceeds the average background contamination (µ) by more than three times the standard deviation (σ), the laboratory must investigate the source and take measures to reduce it.
  Maximum acceptable contamination is therefor: level=$\mu+3\sigma$

> Example:
>
> The blanks' average is 10 with standard deviation of 3.
>
> Maximum acceptable contamination level=10+(3×3)=10+9=19.

- Reanalysis: If the sample cannot be analysed directly on the collection filter, the particle material can be resuspended and transferred to another carrier for analysis.
- Subsampling: When material from filters a or b is transferred to another carrier for analysis (secondary filter or another suitable surface), it should preferably be done without subsampling. If the analysis procedure involves subsampling, the final analysed sample must constitute at least 10% of the material recovered from the original sampled water volume. The materials collected on filters a and b must be analysed separately. If a sample contains too many particles for a reasonable analysis due to time constraints, at least 20% of the with particles covered area needs to be analysed. The total number of particles on the slide should then be extrapolated.
- Documentation: Record additional information for each sample, including the total sampled volume, sampling location and time, sample handling details, the spectroscopic method and instrument used, and any deviations from the methodology.
- Quality Assurance: Conduct experimental verifications to assess material recovery on each filter when applying the methodology. The recovery rate should be between 100% ± 40%. For more information see below.

Figure 1: Cascade of four filters: two primary filters (a: 100 µm, b: < 20 µm) for collecting suspended particles; two secondary filters (c: 100 µm, d: < 20 µm) for procedure blanks. Materials used for this setup can be found in the supplementary information.

Recovery

Experimental verifications must be conducted to assess the recovery of material on each of the filters (a and b) when applying the methodology. This involves:

- Adding Enriched Samples: Add enriched samples with a known amount of clearly identifiable microplastics to the water stream flowing through the cascade of filters. Then, check the amount recovered at the end of the analysis procedure.

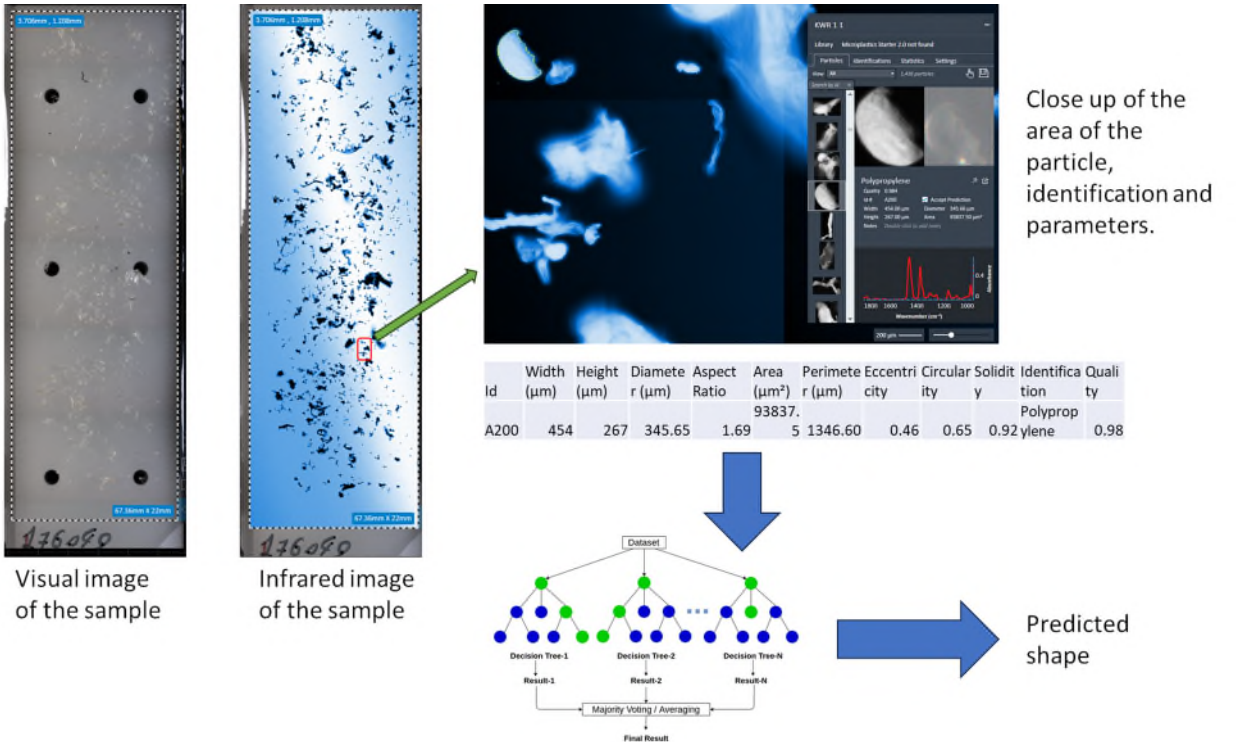
- Sample Characteristics: The enriched samples should contain particles with suitable size, density, and quantities for assessing recovery on filters a and b.

- For filter a: use particles sized 120 to 200 µm.
- For filter b: use particles sized 30 to 70 µm.
- Polymer Types: Use at least two of the priority polymers for the enriched samples. Ensure that:
- At least one polymer has a higher density than water (e.g., PET).
- At least one polymer has a lower density than water (e.g., PE).
- Number of Particles: The number of particles added to the enriched samples should be between 50 and 150.

Acceptable Recovery Rate: The analysis procedure is considered acceptable if the recovery rate is between 100% ± 40%.

1.3 Laser direct infrared spectroscopy

Laser Direct Infrared (LDIR) Imaging is a technique used for chemical imaging and analysis, particularly effective in identifying and characterizing materials based on their chemical composition. This technique makes use of a quantum cascade laser to acquire mid-range infrared spectra (1800 – 975 cm^{-1}). The laser is directed towards the particle, and the interaction between the laser light and the material creates an infrared spectrum of the particle. The data of the image can be analysed to reveal its characteristics. The technique is non-destructive, meaning samples can be analysed multiple times and can subsequently even be analysed by other destructive methods such as pyrolysis or TGA-GC/MS. The LDIR workflow looks as follows: After a chemical work-up to remove as much interfering organic and inorganic material as possible, the sample, fixated on a mirror slide, is placed under the laser. First a visual image of slide is taken. This is followed by a scan of the slide for infrared activity. This information is used to determine where particles are situated on the slide. In the next step a full infrared spectrum is recorded for each detected particle. Based on this spectrum the (polymer composition of the) particle is identified if similar spectra are available in the digital library of the software. Various attributes of the particle, such as circularity and solidity are determined in addition to the infrared spectrum. These are used eventually to determine the particle shape with a help of a random forest model. Additionally to determining the shape using these parameters, also the ratio of width and length are being using in accordance with the delegated act. Figure 2 shows the workflow:



Id	Width (µm)	Height (µm)	Diameter (µm)	Aspect Ratio	Area (µm²)	Perimeter (µm)	Eccentricity	Circularity	Solidity	Identification	Quality
A200	454	267	345.65	1.69	93837.5	1346.60	0.46	0.65	0.92	Polypropylene	0.98

Figure 2: Workflow of the LDIR: First a visual and an infrared picture are taken of the sample. Next each individual particle will be analysed using the infrared laser. Based on the picture also physical parameters such as width and height are determined. These will be used in a random forest model to determine the shape. In this example here Kevley slides are shown. Those can now be replaced with gold filters, which simplify the analysis.

LDIR can detect particles down to the size of 3 µm. However, this is the smallest detectable size that does not allow accurate determination of composition and shape. Realistically, 10 µm is therefore the lower size limit. The upper size limit is about 500 µm with regard to the hight of the particle. If particles are too large focusing the camera and laser is difficult. This is based on experience and advice by the manufacturer. There is no technical limit with regard to the length or width of the particle. The following table shows the information that can be gathered using this technique. The advantage of LDIR method over the more commonly used micro-FTIR method is that LDIR is really fast when only few particles are detected on the slide. This is due to the fact that only a full spectra analysis is executed when particles are actually detected during the screening of the sample. Table 2 shows the most important characteristics of the LDIR.

Table 2: Characteristics of the LDIR

Characteristic	
Size range	10 – 500 µm
Infrared range (wave number)	1800 – 975 cm^{-1}
Step size (wave number)	0.5 cm^{-1}
Shape prediction of particle possible	Yes
Size of each particle	Yes
Colour measurement	No
Identification of polymer/particle	Yes (by supplier as well as KWR software)
Type of information	Number of particles per sample

1.4 Brief introduction to Machine learning

Machine learning (ML), a subset of artificial intelligence, focuses on the development of algorithms that are capable of learning from and making predictions or decisions based on data. This field of ML aims to identify patterns and/or trends to extract insights from labeled or unlabeled datasets, thereby improving their performance on specific tasks, such as making predictions based on new datasets. Machine learning is typically categorized into supervised learning, unsupervised learning, and reinforcement learning, each addressing distinct types of problems and applications. It has largely impacted various sectors, including water, healthcare, and marketing.

Supervised classification, which is used in this research to classify different types of polymers, is a prominent type of ML wherein the algorithm is trained on a labeled dataset, meaning each training example (also referred to as sample or observation in the field of ML) is paired with an output label (e.g., PVC, PA, or PMMA). The primary objective is to train the model by learning the relationship between input (also commonly referred to as features) and output labels, enabling it to predict the label for new, unseen data accurately. This process is divided into three main phases: training, validation, and testing. For this process one data set is used. This is randomly divided into subsets to ensure that training and validation is not done with the same data. During the training phase, the model is trained with a subset of the data, learning the mapping from features to labels. In the subsequent validation and testing phases, the model's performance is evaluated using a separate subset of the same data to ensure the accuracy and generalizability of its predictions. Supervised classification is extensively applied in many research fields where classification is needed.

More specifically, ensemble learning, which is used in this research to improve the performance of classification models, is a relatively sophisticated ML technique that optimally integrates the predictions of multiple models to achieve an improved performance compared to any single ML model (referred to as a weak learner). This approach capitalizes on the strengths of various models while mitigating their weaknesses, resulting in more reliable and accurate predictions. Several ensemble methods exist, including bagging, boosting, and stacking. Bagging entails training multiple weak learners independently and averaging their predictions, whereas boosting involves training models sequentially, with each new model focusing on the errors of the preceding ones. Stacking combines different types of models and utilizes another model to determine the optimal combination of their predictions.

Ensemble learning is particularly effective in enhancing the accuracy and stability of ML models, making it a widely adopted technique in a large number of research-wise applications.

2 The Machine Learning models

2.1 Data preparation

Prior to introducing the ML models, it is important to specify the input and output (see Table 3). In this study, 1,650 features are used as input, which corresponds to absorption rates on wave numbers ranging from 975 to 1800 cm^{-1} with a resolution of 0.5 cm^{-1}. For the model to work it is essential that the model data and the data to be used are formatted in the same way (file format, number of columns, wave number range, etc.). Additionally, the following steps for training data preparation have been done and are crucial:

1. Per class and subclass the dimensionality (wave number entries) of the data was reduced and then the data was clustered with the t-SNE[3] method. The type of polymer is defined at the class level, while the subclass level specifies the different types within the same polymer category, e.g. polystyrene is the class and expanded polystyrene and oxidised polystyrene are the subclasses. At this stage outliers were manually removed from the data (Figure 3). Outliers can be falsely labelled data entries, qualitatively bad measurements or corrupted data. As each data point is labelled, extreme outliers were manually checked (spectra were compared).

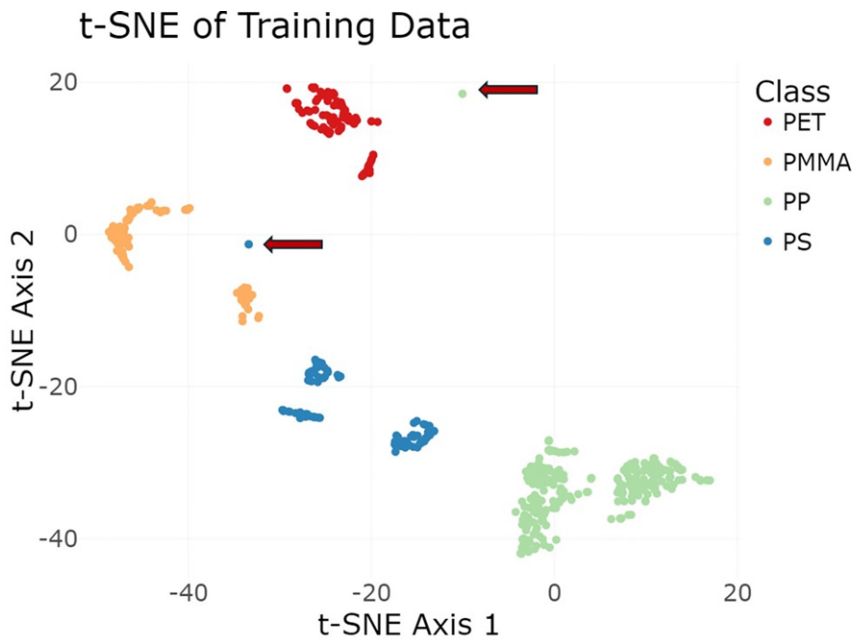


Figure 3: Example of outlier removal per class shown on a fraction of the data. The data is plotted in a 2-dimensional coordinate system. Outliers are manually identified and removed from the data set. The two outliers from PS and PP are removed from the data set. The figure shows the original data set with the outliers.

