Use of the distributions of adamantane acids to profile short-term temporal and pond-scale spatial variations in the composition of oil sands process-affected waters

Sabine K. Lengger[a,*], Alan G. Scarlett[a], Charles E. West[a,#], Richard A. Frank[b], L. Mark Hewitt[b], Craig B. Milestone[b], Steven J. Rowland[a]

[a] Petroleum and Environmental Geochemistry Group, Biogeochemistry Research Centre Plymouth University, Drake Circus, Plymouth PL4 8AA

[b] Aquatic Contaminants Research Division/Water Science & Technology Directorate, Environment Canada, 867 Lakeshore Road, Burlington, ON, Canada L7R 4A6

[*] Present address: Organic Geochemistry Unit, School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

[#] Present address: EXPEC Advanced Research Center, Saudi Aramco, Dhahran 31311, Saudi Arabia

Contact: srowland@plymouth.ac.uk

For submission to: Environmental Science: Processes & Impacts

Table of contents entry:

The tricyclic naphthenic acid distributions of oil sands process-affected waters from two industry tailings ponds showed industry-dependent differences and, within a given industry pond, spatial, but little short-term temporal, variability.

Abstract

Oil industry produced waters, such as the oils sands process-affected waters (OSPW) of Alberta, Canada, represent a challenge in terms of risk assessment and reclamation due to their extreme complexity, particularly of the organic chemical constituents, including the naphthenic acids (NA). The identification of numerous NA in single samples has raised promise for the use of NA distributions for profiling OSPW. However, monitoring of the success of containment is still difficult, due to the lack of knowledge of the homogeneity (or otherwise) of OSPW composition within, and between, different industry containments. Here we used GC×GC-MS to compare the NA of five OSPW samples from each of two different industries. Short-term temporal and pond-scale spatial variations in the distributions of known adamantane acids and diacids and other unknown tricyclic acids were examined and a statistical appraisal of the replicate data made. The presence/absence of individual acids easily distinguished the OSPW NA of one industry from those of the other. The proportions of tricyclic acids with different carbon numbers also varied significantly between the OSPW of the two industries. The pond-scale spatial variation in NA in OSPW samples was higher than the short-term (2 weeks) temporal variations. An OSPW sample from an aged pond was exceptionally high in the proportion of $C_{15,16,17}$ compounds, possibly due to increased biotransformation. Such techniques could possibly also help to distinguish different sources of NA in the environment.

Introduction

Exploitation of many oil reserves requires the use of water for production and processing. For instance, increasing exploitation of the vast reserves of bitumen contained in oil sands deposits in northeastern Alberta, Canada, has led to the generation of large volumes of oil sands process-affected water (OSPW) which are not discharged back into the natural aquatic system due to the lack of knowledge about the effects this could have on the environment[1]. It has been estimated that approximately 840 million m^3 of tailings waters produced as a result of surface mining processes that contain a high loading of fine particles as well as dissolved compounds, are currently contained within

settling basins[2]. There have been concerns regarding the potential environmental impact of any leakage from tailings ponds and future projected extraction activities may further exacerbate any problems associated with the long-term storage of OSPW. This has led to calls for an improved understanding of the potential impacts upon the Athabasca River ecosystem and downstream communities[3–5]. Expert panel reviews concerned with the monitoring of waste materials from the oil sands industry were instigated by the Canadian Federal[6] and Alberta Provincial[7] governments and these have consistently recommended a complete overhaul of existing monitoring programs in order to strengthen the understanding of the potential impacts of oil extraction activities and to allow for future sustainable development. In response to this, a comprehensive monitoring system has been implemented[8]. An objective of the monitoring program is to evaluate the possible migration of contaminants associated with oil sands development into aquatic ecosystems via groundwater[8]. The proximity of some tailings ponds to the Athabasca River and its tributaries is a logical primary focus for these investigations, due to the acute and chronic toxicity of OSPW associated with aquatic organisms[9–16].

