

## Supporting information

### Extraction optimization and pixel-based chemometric analysis of semi-volatile organic compounds in groundwater

Peter Christensen <sup>1\*</sup>, Giorgio Tomasi<sup>1</sup>, Mette Kristensen<sup>1</sup>, Peter E. Holm<sup>1</sup> and Jan H. Christensen<sup>1</sup>

<sup>1</sup> *Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen, Thorvaldsensvej 40,  
1871 Frederiksberg C, Denmark*

\* Corresponding author telephone: +45 35 33 35 23; e-mail: pech@plen.ku.dk

A total of 15 groundwater samples collected from five locations in Denmark were analyzed in this study. A description of the sampling location is provided in Table S-1.

**Table S-1:** Description of sampling sites: DGU no. for individual sampling wells, depth of inlet pipe in meters, location, UTM coordinates and main land-use are listed. Further information about individual wells can be found in the Jupiter database (via DGU no.) at the GEUS website ([www.geus.dk](http://www.geus.dk)). Main land use is identified at a radius of 800 m from the sampling site by the GIS software (digital land use map (1:25,000)) [1].

Sample	DGU no.	inlet pipe (depth in m)	Location/site	UTM (32 or 33*) coordinates	Main land use in %		
					Agriculture	Urban	Forest
Fr1	146.2652	1 (10.5-11.5)	Freltofte	(590891.01, 6126255.82 )	82.5	8.82	0.731
Fr2	146.2653	1 (17 – 18)	Freltofte	(591890.9, 6126293.88)	85.2	9.18	0.110
Fr3	146.2556	1 (10.5-11.5)	Freltofte	(591808.5, 6126485.6)	83.6	9.75	0.470
Th1 †	30.935	1 (32-34)	Thisted	(474225, 6318734)	46.3	3.46	0.448
Th2	30.933	3 (22-24)	Thisted	(477360, 6316747)	75.1	10.2	4.40
Th4	30.933	1 (48-50)	Thisted	(477360, 6316747)	75.1	10.2	4.40
Jy1	190.279	1 (11.5-12.5)	Jyderup Skov	(658402, 6191768)	15.8	0.508	77.9
Jy5	190.272	3 (7-9)	Jyderup Skov	(657844, 6191231)	30.6	0.896	45.9
Jy6	190.278	1 (11.7-12.7)	Jyderup Skov	(658675.05, 6192526.28)	9.81	4.83	52.4
Sm1	247.382	4 (6-20.5)	Smaalyng	(501952.25, 6100839.29)*	76.9	1.64	20.0
Sm3	247.380	4 (14.3-14.4)	Smaalyng	(502047.97, 6100517.17)*	81.7	1.63	15.4
Sm4	247.618	1 (7.8-8.8)	Smaalyng	(502177.5, 6100318.63)*	85.5	1.57	12.3
Ha1	65.1522	1 (3.7-4.7)	Haderup	(497556.7, 6249263.7)	90.1	2.26	5.54
Ha2	65.1523	1 (3.5-4.5)	Haderup	(497971.6, 6249313.01)	96.2	1.67	0.769
Ha3	65.1520	1 (4.5-5.5)	Haderup	(498194.6, 6248799.7)	79.8	1.46	18.7

† The main land use of Th1 also includes 42.1 % water.

A total of 39 compounds were selected for calculation of the extraction efficiency using SPE-DLLME-GC-MS and SPE-SBSE-GC-MS (Table S-2). Numbering, molecular weight, chemical formula, CAS number, mass spectral properties, retention time on a 60 m ZB5 column, and NIST score for the compounds are listed in Table S-2. The compounds that were expected to be suitable for multiple extractions with SBSE and DLLME were selected for the table. Some methyl esters, dimethyl acetals and alkanes were not added to the table since these groups were already represented. In general, compounds with a NIST score  $\geq 90\%$  are included. Additionally, acids and methyl esters are included to follow the results of derivatization and some polar compounds are included to cover a wide polarity range in the assessment of extraction efficiency. Unfortunately, the glassdistilled water (controls) were contaminated with semi-VOCs from the laboratory air, silicone tubing and glass equipment. Thus, only 19 of the 39 compounds were found in higher concentrations (peak area) in the groundwater samples than in the controls. These compounds are marked with an asterisk (\*).

**Table S-2:** Compounds found in Danish groundwater samples and identified by the NIST database. Compounds with higher average peak area in groundwater samples than in glass distilled controls are marked with an asterisk (\*). Listed is peak number, compound, chemical formula, molecular weight (MW), retention time (RT), target ion, qualifier ions, and NIST qualification score.

