Supplementary Information

An automated approach for microplastic detection using focal plane array (FPA) FTIR microscopy and image analysis.

S. Primpke,[a] C. Lorenz,[a] R. Rascher-Friesenhausen[b,c] and G. Gerdts[a]

[a] *Alfred-Wegener-Institute Helmholtz Centre for Polar and Marine Research, Biologische Anstalt Helgoland, Kurpromenade 201, 27498 Helgoland.*

[b] *Hochschule Bremerhaven, Studiengang Medizintechnik, An der Karlstadt 8, 27568 Bremerhaven.*

[c] *Fraunhofer MEVIS, Institute for Medical Image Computing, Universitaetsallee 29, 28359 Bremen.*

Tables

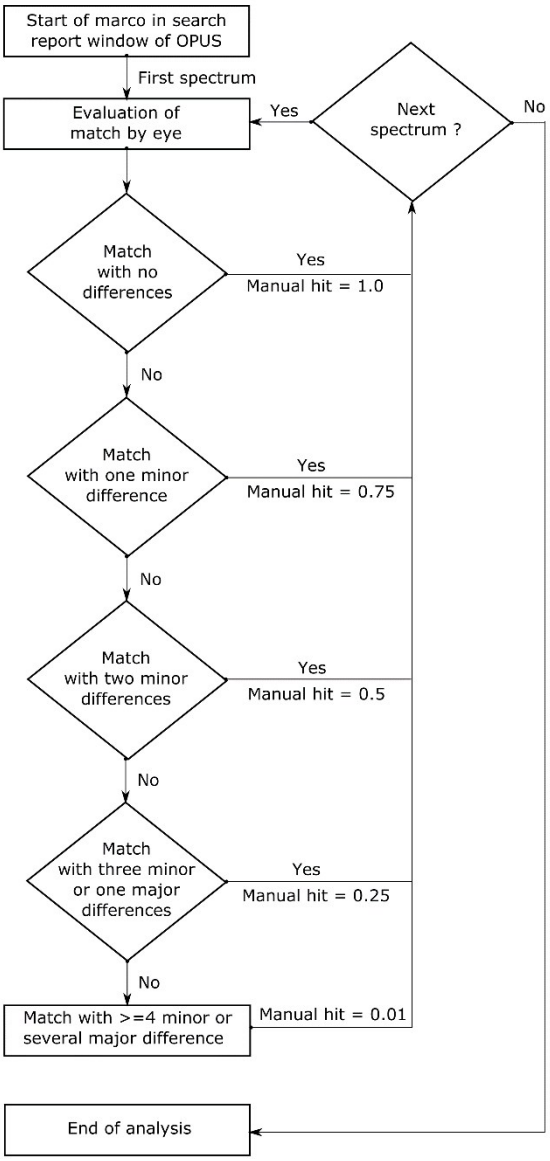
Tab. S1: The different filling matrices applied during pixel approximation for a distance of one and two pixels for image analysis of automated generated data.

Filling method	1 Pixel horizontal/ vertical	1 Pixel diagonal	2 Pixel horizontal/ vertical	2 Pixel diagonal	2 Pixel Staggered Case a)	2 Pixel Staggered Case b)	2 Pixel Staggered Case c)
a) Before closing	1 0 1	1 0 1	1 0 0 1	1 0 0 1	0 0 1 1 0 0	0 1 0 0 0 0 1 0	0 1 0 0 0 0 1 0
b) After closing	1 1 1	1 1 1	1 1 1 1	1 1 1 1	0 1 1 1 1 0	0 1 0 1 0 1 1 0	0 1 1 0 0 1 1 0

Tab. S2: Particle numbers for 12 polymer classes found via automated analysis for the sample H18_21 in dependence of the minimal size class. For full range from 10 to 500 µm and for comparsion with manual data from 25 to 500 µm.

Size / µm	PE	PP	PEST	PS	PU	PVC	PA	PVOH	Varnish	Copolymer	Rubber	Other	Sum
10 – 500	87	241	1	15	1	33	2	4	147	44	42	116	733
25 – 500	12	52	-	9	-	10	1	-	40	14	6	19	163

Schemes:



Scheme S1: Flow chart for manual reanalysis transmission-FPA-FTIR within the Bruker OPUS© software using the macro function.