OSPW contains highly complex mixtures of organic compounds, many of which are so-called naphthenic acids (NA), which are thought to be intermediates and products of hydrocarbon biodegradation pathways [17–19]. NA are a very diverse group of acyclic, alicyclic and aromatic carboxylic acids. Due to their relatively high water solubilities, they may be more likely than more hydrophobic OSPW constituents, such as polycyclic aromatic hydrocarbons (PAH), to migrate via groundwater systems from tailings ponds and could therefore be useful from a monitoring perspective.

Previous attempts to profile OSPW and natural waters have indicated potential chemical markers for differentiation, but definitive assignments of sources have remained elusive. For example, Headley et al.[20] analysed the polar organic compound content of OSPW and natural surface waters by Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS). The relative abundances of sulfur-containing species and species containing O_n, NO_n, and N_2O_n within OSPW from two mines,

Athabasca River water and a reference lake, were subject to principal components analysis (PCA), which showed that sulfur-containing species were useful for distinguishing OSPW, while nitrogen-containing species showed potential for distinguishing natural from industrial sources [20]. A pilot study by Savard et al.[21] illustrated the potential for high-resolution mass spectrometry (HRMS) of ^{13}C isotopic signatures of carboxyl functional groups of NA to differentiate between older, bitumen-derived NAs and the younger, natural organic acids. Ross et al.[22] used HRMS to differentiate polar organic compounds in lakes, the Athabasca River and some of its tributaries and pore water from Athabasca River sediment. Although the observed similarities in compositions of OSPW and river surface waters reported were suggestive of OSPW seepage, distinction of anthropogenic from natural source inputs could not be made and the authors recommended the development of more specific analytical techniques for better differentiation.

The use of known reference compounds which can be identified and then monitored by use of characteristic GC retention times and electron ionisation mass spectra has proved to be the mainstay of environmental chemical analysis for decades (e.g. use of the USEPA 16 PAHs for monitoring hydrocarbon contamination), but until recently this could not be applied to OSPW due to the unresolved nature of the constituents by GC, the unknown composition of individual components and associated lack of authentic reference compounds for comparison. However, analysis of the acid extracts of single OSPW samples and of authentic synthetic or purchased reference compounds, by GC×GC-MS, revealed numerous tricyclic and pentacyclic diamondoid acids[23–26]. This presented an opportunity to apply a proven approach to the challenges associated with the oil sands processing. Rowland et al.[27] therefore suggested that diamondoid NAs could prove useful for monitoring purposes, as such acids are unusual in natural environments. A number of diamondoid acids are now commercially available, are easily resolved by GC×GC and have distinctive mass spectra, enabling the distinguishing of OSPWs from two industries storage ponds[27].

However, whilst the identification of numerous NA in single OSPW samples raises promise for the use of NA distributions for profiling, monitoring containment leakage is still difficult, due to the lack

of knowledge of the homogeneity (or otherwise) of OSPW composition within, and between, different industry containments. Therefore, there remain limitations on what can be concluded from examination of the diamondoid acids of only one or two industry samples. Here, we used GC×GC-MS to compare ten OSPW samples (five from each of two different industries). Short-term temporal and pond-scale spatial variations in the distributions of known adamantane acids and diacids and unknown tricyclic acids, were examined. The NA of a single sample of OSPW collected from a test pond in which it had been stored undisturbed for over 2 decades, was also examined.

Experimental

Sample Preparation

NA were extracted, as described below, from OSPW from two industries, A and B, in 2011. From Industry A, five water samples were collected from the same pond at the same location (a containment receiving fresh OSPW at the time of collection) over a 14-day period (November (7, 10, 14, 17, 21, = D0, 3, 7, 11, 14) 2011; Fig. 1). From Industry B, water from four different locations within a pond was sampled, plus one sample from a recycle pond which was attached to the main pond (September (22) 2011; Fig. 1). All the latter samples were collected within 24 hours of each other. Additionally, a sample of aged OSPW (>20 yr) was collected (October 2012) from an Industry A test pond created in 1993, originally filled with 6000 m^3 of surface water from an active tailings pond, with no subsequent addition, other than precipitation. Samples (100 ml) were all collected by the same method and at the same depth and were filtered through 0.2 μm filter cartridge to remove suspended solids, acidified to pH 2 and cleaned using 200 mg ENV+ SPE cartridges (Biotage, Charlotte, NC, USA). Samples were eluted with 10 mL of acetonitrile, evaporated under N_2 and then made up in 1.5 mL of acetonitrile. An aliquot of 0.5 ml was used for the gas chromatographic analysis. Of this, the acetonitrile was removed under N_2 and esterified by heating with BF_3-MeOH complex (70ºC, >30 minutes), back-extracted into hexane, dried and weighed. The extracts, as methyl esters, were analysed by GC×GC-MS. An aliquot of methylated Industry A sample from November 7 (Day 0, D0) was also dried over 3 h at 70°C under