No.	Compound	Formula	MW	CAS no.	RT, min	Target ion	Qual. ion	NIST
1*	Benzoic acid methyl ester	C <sub>8</sub> H <sub>8</sub> O <sub>2</sub>	136	93-58-3	11.69	136	77 / 105	93
2	Octanoic acid methyl ester	C <sub>9</sub> H <sub>18</sub> O <sub>2</sub>	158	111-11-5	11.85	74	87 / 127	91
3	Octanoic acid	C <sub>8</sub> H <sub>16</sub> O <sub>2</sub>	144	124-07-2	12.22	60	73 / 101	50
4*	Benzoic acid	C <sub>7</sub> H <sub>6</sub> O <sub>2</sub>	122	65-85-0	12.27	122	77 / 105	96
5	Decanal	C <sub>10</sub> H <sub>20</sub> O	156	112-31-2	12.74	41	57 / 112	98
6	Nonanoic acid methyl ester	C <sub>10</sub> H <sub>20</sub> O <sub>2</sub>	172	1731-84-6	12.88	74	87 / 141	95
7	Benzothiazole	C <sub>7</sub> H <sub>5</sub> NS	135	95-16-9	13.21	135	91 / 108	75
8	Nonanoic acid	C <sub>9</sub> H <sub>18</sub> O <sub>2</sub>	158	112-05-0	13.23	60	73 / 115	87
9	Nonanal dimethyl acetal	C <sub>11</sub> H <sub>24</sub> O <sub>2</sub>	188	18824-63-0	13.38	75	83 / 157	91
10	Decanoic acid	C <sub>10</sub> H <sub>20</sub> O <sub>2</sub>	172	334-48-5	14.14	60	73 / 129	83
11	Decanal dimethyl acetal	C <sub>12</sub> H <sub>26</sub> O <sub>2</sub>	202	7779-41-1	14.31	75	83 / 171	86
12	5-Bromo-4-hydroxy-m-cymene	C <sub>10</sub> H <sub>13</sub> BrO	228	100024-50-3	14.79	213	134 / 228	90
13	Octanedioic acid dimethyl	C <sub>10</sub> H <sub>18</sub> O <sub>4</sub>	202	1732-09-8	14.92	74	129 / 171	91

	ester							
14*	2,4-Di-tert-butylphenol	C <sub>14</sub> H <sub>22</sub> O	206	96-76-4	15.57	191	57 / 206	93
15	Dodecanoic acid methyl ester	C <sub>13</sub> H <sub>26</sub> O <sub>2</sub>	214	111-82-0	15.57	74	87 / 214	96
16	Nonanedioic acid dimethyl ester	C <sub>11</sub> H <sub>20</sub> O <sub>4</sub>	216	1732-10-1	15.77	74	152 / 185	91
17	Dodecanoic acid	C <sub>12</sub> H <sub>24</sub> O <sub>2</sub>	200	143-07-7	15.82	73	60 / 129	99
18*	Ibuprofen methyl ester	C <sub>14</sub> H <sub>20</sub> O <sub>2</sub>	220	114376-60-2	15.84	161	177 / 220	98
19*	Mecoprop methyl ester	C <sub>11</sub> H <sub>13</sub> ClO <sub>3</sub>	228	23844-56-6	15.93	228	142 / 169	99
20	Tributyl phosphate	C <sub>12</sub> H <sub>27</sub> O <sub>4</sub> P	266	126-73-8	16.57	99	155 / 211	93
21	2,6-Dibromo-4-(1,1-dimethylethyl)-phenol	C <sub>10</sub> H <sub>12</sub> Br <sub>2</sub> O	308	98-22-6	17.24	293	212 / 308	98
22	Undecanedioic acid dimethyl ester	C <sub>13</sub> H <sub>24</sub> O <sub>4</sub>	244	4567-98-0	17.31	74	98 / 213	91
23*	3,5-di-tert-Butyl-4-hydroxybenzaldehyde	C <sub>15</sub> H <sub>22</sub> O <sub>2</sub>	234	1620-98-0	17.69	219	191 / 234	98
24*	Dibenzothiophene-d <sub>8</sub>	C <sub>8</sub> D <sub>6</sub> S	192	33262-29-2	17.84	146 <sup>†</sup>	96 / 192	n.f.
25*	Phenanthrene-d <sub>10</sub>	C <sub>14</sub> D <sub>10</sub>	188	1517-22-2	18.03	160 <sup>†</sup>	80 / 188	n.f.
26	1,2-Benzenedicarboxylic acid bis(2-methylpropyl) ester	C <sub>16</sub> H <sub>22</sub> O <sub>4</sub>	278	84-69-5	18.27	205 <sup>†</sup>	57 / 223	91
27*	(Z)-9-Hexadecenoic acid methyl ester	C <sub>17</sub> H <sub>32</sub> O <sub>2</sub>	268	1120-25-8	18.4	236	74 / 194	99
28*	Hexadecanoic acid methyl ester	C <sub>17</sub> H <sub>34</sub> O <sub>2</sub>	270	112-39-0	18.52	74	87 / 270	98
29*	7,9-Di-tert-butyl-1-oxaspiro(4,5)deca-6,9-diene-2,8-dione	C <sub>17</sub> H <sub>24</sub> O <sub>3</sub>	276	82304-66-3	18.71	205	217 / 261	99
30*	Hexadecanoic acid	C <sub>16</sub> H <sub>32</sub> O <sub>2</sub>	256	57-10-3	18.75	256	60 / 73	93
31*	3,5-Bis(1,1-dimethylethyl)-4-hydroxybenzenepropanoic acid methyl ester	C <sub>18</sub> H <sub>28</sub> O <sub>3</sub>	292	6386-38-5	18.78	277	147 / 292	92
32*	Eicosane	C <sub>20</sub> H <sub>42</sub>	282	112-95-8	18.99	57	85 / 282	99
33*	(Z)-9-Octadecenoic acid methyl ester	C <sub>19</sub> H <sub>36</sub> O <sub>2</sub>	296	112-62-9	19.67	264	222 / 296	99
34*	Octadecanoic acid methyl ester	C <sub>19</sub> H <sub>38</sub> O <sub>2</sub>	298	112-61-8	19.79	74	87 / 298	99

35*	Octadecanoic acid	C <sub>18</sub> H <sub>36</sub> O <sub>2</sub>	284	57-11-4	20.00	284	60 / 73	98
36*	Docosane	C <sub>22</sub> H <sub>46</sub>	310	629-97-0	20.21	57	85 / 310	94
37	Triphenyl phosphate	C <sub>18</sub> H <sub>15</sub> O <sub>4</sub> P	326	115-86-6	21.82	326	170 / 233	94
38	Octicizer	C <sub>20</sub> H <sub>27</sub> O <sub>4</sub> P	362	1241-94-7	21.88	251	94 / 362	93
39*	1,2-Benzenedicarboxylic acid diisooctyl ester	C <sub>24</sub> H <sub>38</sub> O <sub>4</sub>	390	27554-26-3	22.54	279 †	167 / 149	91

† Less intense qualifier ions chosen due to overloading in the MS detector.