Figures

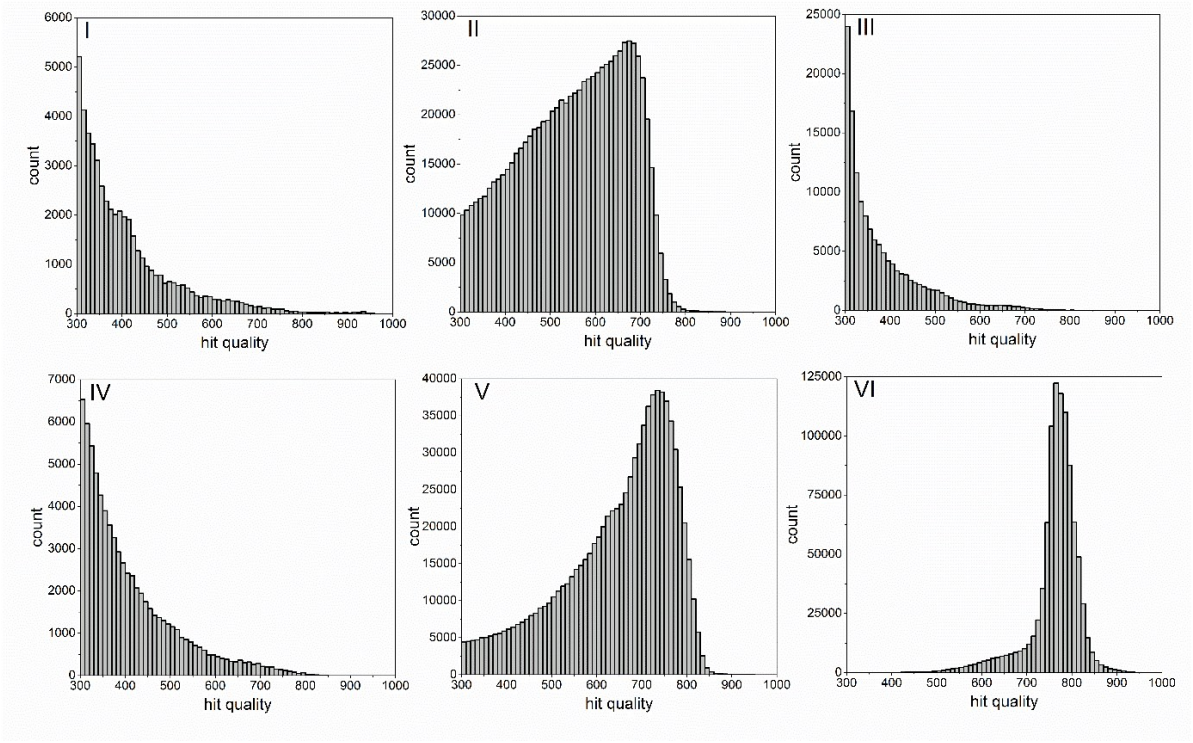


Fig.S1: Histogram of hit quality for the different routines (**I** to **VI**).

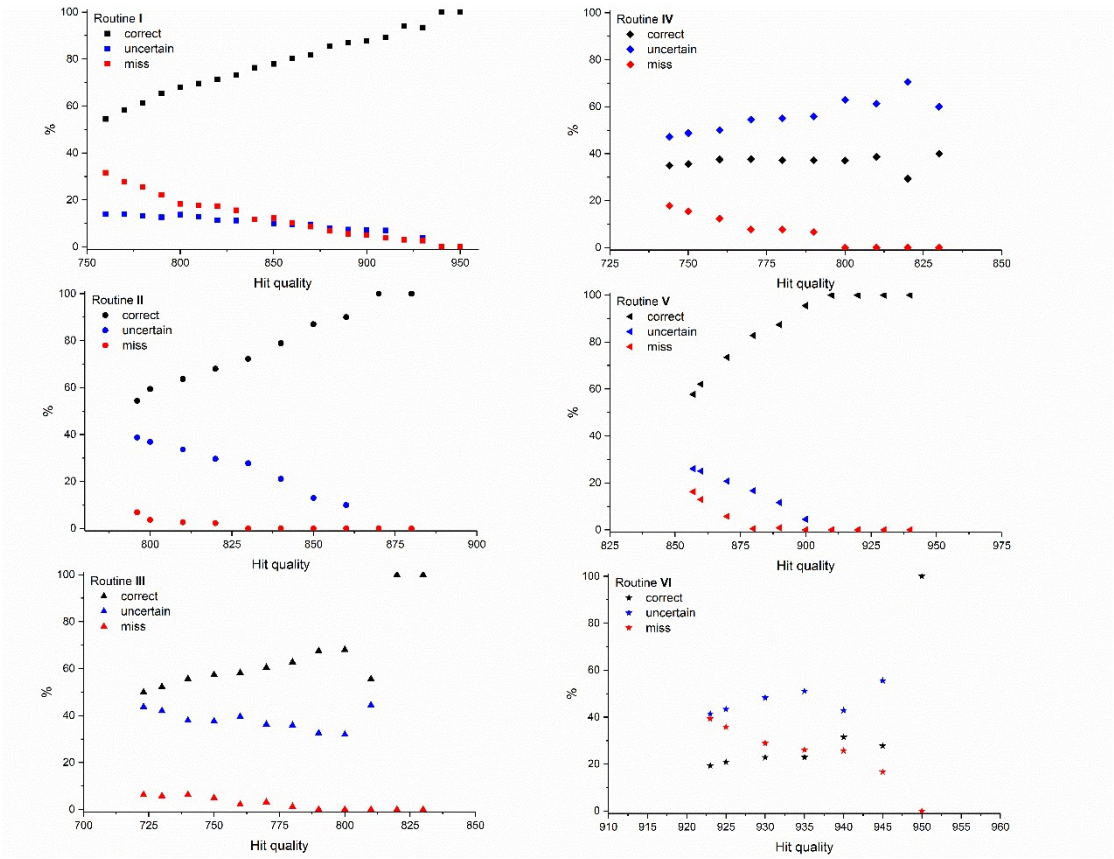


Fig. S2: Reanalysis for the first 541 to 558 hits for each routine (**I** to **IV**) showing the confidence interval at different hit quality thresholds for correct , uncertain and misassignments .