[3] t-distributed stochastic neighbour embedding (t-SNE) a statistical method for visualizing high-dimensional data by giving each datapoint a location in a two dimensional map.

2. If the actual maximum adsorption is too close to the baseline (in our case <0.05), the spectra were removed. This value is based on expert knowledge.

3. Each spectrum is normalised meaning all spectra have a maximum adsorption of 1 after normalisation. This is necessary as particle thickness has an impact on the adsorption but must not play a role in particle identification.

Step 2 and 3 must also be done for the actual data and are automated. Step 1 is only important for the data training the model, meaning the most time-consuming step has only be done when new training data is added to a model.

2.2 Testing models

2.2.1 Selected ML algorithms

The following algorithms were tested and compared in this study. Below their brief introduction gives readers an overview.

Decision Trees

Decision trees are a type of learning algorithm used for classification and regression tasks. They function by recursively splitting the data into subsets based on the value of input features, creating a tree-like structure of decisions. Each internal node represents a test on an attribute, each branch represents the outcome of the test, and each leaf node represents a class label or a regression value. Decision trees are simple to understand and interpret, and they can handle both numerical and categorical data. However, they can be prone to overfitting[4], which is often mitigated by techniques such as pruning [5].

Discriminant Analysis

Discriminant analysis is a statistical technique used for classifying a set of observations into predefined classes. It works by modelling the difference between classes using linear or quadratic combinations of the input features. The most common forms are Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA). LDA assumes that the different classes generate data based on Gaussian distributions with the same covariance matrix, whereas QDA allows each class to have its own covariance matrix. Discriminant analysis is particularly useful when the assumptions about normality and equal variance hold true, providing clear decision boundaries and efficient classification.

Support Vector Machine (SVM)

Support Vector Machines (SVM) are learning algorithms used for classification and regression tasks. SVMs work by finding the hyperplane that best separates the data into different classes with the maximum margin. The algorithm relies on a subset of the training data, known as support vectors, which are the data points that lie closest to the decision boundary. SVMs are effective in high-dimensional spaces and can be adapted to non-linear classification

[4] Overfitting occurs when a machine learning model learns the training data too well, including its noise and outliers, to the extent that it performs poorly on new, unseen data.

[5] Pruning is a technique used to reduce the complexity of a machine learning model by removing parts of the model that are deemed unnecessary or redundant.

using different kernel functions[6]. They are robust to overfitting, especially in high-dimensional settings, and are widely used in applications such as image recognition and text categorization.

K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) is a simple, instance-based learning algorithm used for classification and regression. In KNN, the classification of a sample is determined by the majority class among its K nearest neighbors in the feature space, where K is a user-defined parameter. Similarly, in regression, the predicted value is the average of the values of its K nearest neighbors. KNN is non-parametric, meaning it makes no assumptions about the underlying data distribution. It is easy to implement and understand but can be computationally intensive with large datasets and sensitive to the choice of K and the distance metric.

Neural Networks (NN)

Neural Networks (NN) are a set of algorithms inspired by the human brain, designed to recognize patterns. They consist of layers of interconnected nodes (neurons), where each connection represents a weighted link between neurons. Neural networks can model complex relationships between inputs and outputs and can learn from data in a supervised or unsupervised manner. They are highly flexible and can approximate any continuous function given enough data and computational power. Neural networks are particularly powerful in tasks involving large amounts of data, such as image and speech recognition, where they have achieved state-of-the-art performance in large-scaled applications.

Ensemble Algorithms

In addition to the above-mentioned ML algorithms (except NN), ensemble algorithms are adopted to enhance combine the predictions of multiple machine learning models to produce a better prediction. The main idea is that by aggregating the predictions of various models, the ensemble can reduce variance (bagging), bias (boosting), or improve predictions through a combination of diverse models (stacking). Common ensemble methods include Bagging (e.g., Subspace methods, Random Forests), Boosting (e.g., AdaBoost, Gradient Boosting), and Stacking.

2.2.2 Which model was ultimately chosen?

The No Free Lunch (NFL) Theorem is a fundamental principle in the field of optimization and ML which states that no single algorithm is universally better than others when averaged over all possible problems. This theorem implies that every algorithm has its strengths and weaknesses and can perform well on certain types of problems while performing poorly on others. In other words, the performance of any optimization or learning algorithm depends heavily on the specific problem being addressed. The NFL theorem underscores the importance of tailoring algorithm choice and design to the specific characteristics of the problem at hand, rather than relying on a one-size-fits-all approach.

By comparing the performances of all the candidate models, the Subspace KNN algorithm shows the best performance in our application of polymer classification. Subspace KNN is an ensemble extension of the traditional KNN method, designed to improve performance by considering a considerable number of subsets of the feature space (i.e., wave numbers). Instead of using all available features, subspace KNN selects or constructs multiple subsets (subspaces) of features, performs the KNN classification or regression within these subspaces, and opts for the optimal subset. By focusing on more relevant or informative feature subsets, subspace KNN can achieve better accuracy and computational efficiency, especially in high-dimensional datasets (which is also our case). This

technique becomes particularly useful in our application because the relationship between features and the target variable is complex, where different subsets of features might provide complementary information.

2.2.3 Model output

The model categorizes microplastics by class and subclass as outlined in Table 3. Each type of plastic or polymer in the database is assigned both a class and a subclass; if a subclass does not exist, it defaults to the class itself. The model is trained to recognize these categories based on the distinct and overlapping characteristics of polymers, such as those in the subclasses of polypropylene (e.g., PP-RCT, PP-R), where polypropylene serves as the class, and PP-RCT and PP-R are the subclasses.

Although subclasses often have similar infrared spectra, which can occasionally result in misclassifications between them as shown for PP in Figure 4, it is important to retain subclass distinctions for accurate identification. Typically, misclassifications at the subclass level do not hinder the accurate classification of the polymer at the class level. However, in cases where subclasses significantly differ—as in the example of polystyrene, where expanded polystyrene (Styrofoam) and solid polystyrene have markedly different infrared spectra—the creation of subclasses is crucial for precise identification (Figure 4).

Table 3: Example of the database file for the model. The table shows that polymers are labelled on class and subclass level. More information can be added to the column details. The last two columns are the first columns of the infrared data. For the sake of clarity, the other columns are not shown.

ID	ID2	Date added	Class	Subclass	Details	Type	Source	Data file (csv)	975	975.5
KWR 1 1 A74 (Absorbance)	944	20/03/2024	Nylon	Nylon6	Nylon6	polymer	Hawai Pacific	nylon6	0.10377	0.10220
(Absorbance)…87	1020	20/03/2024	Nylon	Nylon66	Nylon66	polymer	Hawai Pacific	nylon66	0.21247	0.20166
KWR A13 (Absorbance)…1563	1446	28/11/2023	PP	PP-R	PP-R	polymer	TEPPFA	TEPPFA - PPR	1.17486	1.10969
KWR A41 (Absorbance)	1481	28/11/2023	PP	PP-RCT	PP-RCT	polymer	TEPPFA	TEPPFA - PPRCT	0.98155	0.93659
(Absorbance)…122	1568	28/11/2023	PS	expanded PS	eps A	polymer	Hawai Pacific	eps	0.02567	0.02867
KWR combi 1 A5 (Absorbance)…925	1651	28/11/2023	PS	PS	PS	polymer	BAM	BAM - PS	0.29860	0.30432

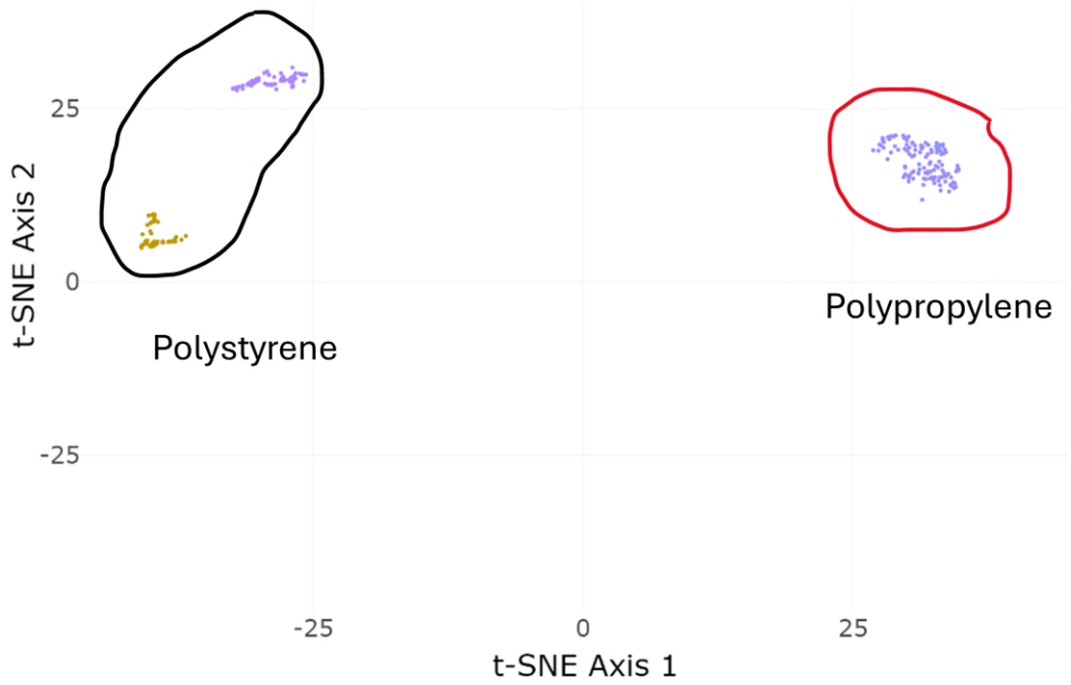


Figure 4: Graphical representation of the similarity of the two types of polystyrene (expanded polystyrene in purple and solid polystyrene in brown) and polypropylene (PP-R and PP-RCT) in a t-SNE coordinate system with the first two representative features. The two polystyrenes can be clearly differentiated, whereas the polypropylene types are indistinguishable.

2.3 Selected Model

To evaluate model performance, four key values are computed first. For each class (i.e., each type of polymer), True Positives (TP) represent the instances where the model correctly predicts the presence of a specific polymer. True Negatives (TN) occur when the model correctly identifies that a certain polymer is not present in the sample. False Positives (FP) are instances where the model incorrectly predicts the presence of a polymer that is not actually present, while False Negatives (FN) are cases where the model fails to detect a polymer that is present in the sample. These four indicators are fundamental in evaluating classification models, as they form the basis for key metrics such as accuracy, precision, and recall, which help in understanding the model's strengths and weaknesses in different scenarios. Based on TP, TN, FP, and FN, the following performance metrics are adopted:

Accuracy

Accuracy is a metric used to evaluate the overall correctness of a classification model. It is defined as the ratio of correctly predicted instances to the total number of instances in the dataset. In other words, accuracy measures how often the model's predictions match the actual labels. It is calculated as follows:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

While accuracy is a useful measure in many cases, it can be misleading in situations where the data is imbalanced, meaning that the number of instances in different classes is not evenly distributed. For example, in a dataset where 95% of the instances belong to one class, a model that always predicts that majority class would have high accuracy but would fail to correctly identify the minority class.

Precision

Precision, also known as positive predictive value, is a measure of the accuracy of the positive predictions made by the model. It is the ratio of true positive predictions to the total number of positive predictions (both true positives and false positives). Precision answers the question: "Of all the instances that were predicted as positive, how many were actually positive?" It is calculated as follows:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

High precision indicates that the model makes few false positive errors, meaning that when it predicts a positive class, it is usually correct. Precision is particularly important in scenarios where the cost of false positives is high, such as in spam detection or medical diagnosis.

Recall

Recall, also known as sensitivity or true positive rate, measures the ability of the model to correctly identify all positive instances in the dataset. It is the ratio of true positive predictions to the total number of actual positive instances (both true positives and false negatives). Recall answers the question: "Of all the actual positive instances, how many were correctly predicted as positive?" It is calculated as follows:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

High recall indicates that the model makes few false negative errors, meaning that it successfully identifies most of the positive instances. Recall is crucial in situations where missing a positive instance is costly, such as in disease screening or fraud detection.