### Method development (supporting information to section 3.1)

The SBSE and DLLME sample preparation methods were optimized to give the largest peak area for the 39 semi-VOCs listed in Table S-2.

Extraction efficiency of the SBSE method was tested with and without salt addition (30 % w/v), and with acidified conditions at pH 2 (with 50 µL 2 M HCl). Salting out is generally acknowledged to result in increased extraction efficiency of polar compounds, e.g. for pesticides with  $\log K_{o/w} < 3.5$  [2]. Yang and Peppard (1994) studied the salting out effect on SBSE extraction for a range of flavor compounds. For most compounds, extraction efficiency increased with increasing salt concentration, but they also saw leveling off at high concentrations and a decrease in efficiency at high salt concentrations [3]. In our study, higher extraction efficiency of polar compounds was obtained with solutions containing salt, while the opposite trend was found for less polar compounds like methyl esters. Mono aliphatic carboxylic acids larger than C12 were more efficiently extracted without salt, whereas smaller ones were extracted most efficiently when salt was added. Acidic conditions resulted in the highest extraction efficiencies for carboxylic acids and high extraction efficiency (> 50 %) for non-hydrolysed compounds in the  $\log K_{o/w}$  range 2.86 - 7.03. Acidic conditions with no salt addition were therefore chosen as the optimal extraction method.

For DLLME, two extraction solvents (tetrachloromethane, CCl<sub>4</sub>, and tetrachloroethene, C<sub>2</sub>Cl<sub>4</sub>) and different salt concentrations (0, 6, 12, 18, 24, 30 % w/v) of the aqueous solution were tested. The addition of 30 % NaCl (w/v) to the extraction blanks resulted in NaCl precipitation. After centrifugation, this sedimented phase consisted of < 20 µL of extraction solvent mixed with > 1 g of NaCl, which made it difficult to transfer the extraction solvent to the GC-MS vials without NaCl crystals. On the other hand, the use of 0 % NaCl (w/v) with CCl<sub>4</sub> as extraction solvent led to precipitation of the dispersed droplets within two seconds. We therefore continue the DLLME

optimization using between 6 % and 24 % NaCl (w/v) combined with ultra-sonication for 2 min to stabilize the dispersion allowing enough time for DLLME.

The two extraction solvents: CCl<sub>4</sub> and C<sub>2</sub>Cl<sub>4</sub> both worked for the DLLME. However, CCl<sub>4</sub> showed to be the most efficient extraction solvent with larger peak areas and less chlorinated impurities. An increase in NaCl concentration from 6 to 24 % w/v had no significant effects on extraction efficiency (< 5 % change in peak area). As a larger film of extraction solvent was formed on the tube sides during centrifugation as a function of NaCl concentration it was therefore decided to use CCl<sub>4</sub> and an aqueous solution with 6 % w/v NaCl for the final method. To achieve a more efficient extraction of carboxylic acids, the pH was lowered to 2 by addition of 25 µL 2 M HCl.

#### *Unintended derivatization*

In this study, an unintended conversion of carboxylic acid and aldehyde functional groups, by reaction with methanol under acid condition, resulted in formation of the corresponding methyl esters and dimethyl acetals. This indicates that drying the SPE cartridges did not completely remove the HCl, resulting in acidic methanol extracts. A similar reaction was not observed for ketones.

#### *Electrospray mass spectrometry analysis (ESI-MS)*

One sample was analysed by direct injection negative electrospray mass spectrometry analysis (ESI-MS) and the spectrum is shown in Figure 1 in the main manuscript to show the presence of complex mixtures of fulvic- and humic acids in the SPE extract before additional DLLME or SBSE cleanup. The sample were injected into a Waters Acquity Ultra Performance Liquid Chromatograph system operating in column-bypass mode and coupled to a mass spectrometer. The mass spectrometer was a Micromass Ultima orthogonal acceleration quadrupole time-of-flight equipped with a Z-configured electrospray source (Waters Micromass). Source settings were as follows: capillary voltage: 2.0 kV (ES-), cone voltage 12 eV, source temperature 120 °C, desolvation temperature 250 °C, cone gas flow 50 L hr<sup>-1</sup>, desolvation gas flow 700 L hr<sup>-1</sup>. Nitrogen was used as cone and desolvation gas. Scan speed was 1 scan s<sup>-1</sup> and interscan delay 0.1 s. The QTOF was operated in scan mode from m/z 100 - 2000 and the quadrupole served as a focussing lens while operated in RF-only mode.