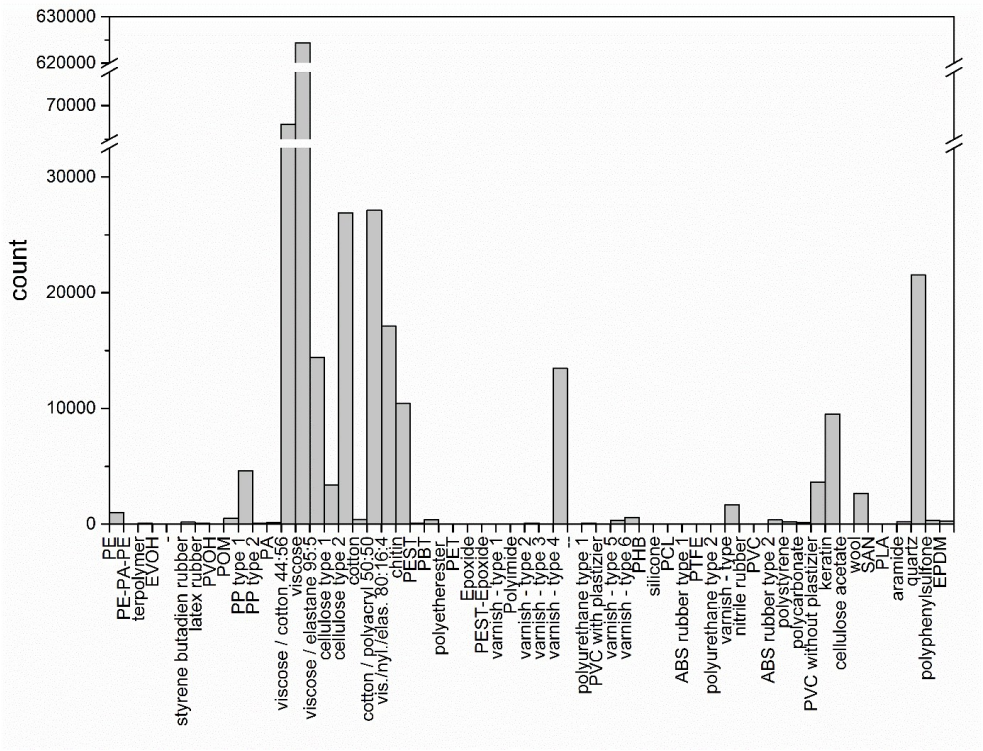


Fig. S3: Histogram of polymer identification of routine **II**.

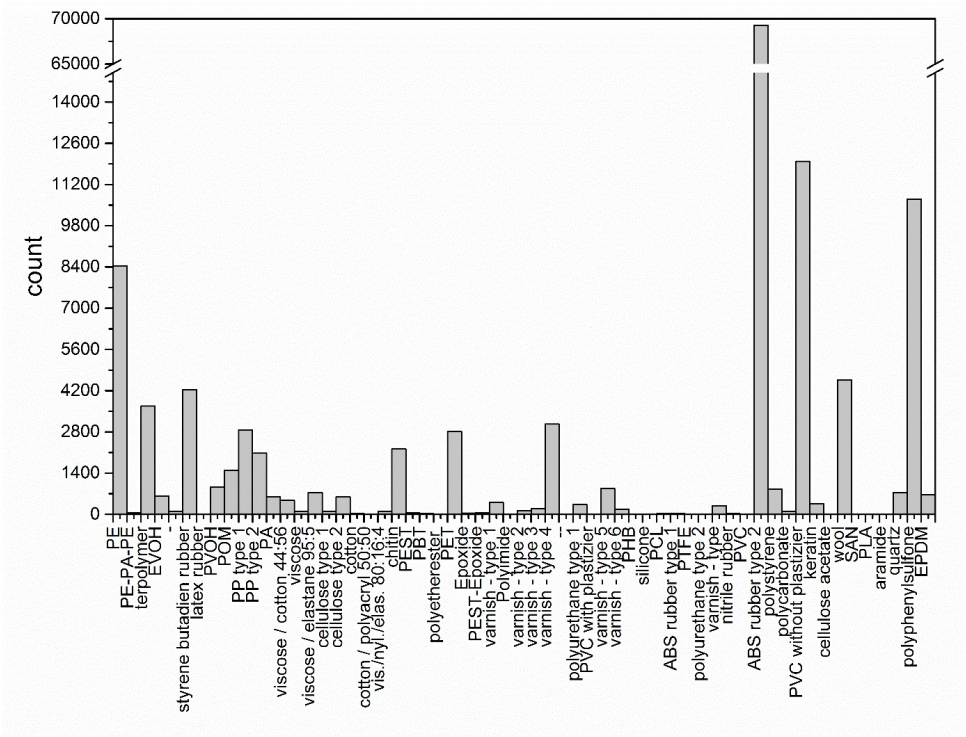


Fig. S4: Histogram of polymer identification of routine **III**.