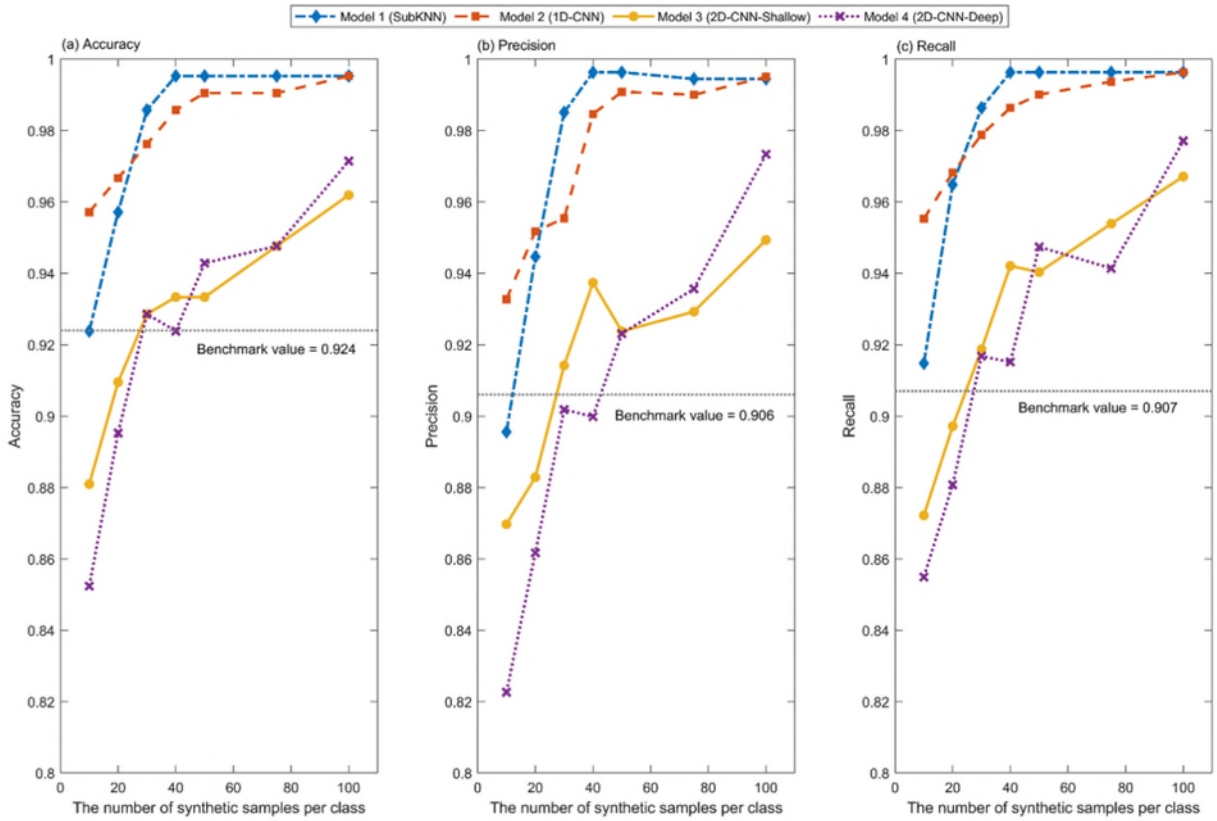


Figure 5: Performance parameters of the different models (Tian et al., 2023).

Based on the performance parameters the subknn model was chosen. In Figure 6 an example of a confusion matrix of the chosen subknn model is shown. The figure shows how often the model predicted the actual class correctly or incorrectly. In this case the model predicted 99% of the microplastics correctly.

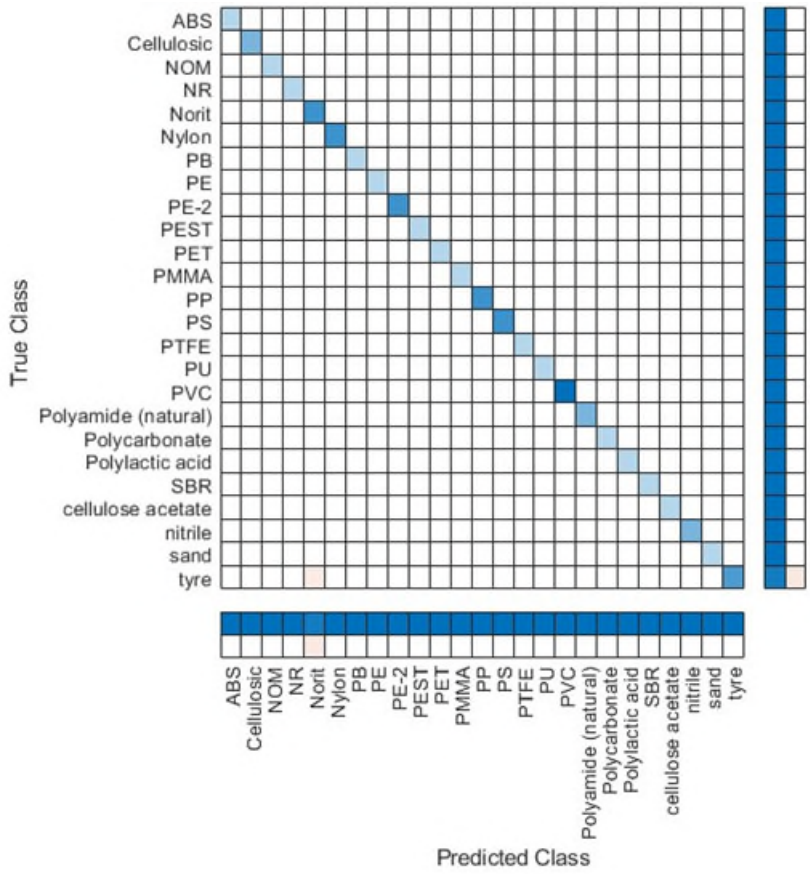


Figure 6: Confusion matrix of the subknn model. The predicted class and the true class are shown. The figure shows how often the true class was correctly and wrongly predicted by the model. The blue colour indicates a correctly identified polymer and the red colour indicates a wrongly labelled polymer.

3 Using the model on real samples

3.1 Introduction

To test the model and also to get insight into the possible release of microplastics from polymer pipes in the drinking water distribution network, four samples were taken at different locations in the network of the production plant Sijmons in Arnhem (Vitens) (Figure 7). The samples in the distribution network were taken in Doetinchem. The main pipe materials between the sampling points are PVC and PE. For the sake of completeness, it is pointed out that rubber is present as well in the sealing rings of the joints between pipes.

Samples were taken using candle filters (details in Paragraph 3.2). The above-mentioned EU sampling device could not be used at the time of field sampling as the specifications for this device where yet published by EU at the time of the sampling.

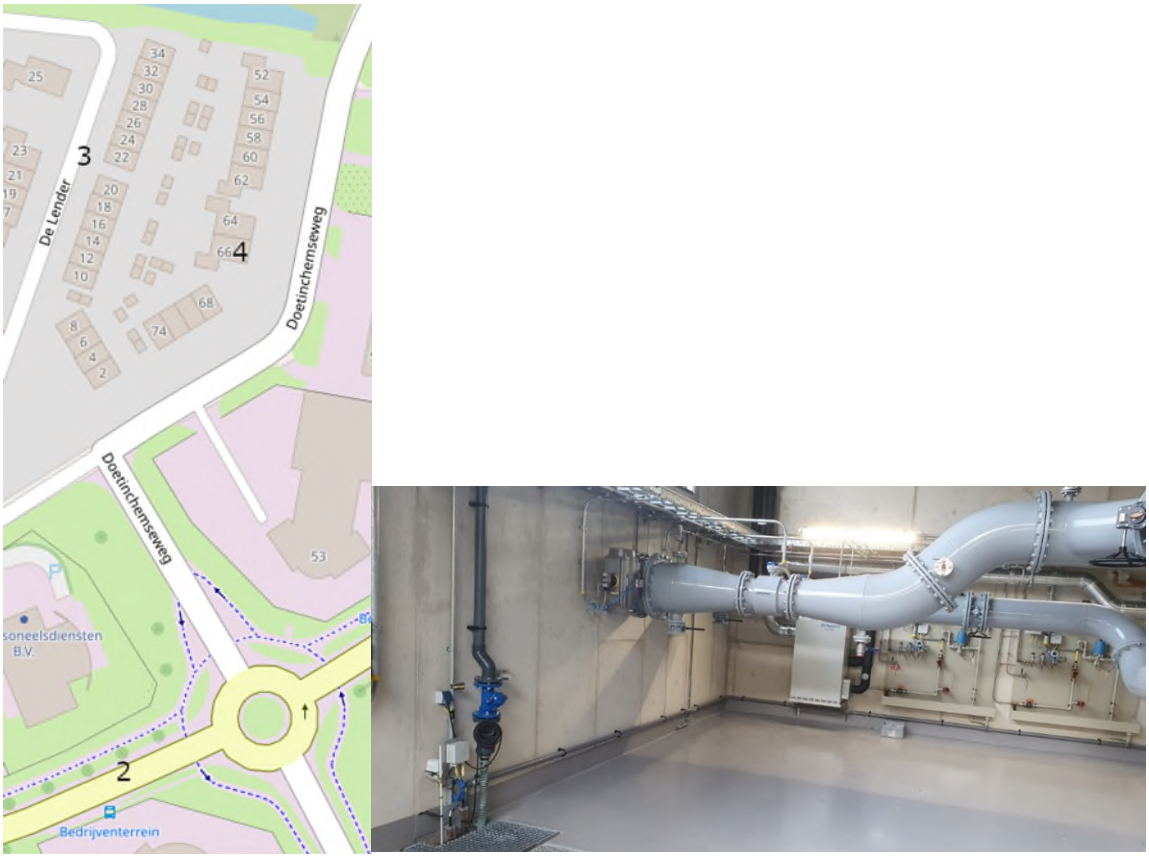


Figure 7: Map with the location of the three sampling points in Doetinchem. Postcode area 7007 (left). Sampling point two was just before the water enters a newly built neighbourhood, sampling point three inside the new neighbourhood and number four in a house. Sampling location one, drinking water production site Sijmons in Arnhem (rechts).

3.2 Material and method

3.2.1 Sample treatment

Each sample underwent an initial cleaning process involving ultrapure water and ethanol rinses in three separate beakers. For a more thorough cleaning, each sample was ultrasonicated in ultrapure water for 5 minutes in an Erlenmeyer flask, followed by ethanol rinses in three beakers, consecutively. The particulate matter was filtered using a 1 µm 47 mm stainless steel (RVS) filter and subsequently rinsed with ultrapure water. The filters containing particles were subjected to a solution of approximately 100 mL of 30% hydrogen peroxide. The suspension was filtered through a 1 µm 47 mm RVS filter and rinsed with ultrapure water. Due to the high particulate content observed, a brief 10-minute exposure to an 18% hydrochloric acid (HCl) solution was performed to dissolve rust particles. After the acid-treatment, the filters were rinsed again with ultrapure water. Samples were transferred to 25 mm 1 µm RVS filters, ultrasonicated, and rinsed with an excess of ultrapure water. The small RVS filters were placed in beakers containing approximately 500 µL of ethanol, ultrasonicated briefly, and then rinsed with about 500 µL of ethanol. Microscopic examination revealed no particles on the RVS filters. For further analysis, samples were transferred to Kevley slides using a Finn-pipet with a cut tip, ensuring the sample was entirely deposited. The slides were then heated to 40°C to evaporate the ethanol.

3.2.2 Recovery

150 polyethylene particles (ca 100 µm) were treated the same way as the real samples. Of these 150 particles 118 were recovered, which means that the recovery rate is 78.7%. This means that the number of particles in the

samples will be adjusted by dividing by 0.79. This is within the acceptable range of 100% ± 40% mentioned in the delegated act.

3.2.3 LDIR measurement

The Kevley slides were placed in the LDIR and particles between 10 – 500 µm were measured. In first instance, the total particle number was determined and then a subsample was measured. This was done as the total particle number was too high for the software to handle. In section 3.4 the statistical evaluation of this choice is described.

3.3 Results and Discussion

3.3.1 Quality criteria

In this section the quality criteria for acceptation particle numbers in the four samples will be defined. Particles and fibre numbers are shown if the counted number of particles in the sample is at least three times the number in the blank or if no particles or fibres were found in the blank. The supporting information shows which particle numbers are excluded based on the above mentioned criteria (SI I). This threshold was chosen as there were insufficient blank samples to statistically determine particle numbers in samples deviated from the blanks. Three times higher than the control sample was chosen as the criterion for quantification, as it is a generally accepted standard in analytical chemistry. Also in the supporting information a list of all the classes of compounds that can be identified is shown. The list also contains the subclasses for each class. E.g. Nylon-6,6 is a subclass of the class Nylon.

3.3.2 All polymers

Table 4 and Table 5 show the particle concentrations for all polymers that can be identified. In total 10 different polymers were found, namely PE, PEST, PET, PMMA, Polyamide, PP, PS, PU, PVC and rubber. The total particle concentration is about 800 particles per cubic metre for particles larger than 10 µm. For particles larger than 20 µm it is about 550 particles per cubic metre. That latter number can be compared to earlier research with the same size cut off. The numbers in this research are similar to earlier findings in Dutch drinking water (200 – 1500 particles per cubic metre) (Bäuerlein et al., 2022b). Recent research from the UK found particle concentrations in drinking water samples (i.e. on average 25 particles per cubic metre for particles larger than 25 µm) (Adediran et al., 2024). This translates to about 0.5 p/L and 0.025 p/L, respectively. The slightly higher concentrations in this study can be consequence of different lower size limits. This research counted particles from 10 µm upwards, the study in the UK from 25 µm upwards. In comparison bottled water can contain between ca. 0.009 to 4 p/L (Adediran et al., 2024).

Of the ten polymers mentioned in the delegated act PC and PTFE were not found. The other eight were detected. However, in case of PET e.g. only one particle per cubic meter is found in one of the samples. The highest concentrations were found for PVC and PEST.

Table 4: All identified polymers found at the four locations. Part 1/2.