#### **Pixel-based analysis**

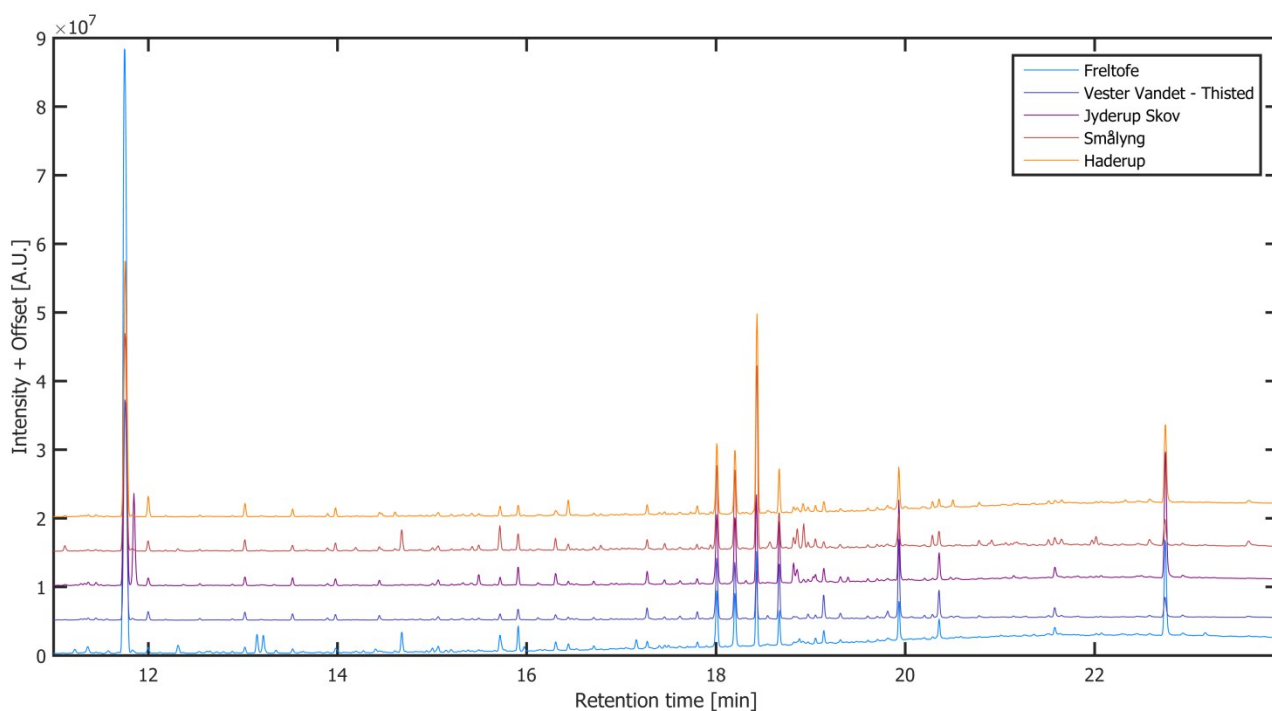
In this study, the pixel-based analysis consist of five steps: (1) Identification and removal of instrumental artifacts, (2) focusing the analysis on relevant information, (3) determination of the optimal signal processing steps, (4) multivariate data analysis, and (5) chemical interpretation. The raw data (total ion chromatograms, TIC's), extracted ion chromatograms (EIC's), and the different pre-processing steps are shown below.

The TIC's and EIC's, before any pre-processing, are shown in Figure S-1 and S-2. Summed extracted ion chromatograms (sEIC's) for nine mix samples (used as QC samples) are shown in Figure S-3. Baseline removal, shift, correlation optimized warping (COW), normalisation, and weighting are shown in Figure S-4 to S-8.

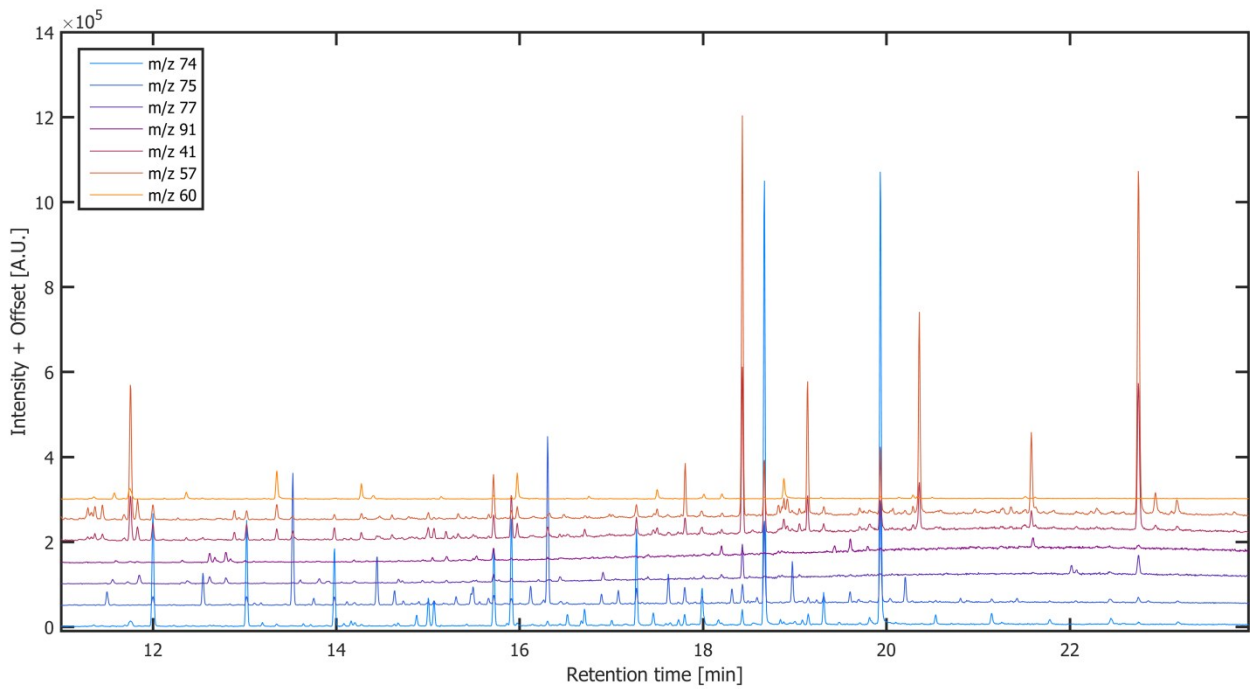
The retention time interval was cropped to include major peak of interest, which can be seen for the interval 18.38 - 18.48 min in figure S-6. Additional cropped sections included; 18.63 - 18.70 min, 19.88 - 19.98 min, and 22.70 - 22.80 min, and resulted in changed y-axis ranges.