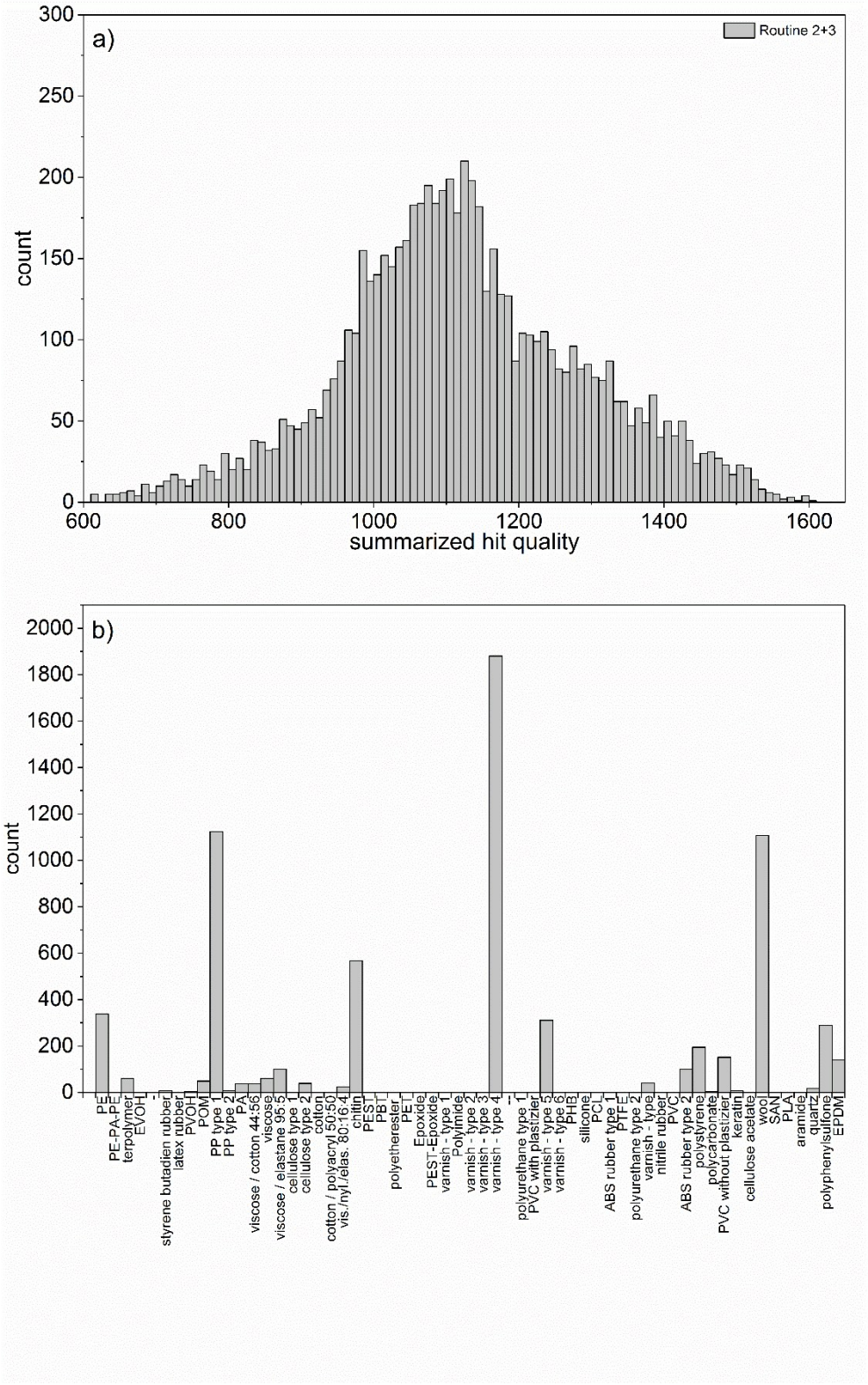


Fig. S5: Result of the combined analysis via OPUS 7.2 and the routines **II** and **III**. a) Histogram of the summarized hit quality; b) Histogram of the polymer hits.

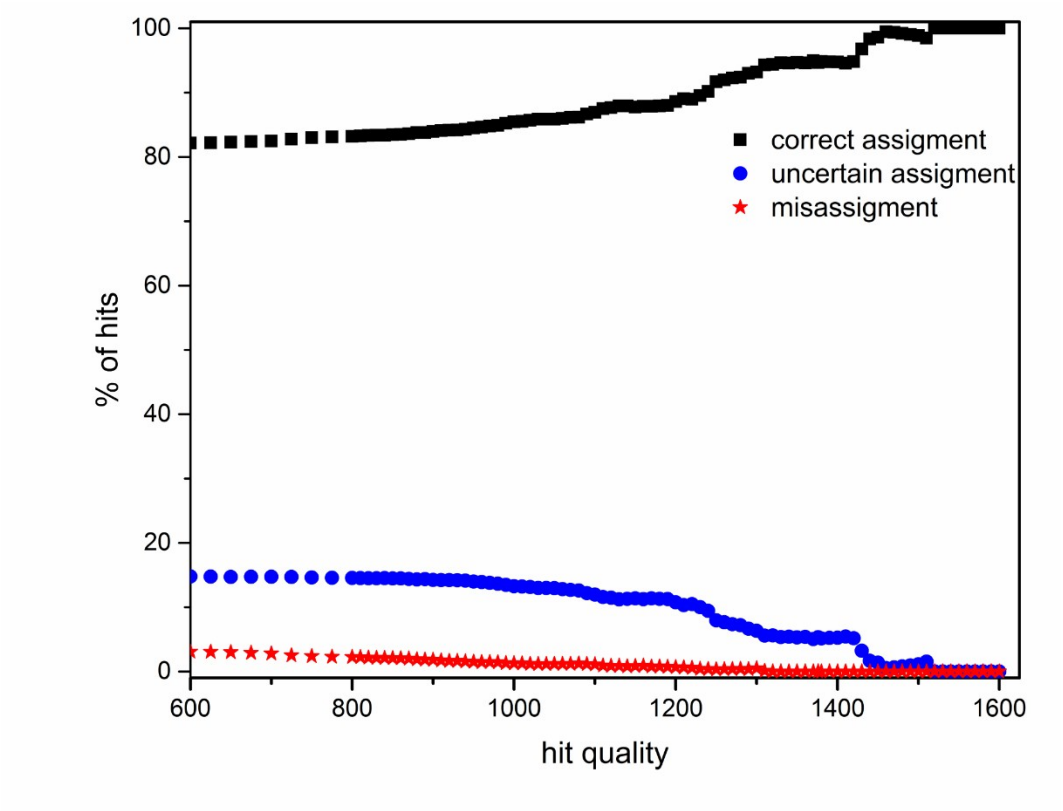


Fig. S6: Confidence interval at different hit quality thresholds by reanalysis of all database hits for the combined routines II and III with certain, uncertain or misassignments.

Data analysis for correlation between hit quality and manual reanalysis

The dataset derived from OPUS was further investigated for a correlation between hit quality and manual certainty assignment from the individual reanalysis to determine the impact of its loss for analytical steps in the future. For the exemplary sample polymer assignments for 34 database entries identified. Only datasets with more than 100 analysed spectra were chosen for a representative analysis (see Table S3).

In summary twelve suitable datasets were found of which eight show a good resemblance of the overall result. Assignments for PP and PS yield highly better results than the average while for PE and PVC show a higher abundances of misassignments. PVC alone contributes nearby 33 % to the total misassignments and direct correlation between the certainty categories and hit quality was found. For hit quality above 1000 only one assignment from 71 was false and below only one of 78 hit was a certain assigment (see Figure S7). This finding indicates that the hit quality is necessary for later analysis. A threshold can be determined for each polymer type and applied afterwards. For PE a similar result was indicated, whereas 11 % were false

assigned which did not show a clear trend for a hit quality dependence (see Figure S8).

Tab. S3: Polymer hits after manual validation stating the hit quality with different values: 1 for a perfect hit; 0.75 for a secure hit; 0.5 for a hit limited by the S/N ratio or the spectral database; 0.25 for a minimal hit; 0.01 for a complete misinterpretation.