a flow of N_2 and subsequently dissolved in 50 µL DCM and analysed in order to test the effects of excessive evaporation on the acid distribution. A method blank was also obtained.

Reference compounds were methylated (as above) for retention time and mass spectral comparison. These compounds included monoacids [C_{11}: Adamantane-1-carboxylic acid (Ia) , adamantane-2-carboxylic acid (Ib); C_{12}: 3-methyladamantane-1-carboxylic acid (II), 2-(1-adamantyl)acetic acid (III); C_{13}: 3,5-dimethyladamantane-1-carboxylic acid (IV), 2-(3-methyl-1-adamantyl)acetic acid (V), 3-(1-Adamantyl)propanoic acid (VI), 3-ethyladamantane-1- carboxylic acid (VII); C_{14}: 3,5,7-trimethyladamantane-1-carboxylic acid (VIII), 2-(3,7-dimethyl-1-adamantyl)acetic acid (IX)] and diacids [C_{12}: Adamantane-1,3-dicarboxylic acid (X), C_{13}: 3-(carboxymethyl)adamantane-1-carboxylic acid (XI)]. Spectra for these compounds are published elsewhere[23,28]. 3-Noradamantane carboxylic acid methyl ester, which was not present in any of the OSPW, was added to all samples as a retention time standard. All acids were purchased from Sigma-Aldrich Company Ltd., Gillingham, UK, except for V and VII which were purchased from Maybridge Chemical Company, Tintagel, UK.

GC×GC/MS analyses

Methyl esters of the OSPW extracts were analysed by GC×GC/MS using an Agilent 7890A gas chromatograph (Wilmington, DE, USA) equipped with a Zoex ZX2 GC×GC cryogenic modulator (Houston, TX, USA) interfaced with an Almsco BenchToFdx™ time of flight mass spectrometer (Almsco International, Llantrisant, UK). Scan speed was 50 Hz. The 1° column was a HP5-MS 30m x 0.25mm x 0.2µm (Agilent) coupled to a 2° column BPX-50 3m x 0.1mm x 0.1µm (SGE). The conditions were: 1° column 80°C (1 min), ramp at 2°C min^{-1} to 340°C, 2° column offset 10°C, hotjet offset 60°C. Helium was used as a carrier gas was with a flow of 2 ml min^{-1}.

Data analyses

Data from GC×GC-MS were processed using ProtoTOF software to .cdf files and analysed using GC-Image (Zoex). Samples of the reference compounds (methylated adamantane acids I-XI) were used to compare retention times and mass spectra to identify individual adamantane acids and adamantane

dicarboxylic acids present in the OSPW extracts (Fig. 2). Deuterated noradamantane was used as a chromatography standard for an exact comparison of retention times. A minimum of three injections per sample were performed to test instrument variability. Extraction of the molecular ions of *m/z* 194, 208, 222, 236, 250, 264 and 278 was performed on three runs of each of the five samples from Industry A (total n = 15) and 3 runs of the SE location sample of Industry B, and 4 of SW, NE, NW and Rec (total n = 19), in order to integrate peaks due to methyl esters of all isomers of the tricyclic acids with 11-17 carbons (Fig. 2). The fractional abundance fC_n was calculated using the intensity *Int* of the extracted ion current (EIC) according to Equation. 1.