Production plant	**Location**	**1**					
Size (µm)	PE	PEST	PET	PMMA	Polyamide	Polycarbonate	Polylactic acid
10-20	3	16	0	7	0	0	0
20-50	3	29	0	10	0	0	0
50-100	5	5	0	0	0	0	0
100-300	0	0	0	0	0	0	0
300-1000	0	0	0	0	0	0	0
1000-5000	0	0	0	0	0	0	0
Rotonde	Location	2					
	PE	PEST	PET	PMMA	Polyamide	Polycarbonate	Polylactic acid
10-20	0	0	0	5	10	0	0
20-50	11	14	0	8	12	0	0
50-100	2	6	0	4	0	0	0
100-300	0	0	0	0	0	0	0
300-1000	0	0	0	0	0	0	0
1000-5000	0	0	0	0	0	0	0
Woonwijk	Location	3					
Size (µm)	PE	PEST	PET	PMMA	Polyamide	Polycarbonate	Polylactic acid
10-20	0	16	0	29	0	0	0
20-50	0	22	0	47	0	0	0
50-100	2	16	0	9	0	0	0
100-300	0	11	0	0	0	0	0
300-1000	0	0	0	0	0	0	0
1000-5000	0	0	0	0	0	0	0
Woonhuis	Location	4					
Size (µm)	PE	PEST	PET	PMMA	Polyamide	Polycarbonate	Polylactic acid
10-20	0	0	1	31	0	0	0
20-50	0	16	0	19	0	0	0
50-100	0	3	0	8	0	0	0
100-300	0	1	0	0	0	0	0
300-1000	0	0	0	0	0	0	0
1000-5000	0	0	0	0	0	0	0

Table 5: All identified polymers found at the four locations. Part 2/2.

Production plant	**Location**	**1**					
Size (µm)	Polyoxymethylene	PP	PS	PTFE	PU	PVC	rubber
10-20	0	0	0	0	0	5	155
20-50	0	0	0	0	8	49	440
50-100	0	3	0	0	0	5	59
100-300	0	0	0	0	0	0	8
300-1000	0	0	0	0	0	0	0
1000-5000	0	0	0	0	0	0	0
Rotonde	Location	2					
Size (µm)	Polyoxymethylene	PP	PS	PTFE	PU	PVC	rubber
10-20	0	0	0	0	0	3	125
20-50	0	5	0	0	4	12	141
50-100	0	6	0	0	0	3	18
100-300	0	0	1	0	0	0	1
300-1000	0	0	0	0	0	0	0
1000-5000	0	0	0	0	0	0	0
Woonwijk	Location	3					
Size (µm)	Polyoxymethylene	PP	PS	PTFE	PU	PVC	rubber
10-20	0	0	0	0	2	47	203
20-50	0	4	0	0	0	82	161
50-100	0	13	0	0	0	51	49
100-300	0	0	7	0	0	2	4
300-1000	0	0	0	0	0	0	0
1000-5000	0	0	0	0	0	0	0
Woonhuis	Location	4					
Size (µm)	Polyoxymethylene	PP	PS	PTFE	PU	PVC	rubber
10-20	0	0	0	0	1	0	153
20-50	0	9	0	0	6	13	133
50-100	0	5	0	0	0	9	5
100-300	0	0	0	0	0	0	1
300-1000	0	0	0	0	0	0	0
1000-5000	0	0	0	0	0	0	0

3.3.3 Polymers from pipe materials

PE, PVC and rubber might be expected to be released from the pipe system, as these are the main components used in the distribution network. The pipe system in this network mainly constitutes of PVC.

The particle and fibre concentrations for the three polymers are shown in Table 6 and Table 7. The data indicate that particle numbers are highest in the water leaving the production plant and in the water from 'Woonwijk.' However, as only one sample per location was taken no statistical analysis is possible to determine the significance of these differences, and no definitive conclusions can be drawn other than confirming the previously stated low

concentrations. Consequently, the question of whether polymers are released from polymer pipes remains unanswered, technically, but it can be said that if particles are indeed released, their numbers are minimal and would not significantly impact the quantities already present when the water leaves the production facility. To put the numbers in perspective, approximately 300 rubber particles per cubic meter are found in the tap water at location 4, meaning a person consuming about 2 litres of this water per day could ingest roughly 0.6 particles daily. Fibres are almost absent in the four samples. A few single fibres have been found. The highest number was six per cubic meter for PVC at Location 3.

Table 6: Polypropylene (PP), Polyvinyl chloride (PVC) and different rubbers particle concentrations per location in particles per m^3, Size = surface-equivalent diameter.

Production plant	Location	1		Rotonde	Location	2	
Size (µm)	PP	PVC	rubber	Size (µm)	PP	PVC	rubber
10-20	0	5	155	10-20	0	3	125
20-50	0	49	440	20-50	5	12	141
50-100	3	5	59	50-100	6	3	18
100-300	0	0	8	100-300	0	0	1
300-1000	0	0	0	300-1000	0	0	0
1000-5000	0	0	0	1000-5000	0	0	0
Woonwijk	Location	3		Woonhuis	Location	4	
Size (µm)	PP	PVC	rubber	Size (µm)	PP	PVC	rubber
10-20	0	47	203	10-20	0	0	153
20-50	4	82	161	20-50	9	13	133
50-100	13	51	49	50-100	5	9	5
100-300	0	2	4	100-300	0	0	1
300-1000	0	0	0	300-1000	0	0	0
1000-5000	0	0	0	1000-5000	0	0	0

Table 7: Polypropylene (PP), Polyvinyl chloride (PVC) and different rubbers fibre concentrations per location in particles per m^3, Size = minimum diameter.

Production plant	**Location**	**1**		**Rotonde**	**Location**	**2**	
Size (µm)	PP	PVC	rubber	**Size (µm)**	PP	PVC	rubber
10-20	0	0	3	**10-20**	0	0	1
20-50	0	5	0	**20-50**	0	0	0
50-100	0	0	0	**50-100**	0	0	0
100-300	0	0	0	**100-300**	0	0	0
300-1000	0	0	0	**300-1000**	0	0	0
1000-5000	0	0	0	**1000-5000**	0	0	0
Woonwijk	**Location**	**3**		**Woonhuis**	**Location**	**4**	
Size (µm)	**PP**	**PVC**	**rubber**	**Size (µm)**	**PP**	**PVC**	**rubber**
10-20	0	0	4	**10-20**	0	0	0
20-50	0	0	0	**20-50**	0	0	0
50-100	0	4	0	**50-100**	0	0	0
100-300	0	2	0	**100-300**	0	0	0
300-1000	0	0	0	**300-1000**	0	0	0
1000-5000	0	0	0	**1000-5000**	0	0	0

3.4 Statistics on subsample measurement.

In this research, subsamples with about 1000 to 4000 particles were analysed due to the time constraints (analysis time) of identifying the entire population microplastics in a sample. Each sample contained in total between 15,000 to 45,000 particles. Note that not all of these particles, however, were microplastics, e.g. also sand was found. A full list of compounds/polymers in the database can be found in the supporting information (SI II).

As subsampling is a necessity when too many particles (more than ca. 4000) are found due to the length of the measurement, the relative error is determine. To the assess the relative error incurred from evaluating only a portion of the sample, four approximately equal sized subsamples were taken from one main sample. The particle counts for these subsamples were 3,480; 4,733; 1,179; and 1,684. The initial step involved calculating the relative proportion (as a percentage) of each type of particle within these counts. It is not possible to control the exact number of particles analysed within each subsample, , as only specific areas can be selected for analysis Next, the mean and the relative standard deviation (RSD) were calculated. Figure 8 shows the RSD for the various mean values per size and polymer combination. With the given samples sizes, it can be seen that for low relative particle numbers in a category (ca. 1% of the entire population) the RSD can be up to 200%. For relative particle numbers between 1 and 5 % of the population and higher the RSD drops and varies between 20 and 75%. For mean values of 7% or higher, the RSD lies between 20% and 30%. This indicates that the error is significantly larger for particles present at low fraction of the total number of particles when only a part of the total sample is evaluated, as one would expect.

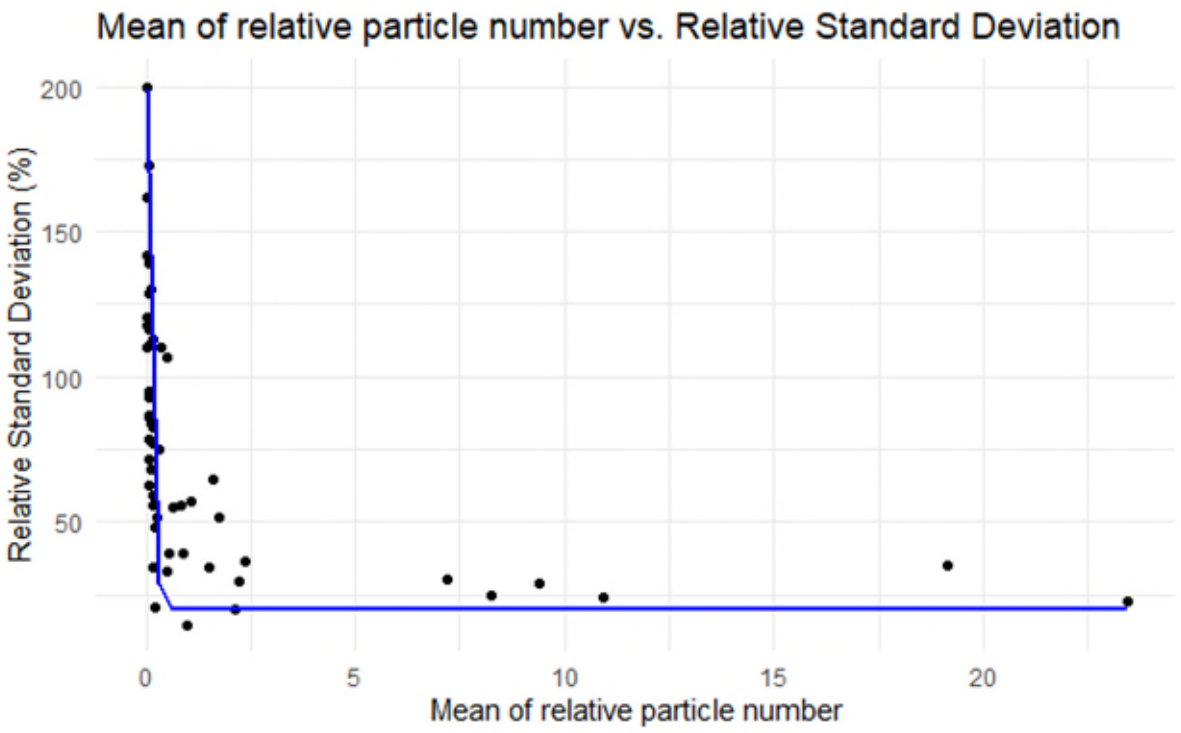


Figure 8: Mean of the percentage across the four subsamples against the relative standard deviation. Each dot is one type of polymer. The blue line is to guide the eye.

Additionally, the largest set of 4733 particles was randomly split into four subsets of 1000 particles each. This approach allowed for the use of actual particle numbers rather than relative ones. The boxplots of the particle numbers are shown in Figure 9, and Table 8 presents the RSD for the identified compounds. The RSD ranges from 2% to 200%, with the highest RSD observed for particles with the lowest counts (Figure 10). These results are comparable to those shown in Figure 8. In Table 9 also the relative particle numbers per subset were calculated, followed by the calculation of the RSD. The trend already shown in Figure 8 can also been seen in Figure 11. For large relative mean values the RSD remains around 25%. In the literature, it is noted that RSD values exceeding 100% are not uncommon in particle counting (Richter et al., 2023). One of the reason is the inevitable loss of particles during work-up, as shown with the recovery rate of about 80%. Another reason is that when subsampling suspensions, homogeneity as known from solutions, cannot as easily achieved. Therefore, an RSD of around 25% for particles with a high count is considered acceptable.

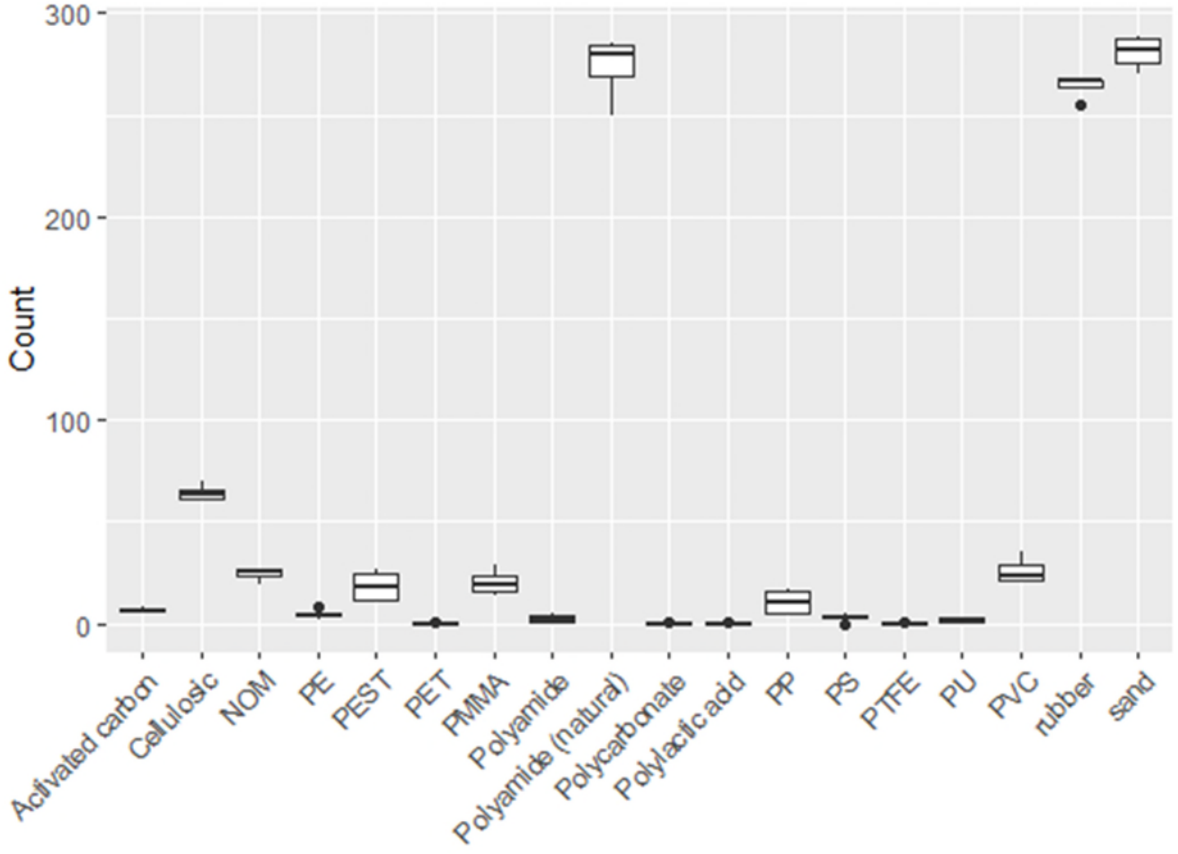


Figure 9: Boxplot of the particle number (Count) found in the four subsamples. Each of the subsamples contained 1000 particles.