Polymer	Hits	Correct assignment	Uncertain assignment	Misassigment
PE	338	71.6	17.6	11.0
PP	1125	90.1	9.2	0.9
Viscose/elastane	101	70.3	28.7	1.0
Chitin	568	71.1	28.2	0.7
Varnish Typ 1	1880	90.5	8.9	0.6
Varnish Typ 2	311	99.7	0.3	-
ABS rubber	100	69.0	29.0	2.0
Polystyrene	195	97.9	1.5	0.5
Polyvinylchloride	152	41.4	13.8	44.7
Wool	1108	77.8	17.8	4.4
PPS	290	93.4	6.2	0.3
EPDM	140	50.7	47.9	1.4

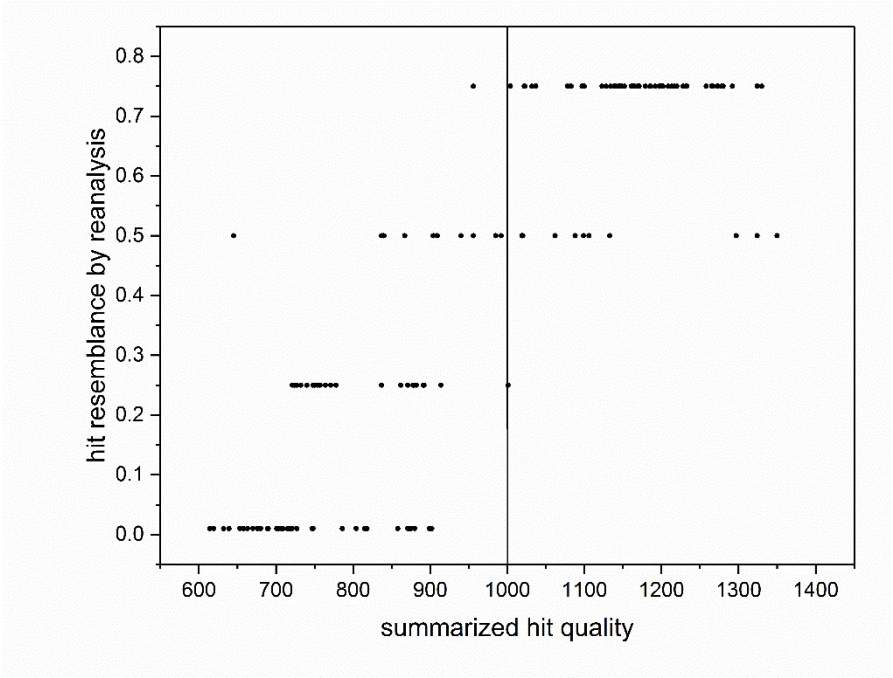


Fig. S7: Hit quality by automated analysis plotted versus the result of manual reanalysis for PVC. The vertical line at 1000 was introduced for optical orientation.

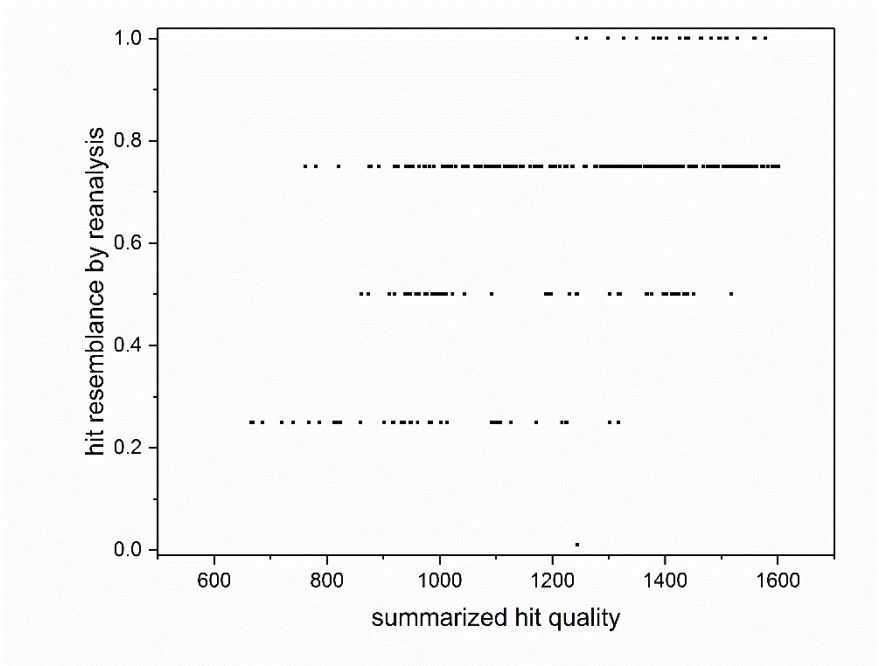


Fig. S8: Hit quality by automated analysis plotted versus the result of manual reanalysis for PE.

OPUS© Macro

General info: The library search need the actual information where the database is placed. Both underlined areas have to be updated with the pathways and file names on the actual system.

LibrarySearchSpectrum ([<File 3>:AB], 0, {LSS=6, SSQ=300, SSH=30, SS1=24584, LB1='D:\Bibliothek\POLYMERSPEKTREN.S01@', MPP=0, LBL='D:\Bibliothek', SSO=0, UFL=0, FRG='3059.2 2419.3 2275.1 0', URG=0, LNC=3});

Single methods:

Routine	SS1= (Parameter in LibrarySearchSpectrum)
I	24584
II	148
III	164
IV	196
V	276
VI	292
VII	324

VARIABLES SECTION

FILE <File 1> = AB/Multiple, Spec/Multiple, Trace/Multiple;

FILE <File 2> = Spec;

FILE <File 3> = AB, AB/Search, AB/Integ;

STRING <OpusPath> = '';

STRING <WorkPath> = '';

STRING <DataPath> = '';

STRING <Filename> = '';