$$fC_n = \frac{Int_{Cn}}{\sum_{n=11}^{17} Int_{Cn}} \qquad \text{(Eq. 1)}$$

Using the presence / absence of individual compounds, a binary cluster analysis was conducted using Ward's method and squared Euclidean distance (IBM ® SPSS ® Statistics). The results were represented in a dendrogram showing the maximum difference between the two main clusters at 25. On the f_{Cn} of the $C_{11\text{-}17}$ acids, a principal component analysis was conducted using R (FactoMineR package[29]). A Pearson correlation analysis was conducted to test whether the variation in f_{Cn} was due to a variation in TIC and thus concentration of the sample injected, and Welch's t-test in order to detect whether the differences in fC_n between the two ponds were significant (95% confidence level, df = 18).

Results

Identified compounds

We identified adamantane acids in all OSPW samples by comparison of spectra and GC×GC retention times with those of reference compounds [23,28] (Fig. 2b, Table 1). None of the monoacids were detected in all samples. Samples from Industry A contained a range of monoacids, while in the NW, SE, SW and Rec samples from Industry B, only VII could be detected. The sample from the NE

location (Fig. 1) showed a different profile, where a range of monoacids could be detected (Ia, Ib, II, III, V, VII). Of the diacids, X was present in all samples, whereas XI was present only in samples from Industry A. As the peaks of the diacids were well separated chromatographically, we could also compare proposed isomers Xa-d and XIa-f using mass spectra and retention times[28]. Also Xa was present in all samples, and Xb, Xc, Xd were present in all samples from Industry B and in most samples from Industry A. Some isomers of XI were present in some samples of Industry A, but, notably, XIa-f were detected in all samples from Industry B even though XI was not detected. The analysis revealed two clusters to be present, both consisting of five samples, pertaining to Industry A and Industry B (Fig. 3). The NE samples, though belonging to the cluster of Industry B samples, were nonetheless distinct from the other samples in cluster B.

Fractional Abundances (f C_n)

The extracted ion currents for the molecular ions for the C_{11}-C_{17} tricyclic monoacids were used to calculate the fC_n according to Eq. 1 (Table 2). The highest ratios observed were for fC_{14} (0.2809 for A, 0.2691 for B), while the lowest ratios observed were for fC_{11} (0.01845 for A, 0.03281 for B) and fC_{17} (0.06306 for A, 0.04333 for B). The means of the fC_n for all monoacids were significantly different between the two ponds on at least a 95 % confidence level in Welch's t-test ($p < 0.0001$, Table 2). The pond-scale spatially-separated samples from Industry B showed a greater range than the short-term (2 week) temporally-separated samples from Industry A (Fig. 4). While fC_{11}, fC_{12} and fC_{13} were higher for Industry B, fC_{14}, fC_{15}, fC_{16} and fC_{17} were higher for Industry A (Fig. 4). In a PCA conducted on the 7 fC_n, it was revealed that two components explained > 88 % of the variance. In fact, the two ponds could be clearly distinguished on only PC1 (77% of total variance, Fig. 5a), with the NE sample plotting lower on PC1 than the other Industry B samples. This variation of PC1 was, as expected from Fig. 4, due to the difference in C_n=11-13 vs. C_n=15-17 ratios, and is illustrated by the loadings of $fC_{11,12,13}$ and $fC_{15,16,17}$ plotting on opposite ends on PC1 (Fig. 5b). Based on this, the sum of $fC_{15,16,17}$ and $fC_{11,12,13}$ was calculated (Fig. 6). $fC_{11,12,13}$ ranged from 0.26 – 0.27 for Industry A and from 0.30 – 0.39 for Industry B, $fC_{15,16,17}$ from 0.44 – 0.47 for A and from 0.33 – 0.44

198 for B. The sample evaporated at high temperatures (Industry A-D0) showed a strongly changed
199 distribution in comparison to the original sample, with $fC_{11,12,13}$ decreased to 0.13 compared to 0.26
200 and $fC_{15,16,17}$ increased to 0.59 from 0.46 (Fig. 6).