Table 8: Relative standard deviation (RSD) for the particle number of various types of identified compounds in the four subsamples of 1000 particles.

Type	mean	sd	RSD
Activated carbon	6.5	1.3	19.9
Cellulosic	64.0	4.3	6.8
NOM	24.0	3.6	14.8
PE	4.5	2.5	55.9
PEST	18.5	8.2	44.2
PET	0.3	0.5	200.0
PMMA	20.3	6.5	32.1
PP	10.8	6.7	61.9
PS	2.8	2.1	75.0
PTFE	0.3	0.5	200.0
PU	1.5	1.3	86.1
PVC	25.5	6.9	26.9
Polyamide	2.5	1.9	76.6
Polyamide (natural)	273.5	16.3	6.0
Polycarbonate	0.3	0.5	200.0
Polylactic acid	0.3	0.5	200.0
rubber	264.3	6.2	2.3
sand	280.5	8.6	3.1

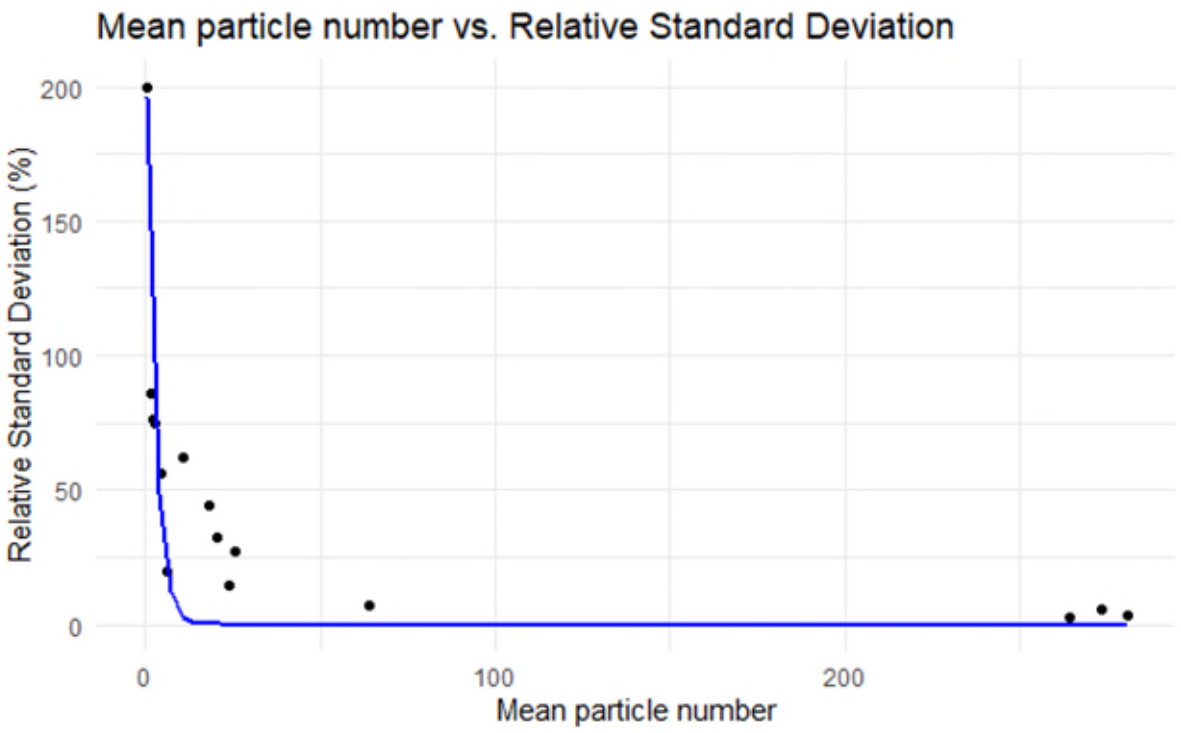


Figure 10: RSD for the mean of the particle numbers of the four sets of 1000 particles. See also Table 8. Each dot is one type polymer. Blue line is to guide the eye.

Table 9: Relative standard deviation (RSD) for the relative particle number of various types of identified compounds in the four subsamples of 1000 particles.

Type	mean_percentage	sd_percentage	RSD (percentage sd of the mean)
Activated carbon	0.6	0.3	53.2
Cellulosic	5.9	1.6	27.4
NOM	1.9	0.6	30.8
PE	0.4	0.4	82.2
PEST	1.4	0.3	23.9
PET	0.0	0.0	91.9
PMMA	3.8	1.7	45.5
PP	0.7	0.3	37.0
PS	0.2	0.1	53.2
PTFE	0.1	0.1	143.5
PU	0.2	0.2	124.0
PVC	1.2	1.0	90.2
Polyamide	0.2	0.1	45.9
Polyamide (natural)	40.9	10.8	26.4
Polycarbonate	0.0	0.0	200.0
Polylactic acid	0.0	0.0	200.0
rubber	21.1	5.7	27.1
sand	21.5	4.8	22.2

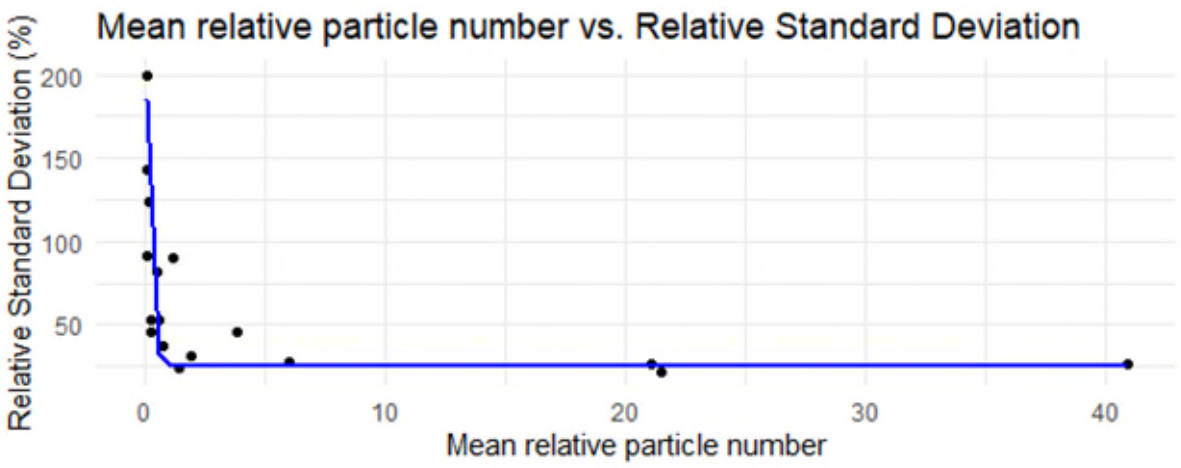


Figure 11: RSD of the relative mean standard deviation of the four sets of 1000 particles. See also Table 9. Blue line is to guide the eye.

In a last step it was investigated how the number of subsamples does effect the error margin and the RSD. For this two types of particles were chosen in a given dataset. One particle type (PS) had a low particle count of 14 in a total population of 4732 and the other one (rubber) a high count of 1267 particles (Figure 12). The RSD of the particle numbers were checked for subsample sizes of 100, 500 and 1000 using a bootstrapping method.

First the stability of the relative standard deviation (RSD) for polystyrene (PS) particles when subsampling from an original sample of 4,732 particles, which includes 14 PS particles was assessed. The results indicate that as the number of subsamples increases, the RSD decreases across all tested subsample sizes (100, 500, and 1000 particles), suggesting greater consistency in particle representation. Subsample sizes of 500 and 1000 exhibit notably lower RSD values compared to a subsample size of 100, particularly as the number of subsamples reaches 3 or more. This suggests that a minimum of 3 subsamples of size 500 particles or more may provide an acceptable representation of the particle population.

Next, for rubber particles within the sample of 4,732 total particles (of which 1,267 were rubber), the RSD is evaluated across various subsample sizes (100, 500, and 1000 particles) and numbers of subsamples. The results also show that the RSD decreases as the number of subsamples increases, with larger subsample sizes (500 and 1000 particles) achieving more stable and lower RSD values. For a subsample size of 100, the RSD stabilizes at a higher subsample numbers compared to larger subsample sizes. Also for higher counts of one type of particles, these findings suggest that using at least 3 subsamples of 500 or more particles each is sufficient to provide an accurate representation of the rubber particle population.

For both high and low particle counts, the results indicate that it is not necessary to measure the entire population. Three subsamples are sufficient to provide a reliable representation of the entire sample. This is important as the analysis time directly correlates with the number of particles in a sample. This does not constitute a problem if only occasionally a sample needs to be measured. However, when multiple samples need to be analysed, subsampling is necessary. E.g., analysing 13.000 particles takes more than

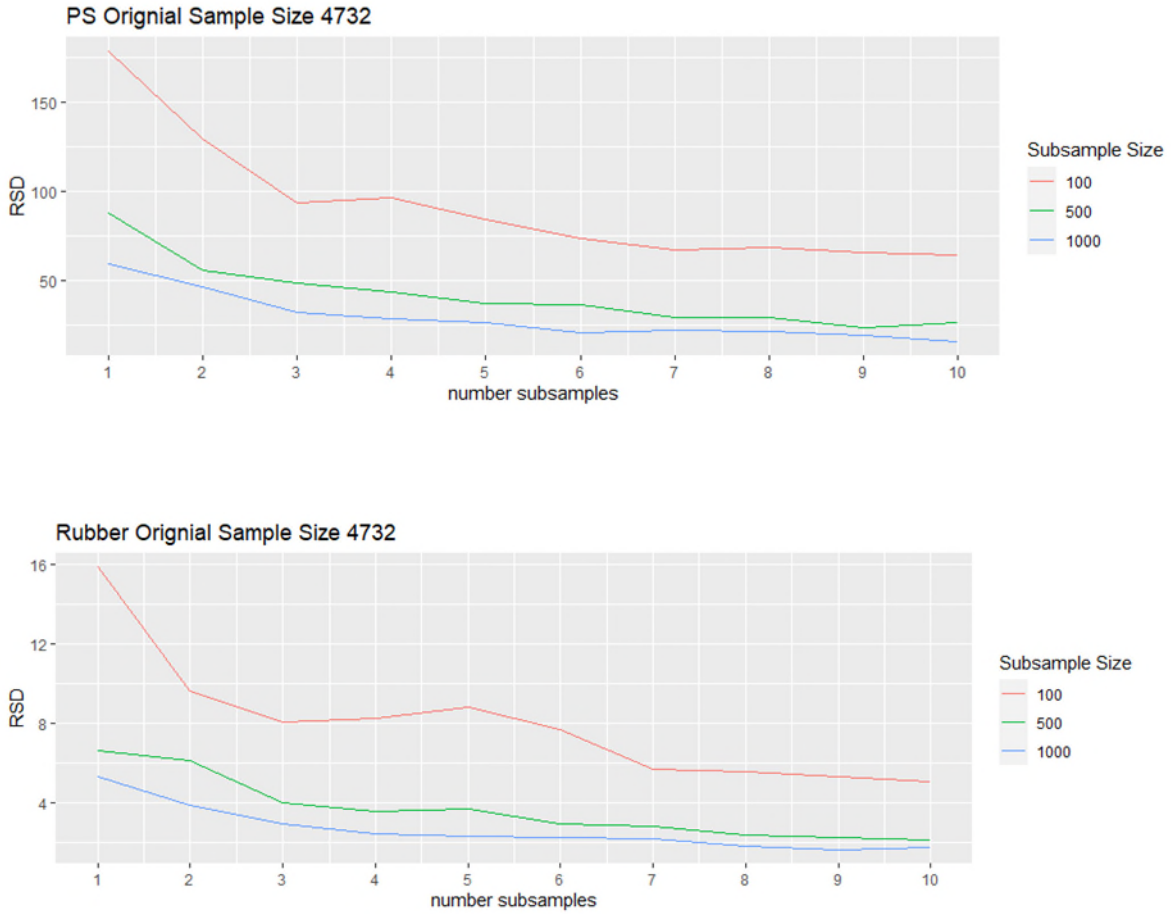


Figure 12: Top: Variation in the RSD for polystyrene (PS) particles, with 14 particles in a sample of 4,732, analysed across three different subsample sizes and varying numbers of subsamples. Bottom: Variation in the RSD for rubber particles, with 1,267 particles in a sample of 4,732, analysed across three different subsample sizes and varying numbers of subsamples.