One clear outlier of the Mix samples is seen in figure S-3 - S-7 (light blue). After all preprocessing steps, we see that this outlier is down scaled (Figure S-8).

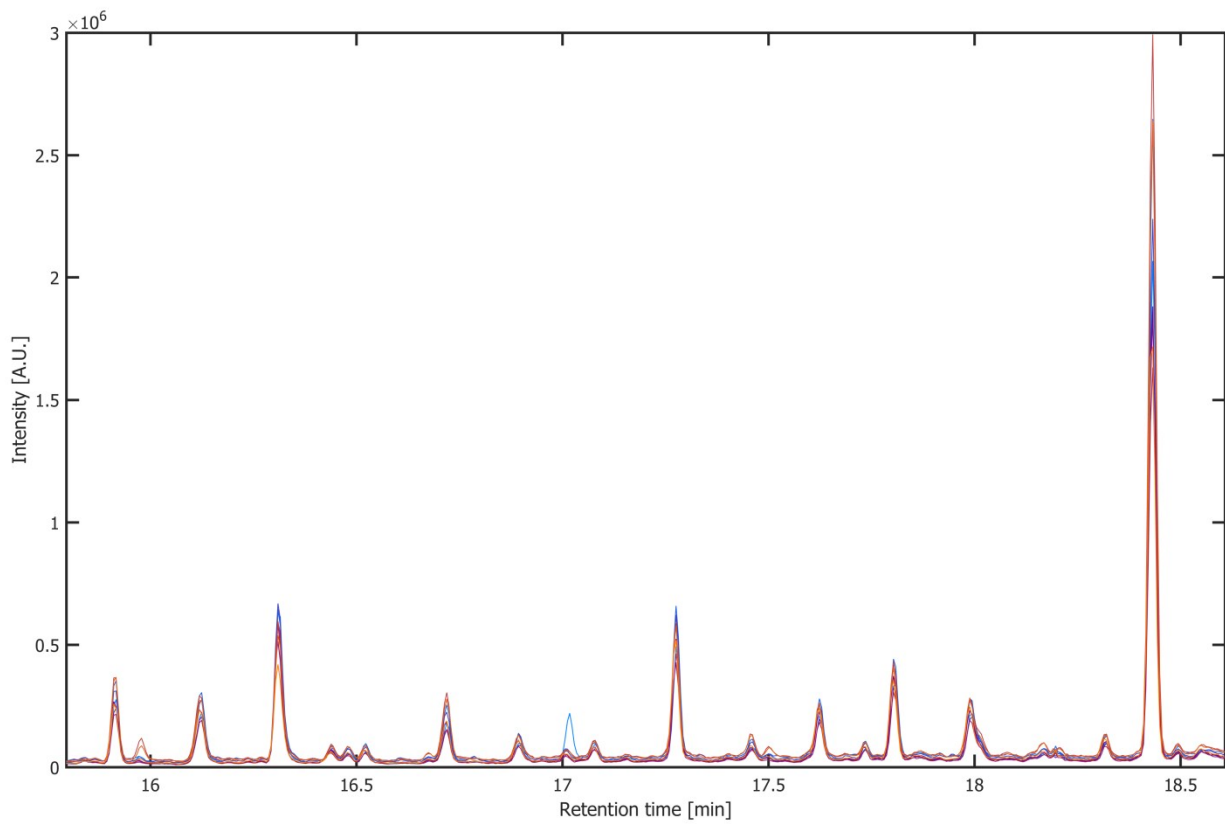
All samples are treated the same way as shown for the Mix samples in figure S-3 - S8.



**Figure S-1:** TIC for the five sample locations (one replicate of each location is shown), plotted with a y-axis offset of  $0.5 \cdot 10^7$  A.U.