NUMERIC <Loopindex0> = 0;

STRING <texttoexport1> = '';

STRING <texttoexport12> = '';

STRING <gesamtext> = '';

STRING <textdatei> = '';

NUMERIC <Max_Spectra> = 0;

NUMERIC <NULL> = 0;

NUMERIC <hit0> = 0;

NUMERIC <one> = 1;

NUMERIC <FPAx> = 0;

NUMERIC <FPAy> = 0;

NUMERIC <binning64> = 0;

NUMERIC <PosKomma> = 0;

NUMERIC <x_row> = 0;

NUMERIC <x_experiment> = 0;

NUMERIC <y_experiment> = 0;

STRING <y_exp> = '';

```
BOOL <Restart> = FALSE;

STRING <backuptext> = '';

NUMERIC <written> = 0;

NUMERIC <writtentest> = 0;

NUMERIC <Folder> = 0;

STRING <WorkPath0> ='';

PROGRAM SECTION


REM The file for analysis will be chosen and input parameter for x and y calculation set

UserDialog ('Routine 1', STANDARD, FILE:[<File 1>:Spec/Multiple], TEXT:'Each spectrum will be compared', TEXT:'Input FPA Field in x and y direction for analysis', EDIT:'<FPAx>', EDIT:'<FPAy>', TEXT:'Input Binning for Analysis', EDIT:'<binning64>', BLANK, BLANK, BLANK, TEXT:'Restart after interruption ?', CHECKBOX:'<Restart>');

REM Determination of file details and data path for documentation

<Max_Spectra> = GetParameter ([<File 1>:Trace/Multiple], NPT);

<OpusPath> = GetOpusPath ();

<DataPath> = GetParameter ([<File 1>:Spec/Multiple], PAT);

<Filename> = GetParameter ([<File 1>:Spec/Multiple], NAM);

<WorkPath> = '<DataPath>\Extracts_autoanalysis';

<WorkPath0> = '<WorkPath>';

REM Check backup file if wished or start a new one

If (<Restart>, .EQ., TRUE);

<backuptext> = ReadTextFile ('<DataPath>\backup');

<Loopindex0> = <backuptext>[0];

<written> = <backuptext>[1];

<Folder> = <backuptext>[2];

Else ();

TextToFile ('<DataPath>', 'backup', '0', REPLACE_TEXT);

TextToFile ('<DataPath>', 'backup', '0', APPEND_TEXT);

TextToFile ('<DataPath>', 'backup', '0', APPEND_TEXT);

REM Header of the textfile

<gesamtext> = 'Analysed File:;<Filename>';

TextToFile ('<DataPath>', '<Filename>.txt', '<gesamtext>', APPEND_TEXT);

<gesamtext> = '';

Endif ();

REM Determination of amount of spectra per x row

<x_row> = <FPAx> * ( 64 / <binning64>);

StartLoop (<Max_Spectra>, 0);

If (<Loopindex0>, .GT., <Max_Spectra>);

Goto (EndeExtrahieren);

Endif ();

StaticMessage (SHOW, {'<Loopindex0>/<Max_Spectra>'});
```

```
TextToFile ('<DataPath>', 'backup', '<Loopindex0>', REPLACE_TEXT);

TextToFile ('<DataPath>', 'backup', '<written>', APPEND_TEXT);

TextToFile ('<DataPath>', 'backup', '<Folder>', APPEND_TEXT);

PostrunExtract ([<File 1>:AB/Multiple], {EXS=0, EXE=0, ENT=0, ENE=0, ECO=0, XTP='<WorkPath>', XTN='<Filename>_<Loopindex0>.0',
XTI=0, EAB=0, ELF=0, EIL='<Loopindex0>', EUL=1, EDM=1});

<File 3> = Load (00, {COF=64, DAP='<WorkPath>', DAF='<Filename>_<Loopindex0>.0', INP='E:\OPUS_7.5\METHODS',
IFP='E:\OPUS_7.5\METHODS', INM='DEFAULT.TXD', IFN='DEFAULT', LB0='SNM CNM NPT', LB6='|| |Load 1 E:\OPUS_7.5\WORK|
|LibrarySearchSpectrum 1| |RunMacro 1| |LoadDataPointTable 1| '});

LibrarySearchSpectrum ([<File 3>:AB], 0, {LSS=6, SSQ=300, SSH=30, SS1=24584, LB1='D:\Bibliothek\POLYMERSPEKTREN.S01@', MPP=0,
LBL='D:\Bibliothek', SSO=0, UFL=0, FRG='3059.2 2419.3 2275.1 0', URG=0, LNC=3});

REM Report is written into textfile

REM Export of Hit 1

<texttoexport1> = FromReportMatrix ([<File 3>:AB/Search], 1, 0, 1, 2);

<hit0> = <texttoexport1>;

REM Check if there is a hit. If not delete file and goto next one

If ('<hit0>', .EQ., <NULL>);

Unload ([<File 3>:AB/Integ/Search], {});

Delete ('<WorkPath>\<Filename>_<Loopindex0>.0');

<texttoexport1> = '';

Goto(NextLoop);

Endif();

<texttoexport12> = FromReportMatrix ([<File 3>:AB/Search], 1, 0, 1, 3);

REM Determination of x and y Positon

<y_experiment> = (<Loopindex0> / <x_row>);

If (<y_experiment>, .EQ., 0);

Else();

<y_exp> = '<y_experiment>';

<PosKomma> = FindString ('<y_exp>', '.', NOCASE);

<y_experiment> = '<[0,<PosKomma>]y_exp>';

Endif();

<x_experiment> = <Loopindex0> - (<x_row>*<y_experiment>);

REM Preparation of output line

<gesamtext> = '<Loopindex0>;<x_experiment>;<y_experiment>;<hit0>;<texttoexport12>';

TextToFile ('<DataPath>', '<Filename>.txt', '<gesamtext>', APPEND_TEXT);

REM XYZ Output

If (<Calculation>, .EQ., TRUE);

<gesamtext> = '';

<gesamtext> = '<x_experiment>;<y_experiment>;<texttoexport1>';

TextToFile ('<DataPath>', '<texttoexport12><polymer>xyz.txt', '<gesamtext>', APPEND_TEXT);

Endif ();

REM Reset of Variables

<texttoexport1> = '0';
```

<texttoexport12> = '';

<gesamtext> = '';

<written> = <written> + 1;

Save ([<File 3>:AB/Search], {DAP='<WorkPath>', OEX='1', SAN='<Filename>_<Loopindex0>.0'});

Unload ([<File 3>:AB/Integ/Search], {});

REM OPUS cannot handle folders with more then 21000 files, therefore each 21000 hit a new folder is created.