3.5 Conclusions

Microplastic release from pipes

At four different locations in the drinking water distribution network samples were taken, starting at the drinking water plant and ending at a tap in a private home. At each location one sample was taken and analysed. There were no large amounts of additional PVC particles found downstream of the production location. Statistically these numbers cannot be supported due to the fact that only one single sample per location could be taken. Nevertheless the outcome of those four locations did not show any prove that PVC-shedding was occurring within the pipe system. Since PVC is the pipe material of choice in the Dutch drinking water practice, this is a worthwhile result to verify. The current results do not have any implications for the drinking water companies as particle numbers are very low across the distribution network (ca. 1 p/L) and there is no sign of increasing numbers of PVC or other particles between the production facility, within the network and at the tap. Also, the numbers are comparable to earlier findings. However, the current study evaluated a single production location, so additional evidence in other distribution networks with a different size, used material, age and pumps, with sufficient numbers of samples per sampling point are required to evaluate this in more detail.

Delegated act requirement to monitor microplastics

The delegated act supplementing Directive (EU) 2020/2184 requires national government to implement a monitoring strategy. This delegated act lays out the requirements that need to be met, such as ways to report data, type of analytical equipment and sampling devices. Within this project, a previously developed ML model was improved and implemented in a software, able to run on any windows PC. The model can identify microplastics based on their infrared spectra. This was chosen to be independent from commercial software. Apart from being independent from a commercial supplier, also the accuracy of the identification was improved as the ML models are better than the current commercially available software. An accuracy of about 99% was achieved with our software. To report the data according to the delegated act, an R-script was written. Additionally, a sampling device was built according to the specifications mentioned in the delegated act. A first test with this device shows that it works expectedly. With the here developed tools, the drinking water companies are able to meet all the requirements laid out in the delegated act.

Statistical analysis

The analysis of microplastics in water is labour intensive, especially when large numbers of particles are collected. Therefore the analysis of subsample size and number on the reliability of particle analysis was evaluated to balance analytical efforts and quality (accuracy) of the results. This study evaluated samples containing 15,000 to 45,000 particles. Due to time constraints (analysis time) it was investigated if subsamples of 1,000–4,000 particles are representative for the entire population. The results indicate that for low-abundance particle types (approximately 1% of the sample), RSD can reach up to 200%, whereas higher-abundance particles (7% or more) show RSD values between 20% and 30%. This variation underscores the higher error associated with low-count particles when analysing only a portion of the sample. These findings align with literature, where RSD values over 100% are common in particle counting (Richter et al., 2023). Therefore we propose that a RSD of around 25% is acceptable for high-count particles. For lower particles numbers a higher RSD can be accepted as long as the absolute error is small.

Additionally, the influence of the number of subsamples for various samples sizes was investigated. A bootstrapping analysis on subsample sizes of 100, 500, and 1000 particles for polystyrene (14 particles) and rubber particles (1267 particles) in a sample of 4,732 shows that increasing the number of subsamples reduces RSD, particularly for subsample sizes of 500 and above. For both high and low particle counts, a minimum of three subsamples of 500 or more particles provided a reliable representation, balancing analytical efforts and data quality. This suggests that that full-sample analysis of samples containing large numbers of particles. More importantly, since the particle number can only be determined after sampling, it is impossible to take an exact amount of water to achieve a suitable particle concentration. Therefore, the ability to analyse subsamples is essential.

3.6 Recommendations

Software and machine learning model

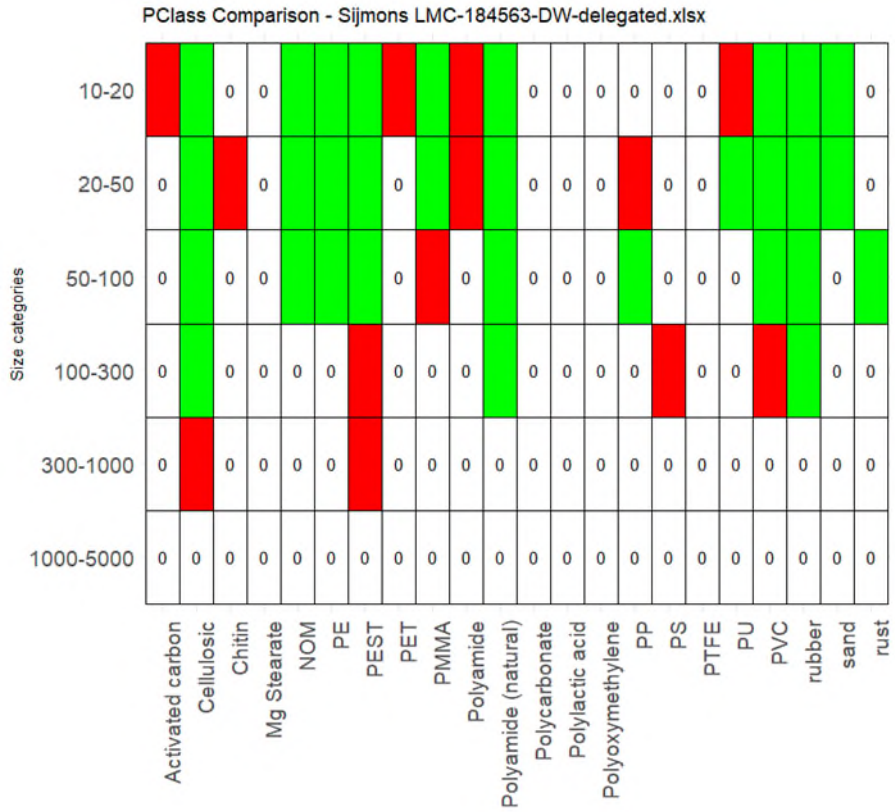
Time does not stand still, especially in the field of ML. For this reason, it is crucial to regularly update both the current model and the associated software. Additionally, evaluating new ML models that may enhance the software is highly recommended. The rapid advancements in this field have been significant. One needs only to consider the improvements made by ChatGPT over the last two to three years to recognize the potential for rapid and vast enhancements. Also, regularly updating and increasing the training data set for the model is highly recommended to cover a broader spectrum of polymer particles and include emerging polymers in water. This includes using the model in interlaboratory studies or round robins.

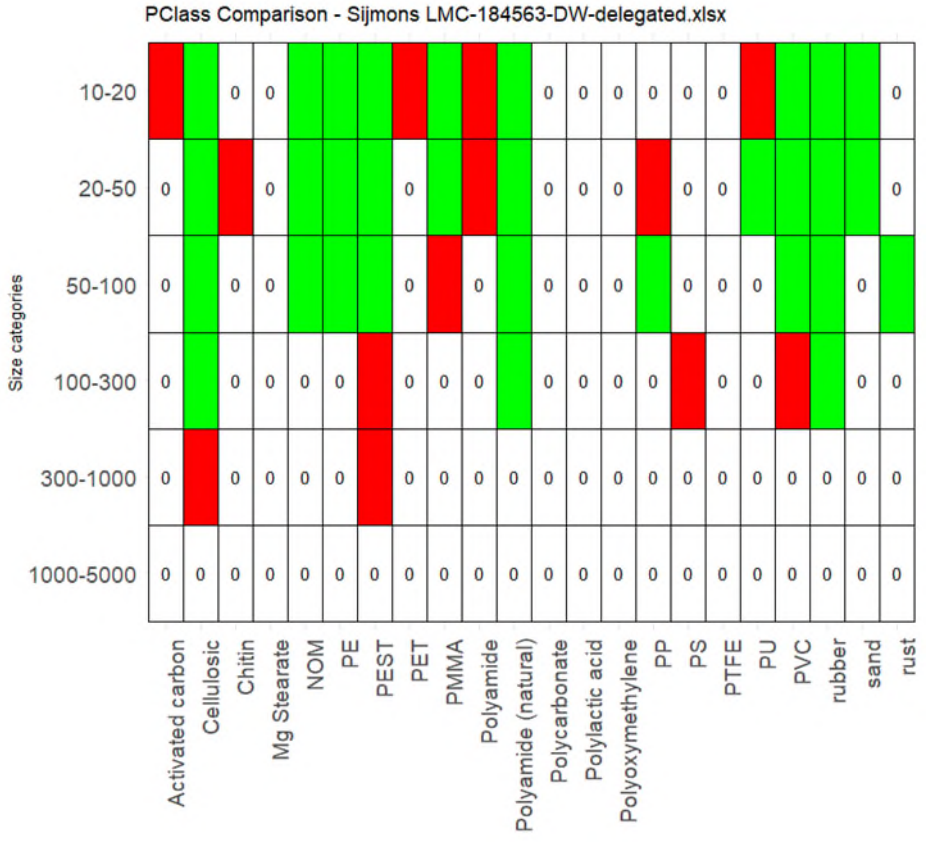
Monitoring of drinking water

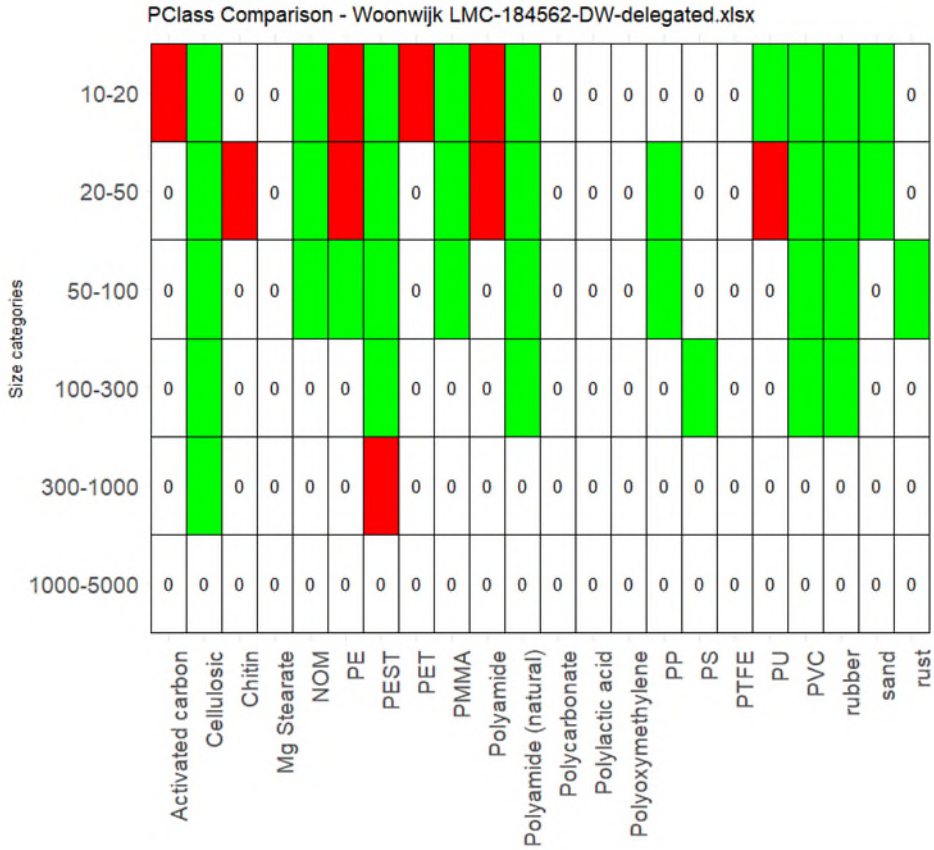
It is recommended to measure the microplastic concentrations in at least two different drinking water distribution networks and take several samples at the same location to allow a statistical analysis of the data collected. By doing so, we can substantiate our statements.