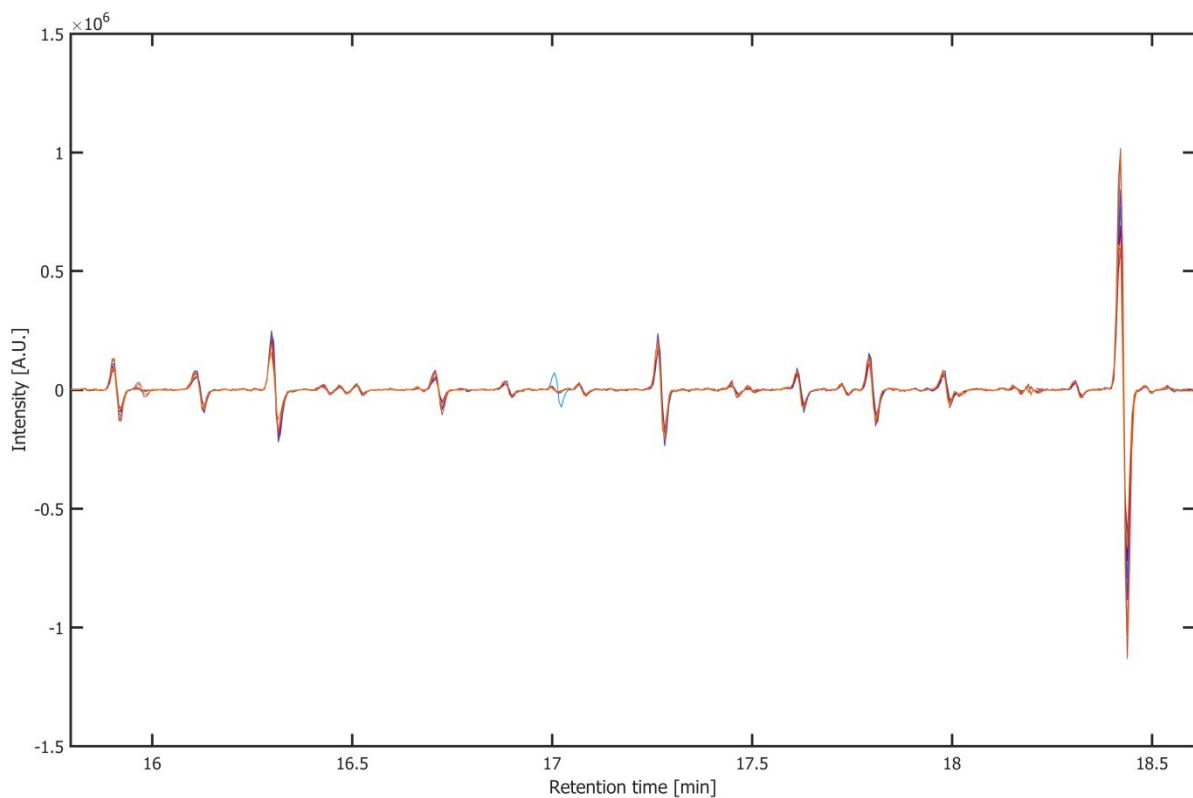


**Figure S-2:** EIC's of seven ions, m/z 41, 57, 60, 74, 75, 77, and 91 for one Freltofte sample (Freltofte 1). Plotted with a y-axis offset of  $1 \cdot 10^4$  A.U.