<writtentest> = <written> / 21000;

If (<writtentest>, .EQ., <one>);

<WorkPath> = '<WorkPath0>_<Folder>';

<Folder> = <Folder> + 1;

<written> = 0;

Endif();

Label (NextLoop);

<Loopindex0> = <Loopindex0>+1;

EndLoop (0);

Label (EndeExtrahieren);

StaticMessage (HIDE, {});

Label(Ende);

PARAMETER SECTION

Full filter analysis

VARIABLES SECTION

FILE <File 1> = AB/Multiple, Spec/Multiple, Trace/Multiple;

FILE <File 2> = Spec;

FILE <File 3> = AB, AB/Search, AB/Integ;

STRING <OpusPath> = '';

STRING <WorkPath> = '';

STRING <DataPath> = '';

STRING <Filename> = '';

NUMERIC <Loopindex0> = 0;

STRING <texttoexport1> = '';

STRING <texttoexport12> = '';

STRING <texttoexport2> = '';

STRING <texttoexport22> = '';

STRING <gesamtext> = '';

```
STRING <textdatei> = '';

NUMERIC <Max_Spectra> = 0;

NUMERIC <NULL> = 0;

NUMERIC <hit0> = 0;

NUMERIC <hit1> = 0;

NUMERIC <zweihits> = 0;

NUMERIC <one> = 1;

NUMERIC <FPAx> = 0;

NUMERIC <FPAy> = 0;

NUMERIC <binning64> = 0;

NUMERIC <PosKomma> = 0;

NUMERIC <x_row> = 0;

NUMERIC <x_experiment> = 0;

NUMERIC <y_experiment> = 0;

STRING <y_exp> = '';

NUMERIC <Threshold> = 600;

BOOL <Restart> = FALSE;

STRING <backuptext> = '';

NUMERIC <written> = 0;

NUMERIC <writtentest> = 0;

NUMERIC <Folder> = 0;

STRING <WorkPath0> ='';
```