201

202 **Discussion**

203 Our results allowed the evaluation of the temporal and spatial variability within a given pond, as well
204 as the comparison of NA distributions between two different industries, Industry A and Industry B.
205 Comparisons were conducted on simple presence/absence of known diamondoid acids, as well as on
206 distributions of their manifold isomers, supported by statistical analyses.

207 The simple presence / absence of the known adamantane acids and diacids in the OSPW samples
208 (Table 1) suggested differences between the samples from Industry A and those from Industry B.
209 Indeed, a cluster analysis based on the occurrence of these acids showed separation of the samples
210 according to the corresponding industry pond source (Fig. 3). These results strongly suggest that the
211 presence/absence of known NA can help to distinguish OSPW from different industrial sources. The
212 present study appears to be the first to achieve this differentiation and to establish target compounds
213 that could be used to characterize sources of OSPW.

214 However, as the simple presence of some of the known adamantane acids could be due to detection
215 limits of the GCxGC-MS method and a bias could arise from the high number of isomers with very
216 similar mass spectra and retention times, a second approach to characterisation of the differences
217 between OSPW samples was also attempted, using the distributions of both known and less rigorously
218 identified, but still tricyclic, acids. In addition to the known adamantane acids, there are many
219 different isomers of unknown tricyclic acids in OSPW, all producing the same molecular ion. The
220 number of isomers increases with increasing molecular weight, due to a higher number of possibilities
221 of permutation. As similar compounds of the same carbon number on a GC×GC elute in a 'tiled'
222 fashion (Fig. 2a), we used this tiling effect and integrated the extracted ion current (EIC) response of

223 the M^+ of monoacids of the corresponding tiles in order to avoid interference from fragments of
224 compounds with higher carbon numbers. We thus calculated the fC_n as specified in Equ. 1.

225 The fC_n of the OSPW of the two industries increased from n=11 to n=14 and decreased from n=14 to
226 n=17 (Fig. 4). This was not unexpected, as the number of isomers increases with n, but at higher
227 molecular weights the solubility in water likely decreases. Interestingly, differences in sampling
228 location (Industry B) seem to cause more variation than sampling at the same location on different
229 days over a two-week period (Industry A). This indicates that individual heterogeneities in OSPW
230 composition within a tailings pond could have an impact on the OSPW composition when samples are
231 taken from different sites. When investigating adamantane acids over the short sampling period, little
232 variation was detected in the tricyclics. However, this could change for other constituents, or with
233 metereological events or changes in production processes. Strikingly, a high spatial variation was
234 detected, which could in part be caused by differences in location such as shaded locations (less UV
235 degradation), distance from the OSPW inlets, dilution by runoff waters or streams or adsorption to
236 suspended particles. This suggests that, for further studies, the spatial heterogeneity of the ponds, and
237 thus the careful selection of locations for repeated sampling, needs to be taken into account.

238 However, even though the intra-variability of OSPW from the Industry B pond was large, a
239 significant difference was also noticed between the acids in the two ponds: the fC_n of n=11-13 acids
240 was lower in OSPW of Industry A than in those of Industry B. This situation was reversed for n=15-
241 17 acids (Table 2, 95% confidence, $P<0.0001$). In other words, samples from Industry A contained
242 relatively more tricyclic acids with higher molecular weights. The sample from the NE location of
243 Industry B was most different from those of Industry A. In order to confirm these differences, a
244 principal component analysis on the fC_n was conducted (Fig. 5). The scores plot (Fig. 5a) showed that
245 the differences were observed on PC1, and the loadings plot (Fig. 5b) that $fC_{11,12,13}$ and $fC_{15,16,17}$
246 plotted on PC1, whereas the TIC and fC_{14} plotted high on PC2. This also showed that the TIC (i.e.
247 reflecting the concentration injected) was not responsible for these differences, so long as it was
248 within the linearity range of the instrument. In order to further test that, linear and Pearson correlation

coefficients were calculated, showing that correlation between TIC and the fC_n was low (Table 3, 0.14 – 0.40 and 0.40 – 0.65); hence the TIC response was thus most probably not causing these differences.