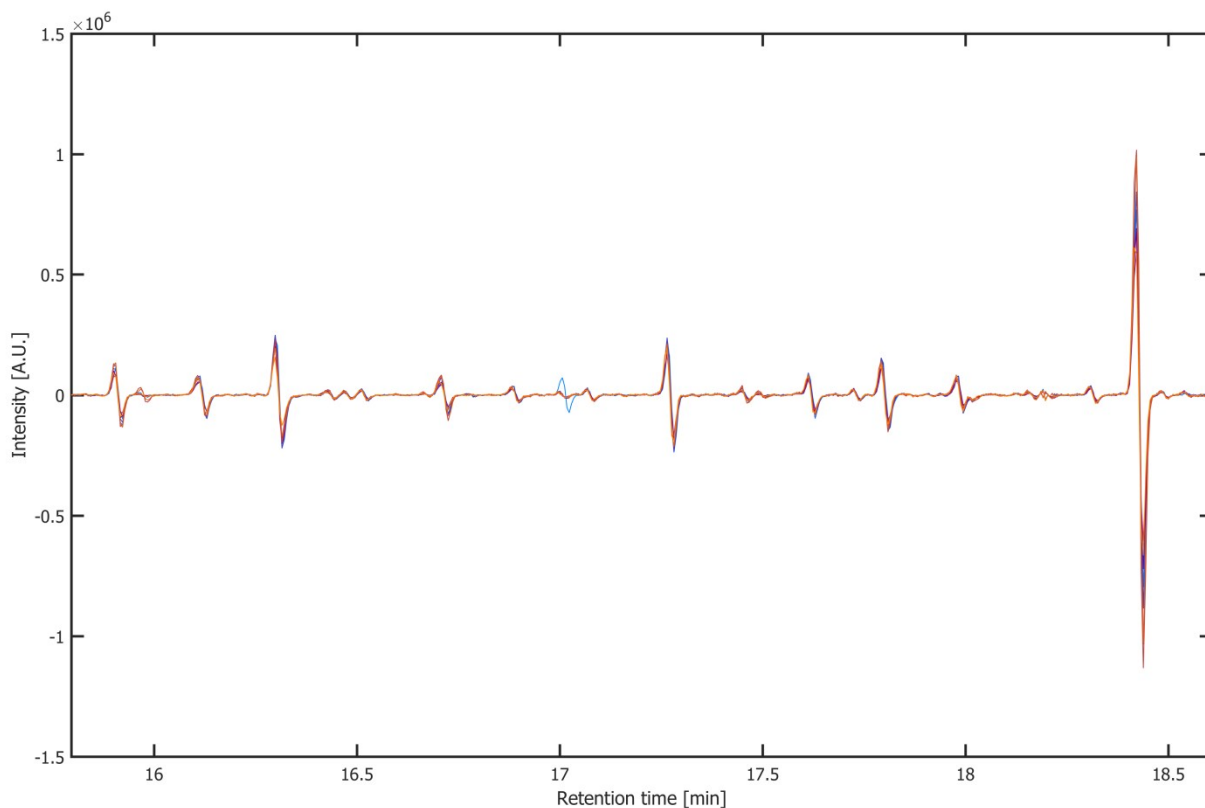


**Figure S-3:** sEIC's of nine Mix samples (QC). Zoom from 15.8 - 18.6 min.

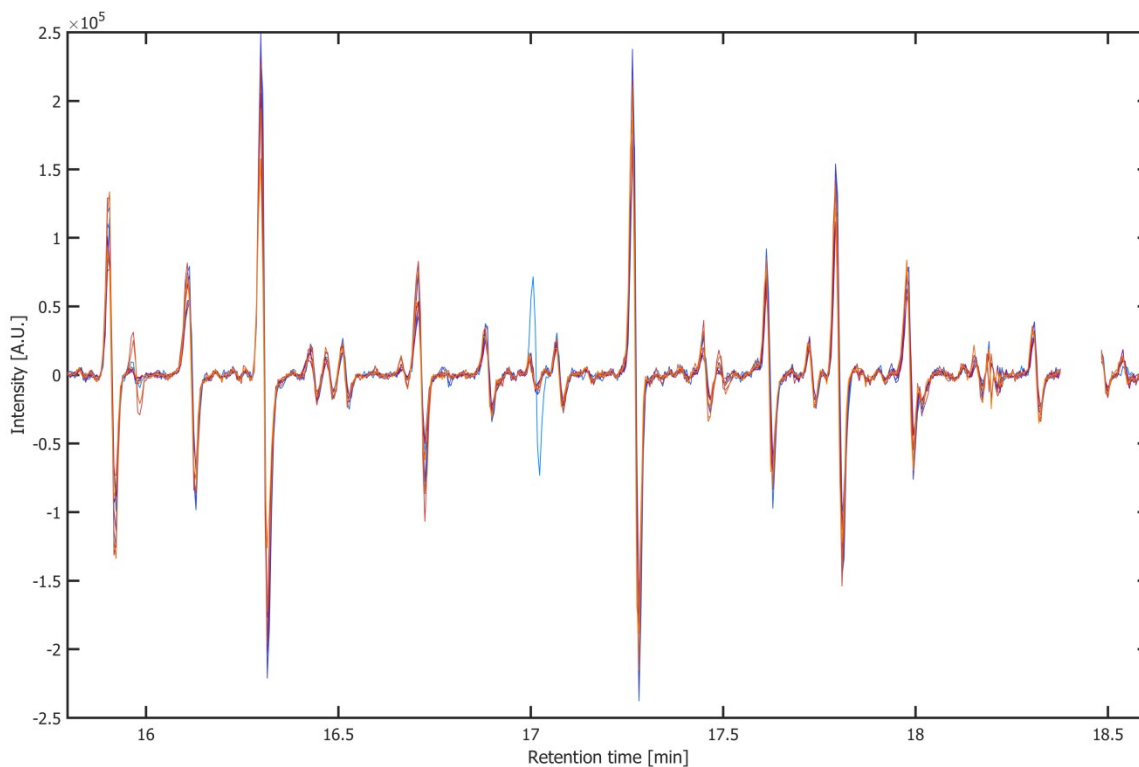




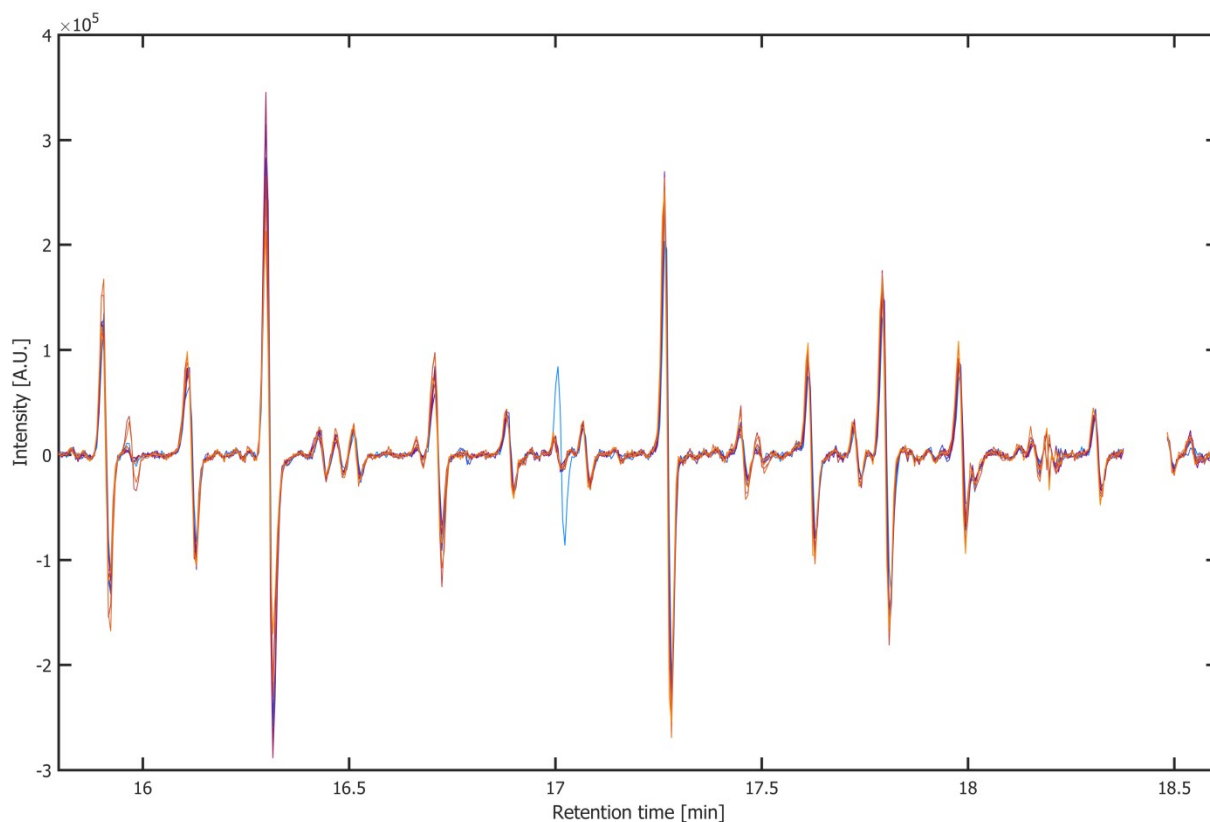
**Figure S-4:** Baseline removal by taking the first derivative of the measured response of nine Mix samples (QC). Zoom from 15.8 - 18.6 min.