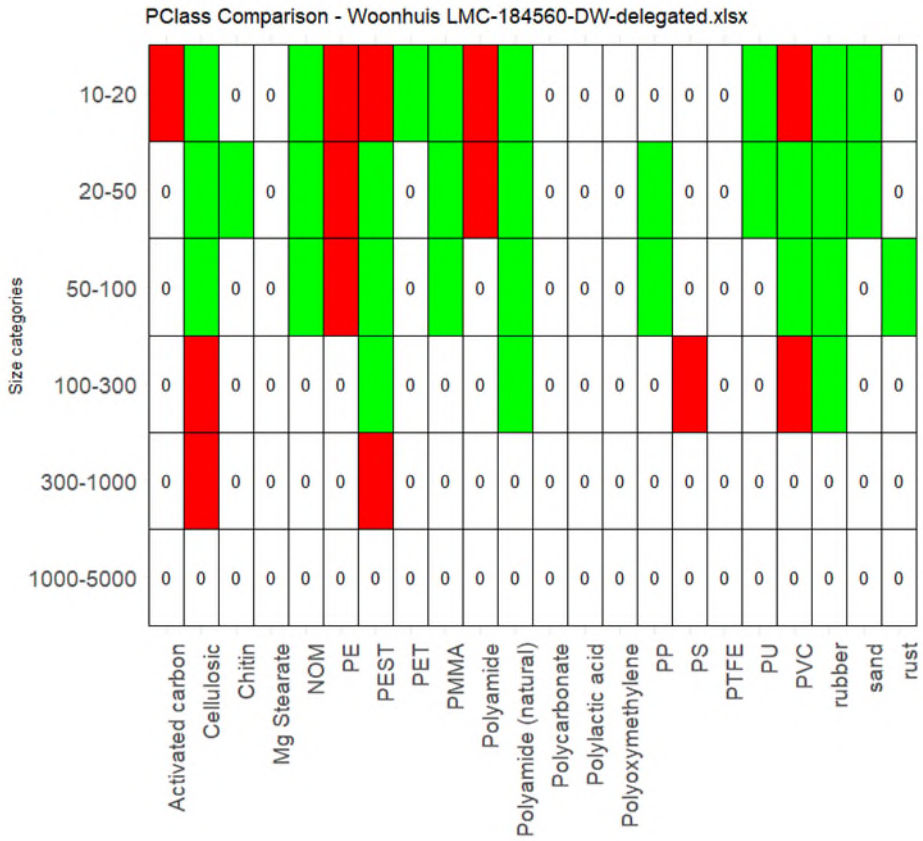
I Comparison to procedural blank

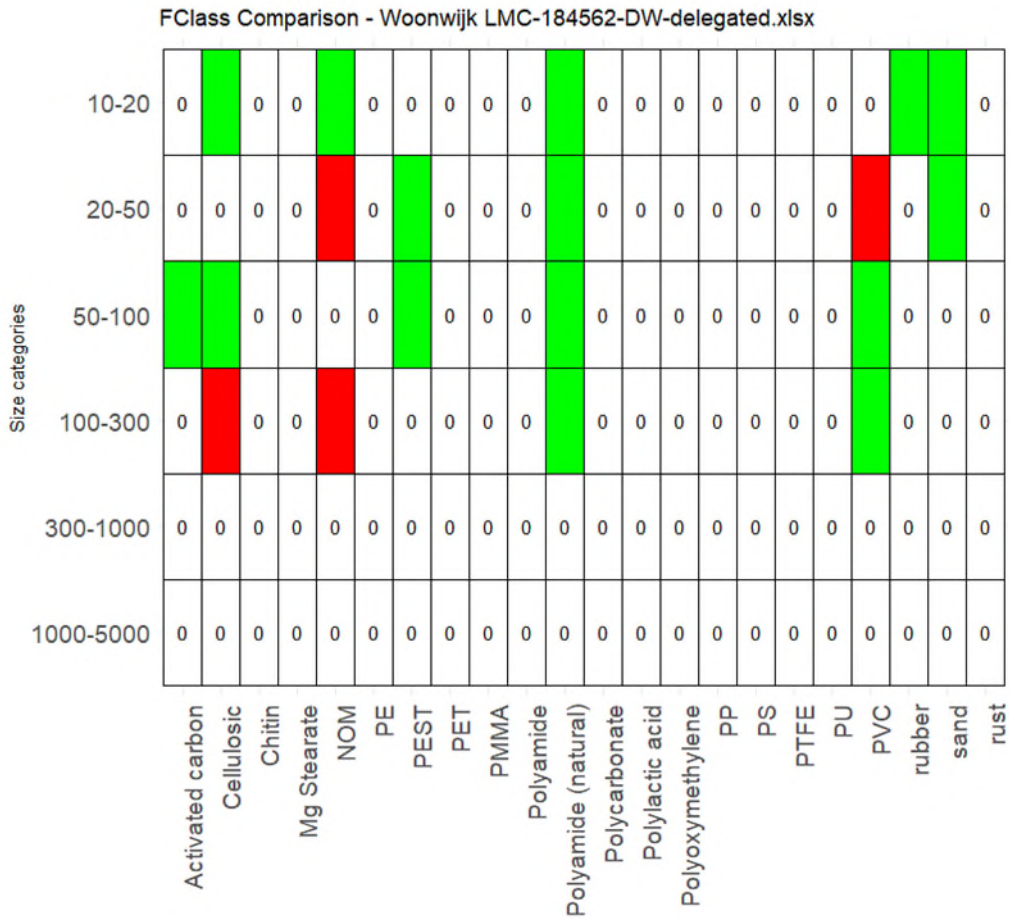
The figures show which of the sample meet the quality criteria (see 3.3.1). A red block means this value was excluded and set to zero as it did not meet the statistical requirement mentioned in the delegated act. Green means that value is reported and a zero means that no particles were found in the sample in the first place.

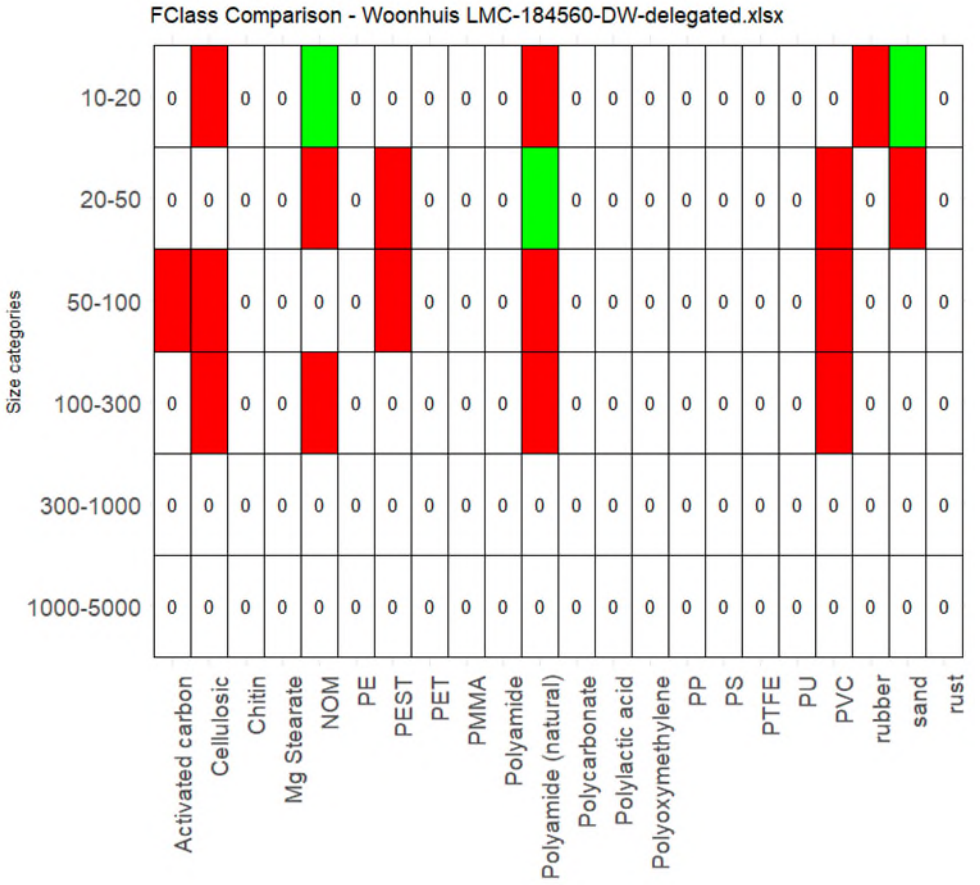


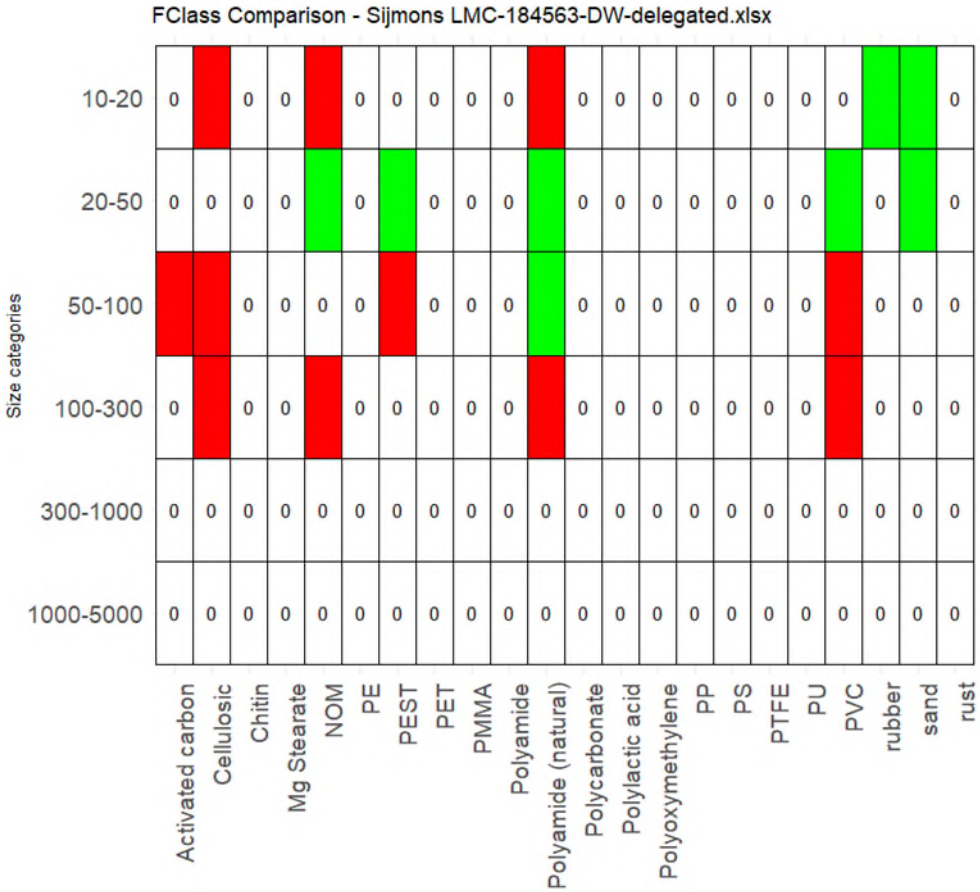

PClass Comparison - Sijmons LMC-184563-DW-delegated.xlsx
Size categories
10-20
20-50
50-100
100-300
300-1000
1000-5000
Activated carbon
Cellulosic
Chitin
Mg Stearate
NOM
PE
PEST
PET
PMMA
Polyamide
Polyamide (natural)
Polycarbonate
Polylactic acid
Polyoxymethylene
PP
PS
PTFE
PU
PVC
rubber
sand
rust

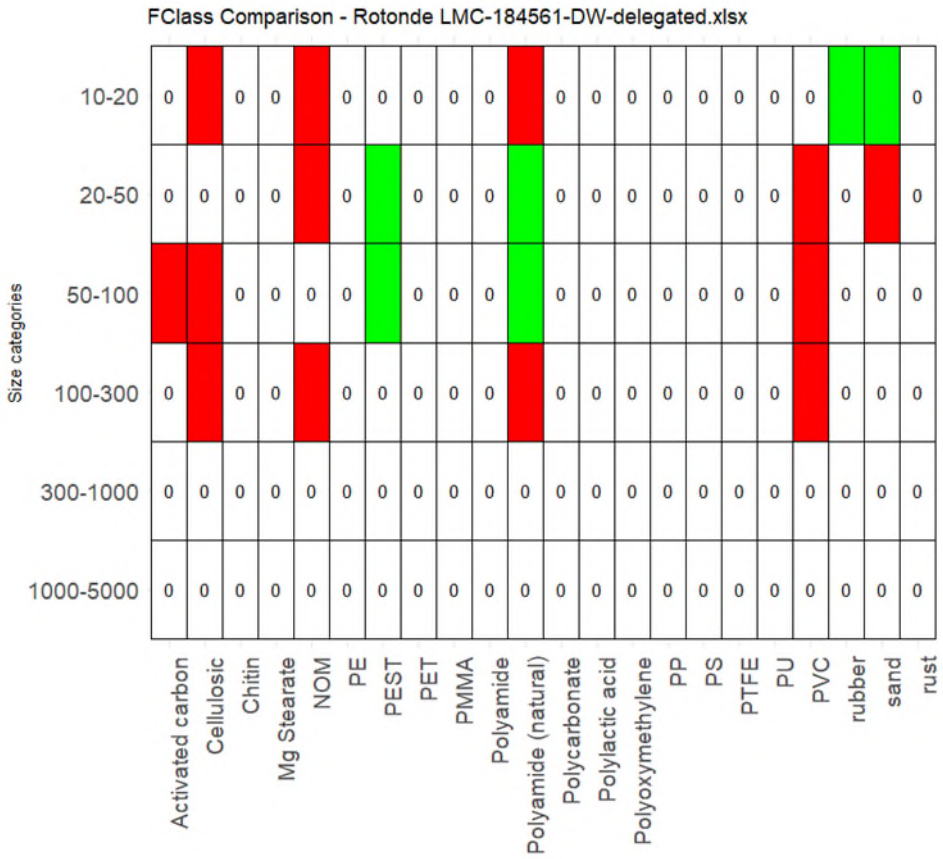

PClass Comparison - Woonwijk LMC-184562-DW-delegated.xlsx
Size categories
10-20
20-50
50-100
100-300
300-1000
1000-5000
Activated carbon
Cellulosic
Chitin
Mg Stearate
NOM
PE
PEST
PET
PMMA
Polyamide
Polyamide (natural)
Polycarbonate
Polylactic acid
Polyoxymethylene
PP
PS
PTFE
PU
PVC
rubber
sand
rust


PClass Comparison - Woonhuis LMC-184560-DW-delegated.xlsx
Size categories
10-20
20-50
50-100
100-300
300-1000
1000-5000
Activated carbon
Cellulosic
Chitin
Mg Stearate
NOM
PE
PEST
PET
PMMA
Polyamide
Polyamide (natural)
Polycarbonate
Polylactic acid
Polyoxymethylene
PP
PS
PTFE
PU
PVC
rubber
sand
rust


FClass Comparison - Woonwijk LMC-184562-DW-delegated.xlsx
Size categories
10-20
20-50
50-100
100-300
300-1000
1000-5000
Activated carbon
Cellulosic
Chitin
Mg Stearate
NOM
PE
PEST
PET
PMMA
Polyamide
Polyamide (natural)
Polycarbonate
Polylactic acid
Polyoxymethylene
PP
PS
PTFE
PU
PVC
rubber
sand
rust


FClass Comparison - Woonhuis LMC-184560-DW-delegated.xlsx
Size categories
10-20
20-50
50-100
100-300
300-1000
1000-5000
Activated carbon
Cellulosic
Chitin
Mg Stearate
NOM
PE
PEST
PET
PMMA
Polyamide
Polyamide (natural)
Polycarbonate
Polylactic acid
Polyoxymethylene
PP
PS
PTFE
PU
PVC
rubber
sand
rust


FClass Comparison - Sijmons LMC-184563-DW-delegated.xlsx
Size categories
10-20
20-50
50-100
100-300
300-1000
1000-5000
Activated carbon
Cellulosic
Chitin
Mg Stearate
NOM
PE
PEST
PET
PMMA
Polyamide
Polyamide (natural)
Polycarbonate
Polylactic acid
Polyoxymethylene
PP
PS
PTFE
PU
PVC
rubber
sand
rust



II List of compounds used for these measurements (class and subclass)

This the list of polymers and other compounds in the library at the time of this research project. EU means that particle needs to be reported according to the delegated act.