PROGRAM SECTION

```
UserDialog ('Makro for image analysis file, STANDARD, FILE:[<File 1>:Spec/Multiple], TEXT:'Each spectrum will be analyszed', TEXT: 'Input FPA Field in x and y direction for analysis Auswertung', EDIT:'<FPAx>', EDIT:'<FPAy>', TEXT:'Input Binning for analysis', EDIT:'<binning64>', TEXT:'Higher Threshold necessary (Standard is 600)', EDIT:'<Threshold>', TEXT:'Each file will be documented in a textfile', BLANK, BLANK, TEXT: 'Restart after interruption ?', CHECKBOX:'<Restart>');

REM Determination of file details and data path for documentation

<Max_Spectra> = GetParameter ([<File 1>:Trace/Multiple], NPT);

<OpusPath> = GetOpusPath ();

<DataPath> = GetParameter ([<File 1>:Spec/Multiple], PAT);

<Filename> = GetParameter ([<File 1>:Spec/Multiple], NAM);

<WorkPath> = '<DataPath>\Extracts_autoanalysis';

<WorkPath0> = '<WorkPath>';

REM Backupfile will be generated or loaded into the program

If (<Restart>, .EQ., TRUE);

<backuptext> = ReadTextFile ('<DataPath>\backup');

<Loopindex0> = <backuptext>[0];

<written> = <backuptext>[1];

<Folder> = <backuptext>[2];
```

```
Else ();

TextToFile ('<DataPath>', 'backup', '0', REPLACE_TEXT);

TextToFile ('<DataPath>', 'backup', '0', APPEND_TEXT);

TextToFile ('<DataPath>', 'backup', '0', APPEND_TEXT);

REM Fileheader is written

<gesamtext> = 'x;y;g;p';

TextToFile ('<DataPath>', '<Filename>.csv', '<gesamtext>', APPEND_TEXT);

<gesamtext> = '';

Endif ();

REM Determination of amount of spectra per x row

<x_row> = <FPAx> * ( 64 / <binning64>);

REM Start of analysis

StartLoop (<Max_Spectra>, 0);

If (<Loopindex0>, .GT., <Max_Spectra>);

Goto (EndeExtrahieren);

Endif ();

StaticMessage (SHOW, {'<Loopindex0>/<Max_Spectra>'});

REM Backup of actual postion

TextToFile ('<DataPath>', 'backup', '<Loopindex0>', REPLACE_TEXT);

TextToFile ('<DataPath>', 'backup', '<written>', APPEND_TEXT);

TextToFile ('<DataPath>', 'backup', '<Folder>', APPEND_TEXT);

REM Spectrum is extracted

PostrunExtract ([<File 1>:AB/Multiple], {EXS=0, EXE=0, ENT=0, ENE=0, ECO=0, XTP='<WorkPath>', XTN='<Filename>_<Loopindex0>.0',
XTI=0, EAB=0, ELF=0, EIL='<Loopindex0>', EUL=1, EDM=1});

<File 3> = Load (00, {COF=64, DAP='<WorkPath>', DAF='<Filename>_<Loopindex0>.0', INP='F:\OPUS_7.5\METHODS',
IFP='F:\OPUS_7.5\METHODS', INM='DEFAULT.TXD', IFN='DEFAULT', LB0='SNM CNM NPT', LB6='|| |Load 1 F:\OPUS_7.5\WORK|
|LibrarySearchSpectrum 1| |RunMacro 1| |LoadDataPointTable 1| '});

REM Spectral search routine II

LibrarySearchSpectrum ([<File 3>:AB], 0, {LSS=6, SSQ=300, SSH=30, SS1=148, LB1='D:\Bibliothek\POLYMERSPEKTREN.S01@', MPP=0,
LBL='D:\Bibliothek', SSO=0, UFL=0, URG=0, LNC=3});

REM Export of Hit 1

<texttoexport1> = FromReportMatrix ([<File 3>:AB/Search], 1, 0, 1, 2);

<hit0> = <texttoexport1>;

REM Is a hit found?

If ('<hit0>', .EQ., <NULL>);

REM If not, delete file and start next loop

Unload ([<File 3>:AB/Integ/Search], {});

Delete ('<WorkPath>\<Filename>_<Loopindex0>.0');

<texttoexport1> = '';

Goto(NextLoop);

Endif();

<texttoexport12> = FromReportMatrix ([<File 3>:AB/Search], 1, 0, 1, 3);
```

```
REM Spectral search routine III

LibrarySearchSpectrum ([<File 3>:AB], 0, {LSS=6, SSQ=300, SSH=30, SS1=164, LB1='D:\Bibliothek\POLYMERSPEKTREN.S01@', MPP=0,
LBL='D:\Bibliothek', SSO=0, UFL=0, URG=0, LNC=3});

<texttoexport2> = FromReportMatrix ([<File 3>:AB/Search], 1, 0, 1, 2);

<hit1> = <texttoexport2>;

REM Is a hit found?

If ('<hit1>', .EQ., <NULL>);

REM If not, delete file and start next loop

Unload ([<File 3>:AB/Integ/Search], {});

Delete ('<WorkPath>\<Filename>_<Loopindex0>.0');

<texttoexport2> = '';

Goto(NextLoop);

Endif();

<texttoexport22> = FromReportMatrix ([<File 3>:AB/Search], 1, 0, 1, 3);

REM Is the same polymer found?

If ('<texttoexport12>', .EQ., '<texttoexport22>');

<zweihits> = <hit0>+<hit1>;

REM Is the combined hit higher as the threshold?

If ('<zweihits>', .GTEQ., '<Threshold>');

Goto(Auswerten);

Else();

REM If not, delete the file, reset variables and start next loop

Unload ([<File 3>:AB/Integ/Search], {});

Delete ('<WorkPath>\<Filename>_<Loopindex0>.0');

<texttoexport1> = '0';

<texttoexport12> = '';

<texttoexport2> = '0';

<texttoexport22> = '';

<zweihits>=0;

Goto(NextLoop);

Endif();

Else();

REM If not, delete the file, reset variables and start next loop

Unload ([<File 3>:AB/Integ/Search], {});

Delete ('<WorkPath>\<Filename>_<Loopindex0>.0');

<texttoexport1> = '0';

<texttoexport12> = '';

<texttoexport2> = '0';

<texttoexport22> = '';

Goto(NextLoop);

Endif();
```

```
REM Preparation of file entry
Label(Auswerten);
REM Determination of the x and y position
<y_experiment> = (<Loopindex0> / <x_row>);
If (<y_experiment>, .EQ., 0);
Else();
<y_exp> = '<y_experiment>';
<PosKomma> = FindString ('<y_exp>', '.', NOCASE);
<y_experiment> = '<[0,<PosKomma>]y_exp>';
Endif();
<x_experiment> = <Loopindex0> - (<x_row>*<y_experiment>);
REM File entry
<gesamtext> = '<x_experiment>;<y_experiment>;<zweihits>;<texttoexport12>';
TextToFile ('<DataPath>', '<Filename>.csv', '<gesamtext>', APPEND_TEXT);
REM Reset of variables
<texttoexport1> = '0';
<texttoexport12> = '';
<texttoexport2> = '0';
<texttoexport22> = '';
<zweihits>=0;
<hit0>=0;
<hit1>=0;
<gesamtext> = '';
<written> = <written> + 1;
REM File is saved with SEARCH Report
Save ([<File 3>:AB/Search], {DAP='<WorkPath>', OEX='1', SAN='<Filename>_<Loopindex0>.0'});
Unload ([<File 3>:AB/Integ/Search], {});
REM OPUS cannot handle folders with more then 21000 files, therefore each 21000 hit a new folder is created.
<writtentest> = <written> / 21000;
If (<writtentest>, .EQ., <one>);
<WorkPath> = '<WorkPath0>_<Folder>';
<Folder> = <Folder> + 1;
<written> = 0;
Endif();
Label (NextLoop);
<Loopindex0> = <Loopindex0>+1;
EndLoop (0);
Label (EndeExtrahieren);
StaticMessage (HIDE, {});
Label(Ende);
```

Image Analysis

Installation of Python for Image Analysis: We recommend for application of the image analysis tool the installation of Anaconda (https://www.continuum.io). After installation use the Anaconda tools to install Python 3.4 in an new environment (in this example named sitk) and the following packages: numpy, pandas, matplotlib, ipython and sitk.

To install all packages open the windows command prompt use the following steps:

activate sitk

conda install numpy

conda install pandas

conda install matplotlib

conda install ipython

conda install –c https://conda.anaconda.org/simpleitk SimpleITK

Copy the python files in a folder of your choice and create a short cut on the desktop:

Target: C:\Anaconda3\envs\sitk\python.exe C:\(chosen file path)\ MP_analysis_from_OPUS_macros.py

Execute in C:\(chosen file path)

Click on the shortcut to start the script. On file name a .csv file from the OPUS© macro can be chosen and under result dir an existing subfolder for the file export of the script. The files generated from the macro should be in the same folder as the scripts. Start the analysis by clicking start. After analysis the different files are located the chosen folder for the results.