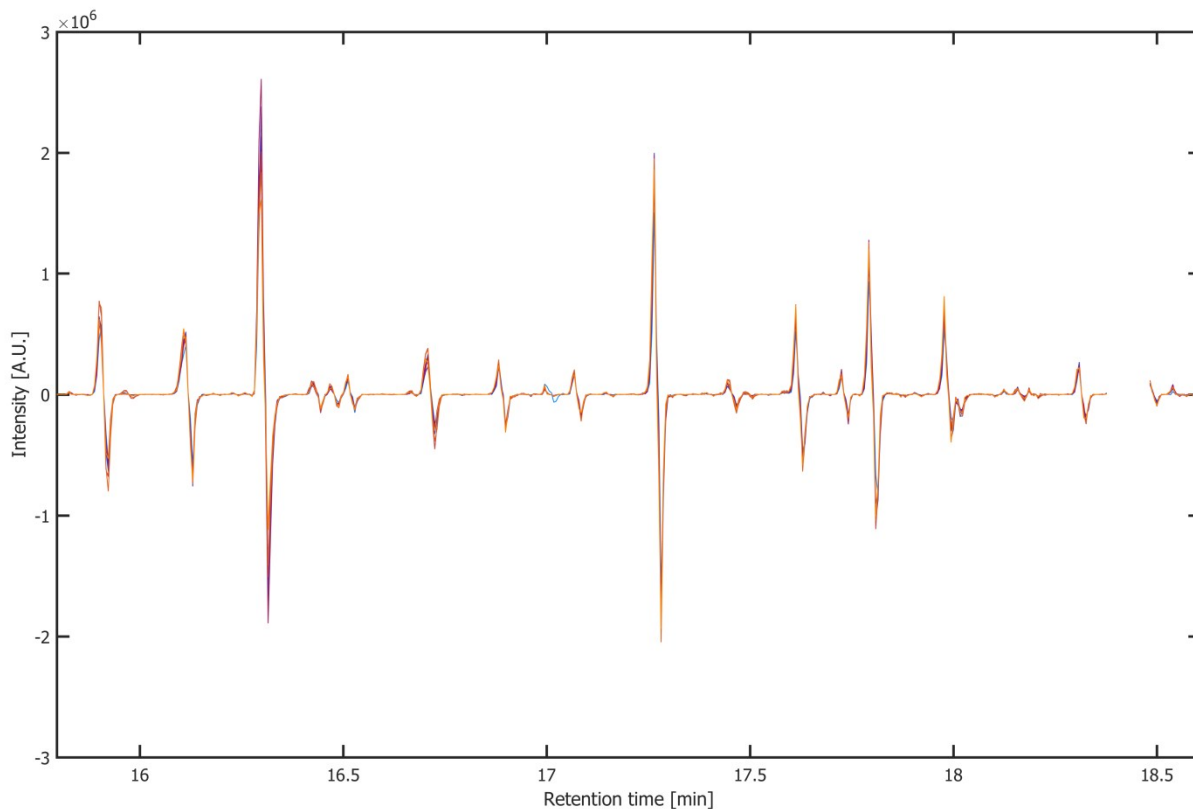
**Figure S-5:** Corrections for rigid shift in nine Mix samples (QC). Zoom from 15.8 - 18.6 min.



**Figure S-6:** Alignment by Correlation Optimal Warping (COW) of nine Mix samples (QC). Four intervals were aligned using iCOSHift (11.80-11.90 min, 18.35-18.50 min, 18.60-18.75 min, and 19.88-20.00 min). Zoom from 15.8 - 18.6 min.



**Figure S-7:** Nine Mix samples (QC) normalised to their Euclidean norm. Zoom from 15.8 - 18.6 min.



**Figure S-8:** Nine Mix samples (QC) weighted with the inverse of the relative standard deviations,  $RSD_{\text{mix}}^{-1}$ . Zoom from 15.8 - 18.6 min.

## References

- [1] Peter R. Jakobsen, Bjørn Hermansen & Lisbeth Tougaard (2011). Danmarks digitale jordartskort. Version 3.1. Danmarks og Grønlands Geologiske Undersøgelse Rapport 2011/40.
- [2] Nakamura, S.; Daishima, S.: Simultaneous determination of 64 pesticides in river water by stir bar sorptive extraction and thermal desorption-gas chromatography-mass spectrometry. *Anal. Bioanal. Chem.* **2005**, *382*, 99-107.
- [3] Yang, X. G.; Peppard, T.: Solid-phase microextraction for flavor analysis. *J. Agric. Food Chem.* **1994**, *42*, 1925-1930.