Class	Subclass	Category
Activated carbon	Activated carbon	nonpolymer
Cellulosic	cellulose acetate	nonpolymer
Cellulosic	Cellulosic	nonpolymer
Cellulosic	tissue	nonpolymer
Chitin	Chitin	nonpolymer
Mg Stearate	Mg Stearate	nonpolymer
NOM	NOM	nonpolymer
PE	PE	EU
PE	PE100	EU
PE	PE-RT	EU
PE	PE-Xa	EU

PEST	PEST	polymer
PET	PET	EU
PMMA	PMMA	EU
Polyamide	Nylon	EU
Polyamide (natural)	animal polyamide	nonpolymer
Polyamide (natural)	finger nail	nonpolymer
Polyamide (natural)	hair	nonpolymer
Polycarbonate	Polycarbonate	EU
Polylactic acid	Polylactic acid	polymer
Polyoxymethylene	Polyoxymethylene	polymer
PP	PP	EU
PP	PP-R	EU
PP	PP-RCT	EU
PS	expanded PS	EU
PS	PS	EU
PS	PS oxidised	EU
PTFE	PTFE	EU
PU	PU	EU
PVC	PVC	EU
PVC	PVC-C	EU
PVC	PVC-Phthalate	EU
PVC	PVC-U	EU
rubber	ABS	rubber
rubber	nitrile	rubber
rubber	NR	rubber
rubber	PB	polymer
rubber	rubber	rubber
rubber	rubber1	rubber
rubber	rubber2	rubber
rubber	SBR	rubber
rubber	silicone	rubber
rubber	tyre	rubber
rubber	used car tyre	rubber
sand	sand	nonpolymer
rust	rust	nonpolymer

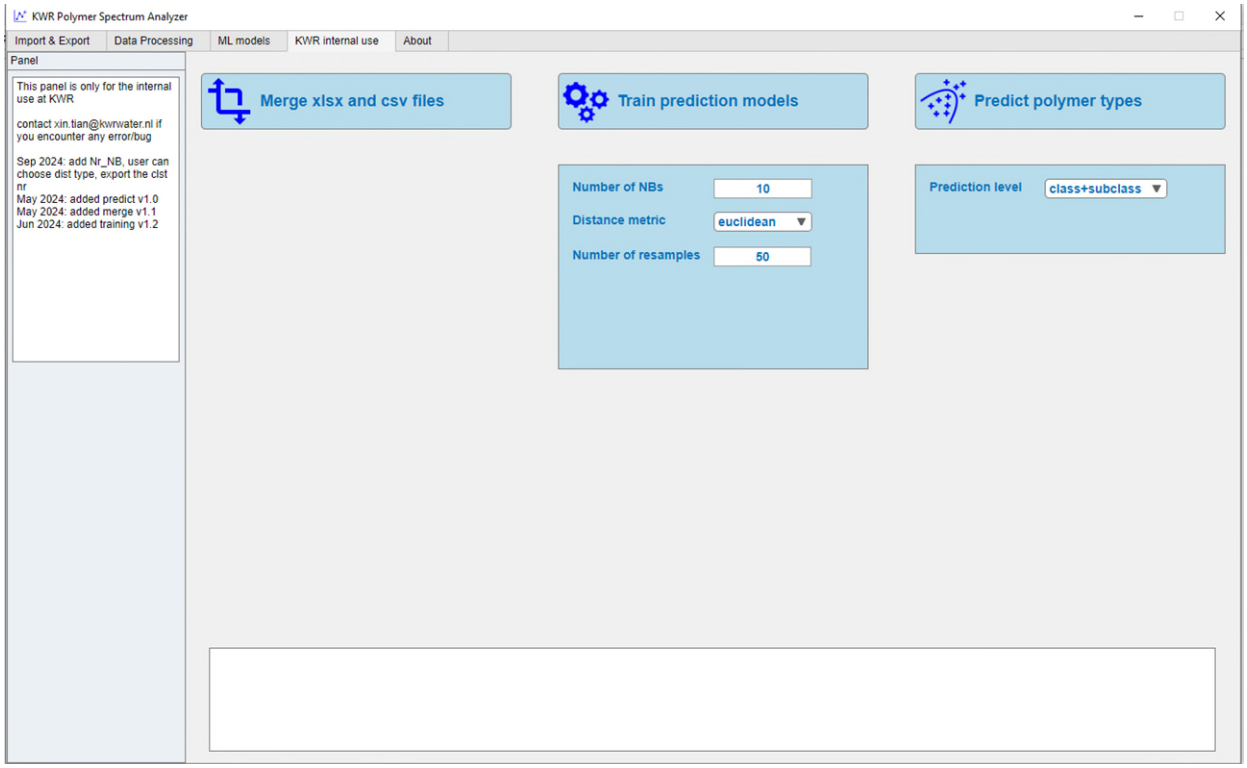
III Database v13 (current version)

The current version of the database contains the following plastics and polymers. The database can be updated easily to include more types of particles.

Class	Subclass
cellulose acetate	cellulose acetate
Cellulosic	Cellulosic
Cellulosic	tissue
NOM	NOM
PE	PE
PE-2	PE100
PE-2	PE-RT
PE-2	PE-Xa
PEST	PEST
PET	PET
PMMA	PMMA
Nylon	Nylon66
Nylon	Copolyamide
Nylon	Nylon11
Nylon	Nylon12
Nylon	Nylon6
Nylon	Nylon 63
Nylon	Nylon 612
Nylon	Nylon 69
Polyamide (natural)	animal polyamide
Polyamide (natural)	finger nail
Polycarbonate	Polycarbonate
Polylactic acid	Polylactic acid
Polyoxymethylene	Polyoxymethylene
PP	PP
PP	PP-R
PP	PP-RCT
PS	expanded PS
PS	PS
PS	PS oxidised
PTFE	PTFE
PU	PU
PVC	PVC
PVC	PVC-C
PVC	PVC-Phthalate

PVC	PVC-U
ABS	ABS
nitrile	glove1
nitrile	glove2
NR	NR
PB	PB
SBR	SBR
tyre	tyre
tyre	goodyear
tyre	used car tyre
sand	sand
Norit	norit B
Norit	norit C
Norit	Norit

IV Manual ML software



1. Interface Overview

1.1. Main Panels and Buttons

- Merge xlsx and csv files: Allows users to merge multiple Excel and CSV files (produced by the LDIR machine).

- Train prediction models: Facilitates the training of machine learning models for prediction tasks (only supports Subspace KNN at this moment). This step takes long processing time if the dataset is large.

- Predict polymer types: Enables the prediction of polymer types using trained models.

1.2 Import & Export

Merging Files

Click on the Merge xlsx and csv files button.

A dialog box will appear prompting you to select the files you want to merge.

Select the desired Excel and CSV files from your directory.

Click Open to import the files.

The interface will merge the files and display a confirmation message once the process is complete.

Training Prediction Models

Click on the Train prediction models button.

Upload the dataset you want to use for training. The dataset should be in Excel format.

The model training will start automatically.

A progress bar will indicate the training status. Once completed, a message will confirm that the model is trained.

Predicting Polymer Types

Using Trained Models for Prediction

Click on the Predict polymer types button.

Upload the dataset for which you want to predict polymer types. Ensure the data format matches the one used during training. (see the example)

Select the trained model you wish to use for predictions (only support Subspace KNN at this moment).

The prediction will start automatically.

The results will be displayed on the screen, and results will be saved in the same folder where the imported data are stored.

Common Issues

File Import Errors: Ensure that the files are in the correct format (xlsx or CSV) and are not corrupted.

Model Training Failures: Check that the dataset is properly formatted and contains all necessary features.

Prediction Errors: Verify that the input data matches the format and features used during model training.

Memory requirement: 16 GB at least

Note: this is a test version, which is only for internal use at KWR

Code: https://github.com/KWR-Water/Microplastics_402045_351 (Not sharable outside KWR, contact Patrick Bauerlein or Xin Tian)

V Material used for the sampling device

Part (Nederlands)	Amount	Order number	Leverancier
Tri-clamp Opschroef ferrule 316L 2.0" FL=64.0 x 0.50" BSP BI	24	184352003012	Mulder Stainless
Screengasket 316L/EPDM 40 MESH 2.0"	12	189992000300004	Mulder Stainless
Klembeugel (1 scharnierpunt) SP 304 64.00 mm DN 50 - 2.00"	12	188SP12001	Mulder Stainless
Pijpnippel BSP 316 1/2" x 060 mm	15	28620007060	Mulder Stainless
Kogelkraan	2		
10 µm filter	2		
20 µm filter	2		
100 µm filter	2		
Watermeter	1		Gardena

VI Literature

Adediran, G.A., Cox, R., Jürgens, M.D., Morel, E., Cross, R., Carter, H., Pereira, M.G., Read, D.S. and Johnson, A.C. 2024. Fate and behaviour of Microplastics (> 25µm) within the water distribution network, from water treatment works to service reservoirs and customer taps. Water Research 255, 121508.

Bäuerlein, P.S., Erich, M.W., van Loon, W.M.G.M., Mintenig, S.M. and Koelmans, A.A. 2022a. A monitoring and data analysis method for microplastics in marine sediments. Marine Environmental Research 183, 105804.

Bäuerlein, P.S., Hofman-Caris, R.C.H.M., Pieke, E.N. and Ter Laak, T.L. 2022b. Fate of microplastics in the drinking water production. Water Res 221, 118790.

Carruthers, H., Clark, D., Clarke, F.C., Faulds, K. and Graham, D. 2022. Evaluation of laser direct infrared imaging for rapid analysis of pharmaceutical tablets. Analytical Methods 14(19), 1862-1871.

Mintenig, S.M., Kooi, M., Erich, M.W., Primpke, S., Redondo- Hasselerharm, P.E., Dekker, S.C., Koelmans, A.A. and van Wezel, A.P. 2020. A systems approach to understand microplastic occurrence and variability in Dutch riverine surface waters. Water Research 176, 115723.

Phan, S., Padilla-Gamiño, J.L. and Luscombe, C.K. 2022. The effect of weathering environments on microplastic chemical identification with Raman and IR spectroscopy: Part I. polyethylene and polypropylene. Polymer Testing 116.

Richter, S., Horstmann, J., Altmann, K., Braun, U. and Hagendorf, C. 2023. A reference methodology for microplastic particle size distribution analysis: Sampling, filtration, and detection by optical microscopy and image processing. Applied Research 2(4), e202200055.

Schymanski, D., Oßmann, B.E., Benismail, N., Boukerma, K., Dallmann, G., von der Esch, E., Fischer, D., Fischer, F., Gilliland, D., Glas, K., Hofmann, T., Käppler, A., Lacorte, S., Marco, J., Rakwe, M.E.L., Weisser, J., Witzig, C., Zumbülte, N. and Ivleva, N.P. 2021. Analysis of microplastics in drinking water and other clean water samples with micro-Raman and micro-infrared spectroscopy: minimum requirements and best practice guidelines. Analytical and Bioanalytical Chemistry 413(24), 5969-5994.

Shruti, V.C., Pérez-Guevara, F., Elizalde-Martínez, I. and Kutralam-Muniasamy, G. 2020. First study of its kind on the microplastic contamination of soft drinks, cold tea and energy drinks - Future research and environmental considerations. Science of The Total Environment 726, 138580.

Tian, X., Beén, F. and Bäuerlein, P.S. 2022. Quantum cascade laser imaging (LDIR) and machine learning for the identification of environmentally exposed microplastics and polymers. Environmental Research 212, 113569.

Tian, X., Beén, F., Sun, Y., van Thienen, P. and Bäuerlein, P.S. 2023. Identification of Polymers with a Small Data Set of Mid-infrared Spectra: A Comparison between Machine Learning and Deep Learning Models. Environmental Science & Technology Letters.