There are several possible reasons for the differences in OSPW composition of industries A and B. Firstly, it could be that the ores used by industry A and B have different origins. Secondly, processing of oil sands ore by Industry A may result in dissolution of the higher molecular weight tricyclic acids than does the processing of ore by Industry B. This may also reflect differences in the NA composition of the ores. Thirdly, it is possible that, with ageing of the OSPW, the fractional abundance of $C_{15,16,17}$ condensed tricyclic acids relative to the lower molecular weight acids, increases (i.e. a shift to higher molecular weight compounds occurs). The OSPW from the pond of Industry A may be more 'aged' than those of Industry B. It is unlikely that the lower molecular weight acids might evaporate more during storage in the ponds or after sampling, especially as the acids are present as sodium salts in OSPW. However, once esterified for analysis, prolonged high temperature evaporation might indeed influence the distributions, so care is needed in order to avoid this. Intentionally prolonged evaporation of an aliquot of esterified NA from an OSPW from Industry A (sample D0) confirmed this effect (Fig. 6). However, this was unlikely to have caused the differences in the other samples examined herein, as these were evaporated to just dryness with care and all samples were handled identically. Future studies might usefully employ controlled evaporation by Kuderna-Danish apparatus to obviate this possibility.

In order to investigate possible environmental causes for the differences in $fC_{11,12,13}$ and $fC_{15,16,17}$, we therefore examined an OSPW sample from a greatly aged pond (>20 y storage) and again determined the fractional abundances of tricyclic acids. This "aged" source was from a test pond that was filled with OSPW from an active tailings pond in 1993, with no further OSPW addition. The high fractional abundance of tricyclic acids with n=15,16,17 compared to n=11,12,13 indicated that the differences observed could indeed be due to effects associated with increased ageing of the OSPW, presumably resulting in further biotransformation of the NA (Fig. 5).

The results from this study suggest the introduction of fC_n of condensed tricyclic acids as a characterisation parameter for OSPW might be worthy of further study. This can be conducted by GC×GC-MS, a powerful technique which is becoming increasingly common in the field of petroleum geochemistry. Furthermore, a calibration of other techniques with known reference acids (e.g. adamantane acids) could also lead to useful results. Using these parameters could allow characterisation of OSPW and other oil process waters in more detail and may also lead to a better understanding of the natural biodegradation processes.

Conclusions

OSPW from ponds from two different industries could be distinguished from the presence/absence of known adamantane acids, as well as by comparing the fractional abundances of related tricyclic acids with carbon numbers from 11 to 17 (fC_n). Negligible short-term temporal variations were detected, while considerable spatial variations occurred within one given pond. The distributions were shifted towards relatively higher molecular weight compounds in OSPW from a pond in which OSPW had been stored for >20y without further addition, suggesting that this may be due to biotransformation of the NA. This suggests that the ratios of $fC_{15,16,17}$ vs. $fC_{11,12,13}$ can indicate to some extent the aging of oil industry produced waters and could potentially present a useful variable for distinguishing natural leaching of NA from bitumen-containing soils from NA due to leakage of active ponds containing less aged OSPW.

Acknowledgements

N. Piechaud-Boura provided assistance with R packages. Research was enabled by an Advanced Investigators Grant (OUTREACH) awarded to S.J.R. by the European Research Council (ERC).

References

1 Energy Resources Conservation Board, *ST98-2012. Alberta's Energy Reserves 2011 and Supply/Demand Outlook 2012-2021*, Calgary, Alberta, 2012.

2 Energy Resources Conservation Board, *News Release - ERCB approves Fort Hills and Syncrude Tailings Pond plans with conditions*, 2010.

3 D. Schindler, *Nature*, 2010, **468**, 499–501.

4 D. J. Tenenbaum, *Environ. Health Perspect.*, 2009, **117**, A150–A156.

5 K. P. Timoney and P. Lee, *Open Conserv. Biol. J.*, 2009, **3**, 65–81.

6 L. Dowdeswell, P. Dillon, S. Ghoshal, A. Miall, J. Rasmussen and J. P. Smol, *A foundation for the future. Building an Environmental Monitoring System for the Oil Sands*, Environment Canada, Ottawa, Ontario, 2010.

7 H. Kvisle, H. Tennant, J. Doucet, W. Kindierski, A. D. Miall, D. Pryce, J. Rasmussen, G. Taylor, R. Wallace, H. Wheater and D. Williams, *A world class environmental monitoring, evaluation and reporting system for Alberta: The report of the Alberta Environmental Monitoring Panel*, Alberta Environmental Monitoring Panel, Edmonton, Alberta, 2011.

8 F. J. Wrona and Environment Canada., *Lower Athabasca water quality monitoring program aphase I: Athabasca River mainstem and major tributaries*, Environment Canada, Gatineau, Québec, 2011.

9 J. Anderson, S. B. Wiseman, A. Moustafa, M. Gamal El-Din, K. Liber and J. P. Giesy, *Water Res.*, 2012, **46**, 1662–1672.

10 R. J. Kavanagh, R. A. Frank, K. D. Oakes, M. R. Servos, R. F. Young, P. M. Fedorak, M. D. MacKinnon, K. R. Solomon, D. G. Dixon and G. Van Der Kraak, *Aquat. Toxicol.*, 2011, **101**, 214–220.

11 M. D. MacKinnon and H. Boerger, *Water Pollut. Res. J. Can.*, 1986, **21**, 496–512.

12 V. Nero, A. Farwell, L. E. J. Lee, T. Van Meer, M. D. MacKinnon and D. G. Dixon, *Ecotoxicol. Environ. Saf.*, 2006, **65**, 252–264.

13 L. E. Peters, M. MacKinnon, T. Van Meer, M. R. van den Heuvel and D. G. Dixon, *Chemosphere*, 2007, **67**, 2177–2183.

14 H. C. Reinardy, A. G. Scarlett, T. B. Henry, C. E. West, L. M. Hewitt, R. A. Frank and S. J. Rowland, *Environ. Sci. Technol.*, 2013, 6614–6620.

15 A. G. Scarlett, H. C. Reinardy, T. B. Henry, C. E. West, R. A. Frank, L. M. Hewitt and S. J. Rowland, *Chemosphere*, 2013, **93**, 415–420.

16 M. A. Warith and R. N. Yong, *Environ. Technol.*, 1994, **15**, 381–387.

17 H. W. Beam and J. J. Perry, *J. Bacteriol.*, 1974, **118**, 394–399.

18 M. Mishina, S. Yanaga, A. Tanaka and S. Fukui, *Agric. Biol. Chem.*, 1973, **37**, 863–870.

19 J.-F. Rontani and P. Bonin, *Chemosphere*, 1992, **24**, 1441–1446.

20 J. V. Headley, M. P. Barrow, K. M. Peru, B. Fahlman, R. A. Frank, G. Bickerton, M. E. McMaster, J. Parrott and L. M. Hewitt, *Rapid Commun. Mass Spectrom.*, 2011, **25**, 1899–1909.

21 M. M. Savard, J. M. E. Ahad, P. Gammon, A. I. Calderhead, A. Rivera, R. Martel, M. Klebek, R. Lefebvre, J. V. Headley, B. Welsh, A. Smirnoff, H. Pakdel, N. Benoit, S. Liao, J. Jautzy, C. Gagnon, J. Vaive, I. Girard and K. Peru, *A local test study distinguishes natural from anthropogenic groundwater contaminants near an Athabasca Oil Sands mining operation*, Geological Survey of Canada, Natural Resources Canada, Ottawa, Ontario, 2012.

22 M. S. Ross, A. dos S. Pereira, J. Fennell, M. Davies, J. Johnson, L. Sliva and J. W. Martin, *Environ. Sci. Technol.*, 2012, **46**, 12796–12805.

23 S. J. Rowland, A. G. Scarlett, D. Jones, C. E. West and R. A. Frank, *Environ. Sci. Technol.*, 2011, **45**, 3154–3159.

24 S. J. Rowland, C. E. West, A. G. Scarlett, D. Jones and R. A. Frank, *Rapid Commun. Mass Spectrom.*, 2011, **25**, 1198–1204.

25 C. E. West, J. Pureveen, A. G. Scarlett, S. K. Lengger, M. J. Wilde, F. Korndorffer, E. W. Tegelaar and S. J. Rowland, *Rapid Commun. Mass Spectrom.*, 2014, **28**, 1023–1032.

26 D. T. Bowman, G. F. Slater, L. A. Warren and B. E. McCarry, *Rapid Commun. Mass Spectrom.*, 2014, **28**, 2075–2083.

27 S. J. Rowland, C. E. West, A. G. Scarlett, C. Ho and D. Jones, *Rapid Commun. Mass Spectrom.*, 2012, **26**, 572–576.

28 S. K. Lengger, A. G. Scarlett, C. E. West and S. J. Rowland, *Rapid Commun. Mass Spectrom.*, 2013, **27**, 2648–2654.

29 S. Le, J. Josse and F. Husson, *J. Stat. Softw.*, 2008, **25**, 1–18.

355 **Figure legends**

356 **Figure 1.** Sampling strategy for this study. Samples from Industry A pond, were taken from the same
357 location over a period of two weeks (D0, 3, 7, 10 and 14), samples from Industry B pond were taken
358 on the same day but at different locations (NE, NW, SE, SW corners, and a recycle pond).

359 **Figure 2.** Structures and retention positions of the tricyclic NA (a) Extracted ion chromatogram of a
360 sample from Industry A, D14, (ions chosen to illustrate the identified compounds: *m/z* 149, 194, 222,
361 236, 252, 266) showing the retention position of the compounds I – XI and the tiling of the C_{11} – C_{14}
362 tricyclic acids. * Compounds were identified in some samples, but could not be unambiguously
363 verified in all samples due to high amounts of co-elution / low signal, and were thus excluded from
364 the further analyses presented in this manuscript. # Compound was present in some samples of this
365 study, but could not be detected in this sample. (b) Structures of the molecules identified with
366 reference compounds.

367 **Figure 3.** Binary cluster analysis on presence/absence of diagnostic compounds. Analysis of the
368 pattern of present/absent compounds showed that all samples from pond A and all samples from B
369 were clustering together. The Y-axis represents distance, with 25 being the maximum distance
370 between the two clusters.

371 **Figure 4.** Fractional abundances of known and unknown tricyclic acids. Boxplots of the fractional
372 abundance of C_{11}-C_{17} monoacids compared to all monoacids $f\,C_n$, calculated using Eq. 1, showing the
373 median (solid line), interquartile ranges (boxes) and extreme values (whiskers). Extreme values below
374 and above 1.5 IQS were plotted as outliers.

375 **Figure 5.** Results of the statistical analysis of the fractional abundance of the C_{11}-C_{17} monoacids. (a)
376 Scores plot of the samples from the ponds from Industry A and Industry B showing variation on PC2.
377 (b) Loadings plot for the different $f\,C_n$, showing that TIC and C_{14} were responsible for the variation on
378 PC1 (i.e. injection concentration), and that the differences in $C_{11\text{-}13}$ vs. $C_{15\text{-}17}$ were causing the
379 variation on PC2.

Figure 6. fC_n of OSPW samples. Fractional abundance fC_n of higher molecular weight tricyclic acids ($C_{15,16,17}$) vs. lower molecular weight tricyclic acids ($C_{11,12,13}$). "Aged" indicates the sample from a test pond which had not received "fresh" OSPW for 20 years, and "Industry A-D0 evaporated" the results for an aliquot left to evaporate for a prolonged time at 70°C.

Table 1. Presence (+) and absence (-) of diagnostic compounds (see Fig. 2) as determined by GC×GC-MS in samples from Industry A and B.

Industry	Sample	Compounds (see Fig. 2) + detected in sample, - not detected in sample																				
		Ia	Ib	II	III	IV	V	VI*	VII	VIII*	IX	X	Xa	Xb	Xc	Xd	XI	XIa	XIb	XIc	XId	XIe	XIf
A	D0	-	-	-	-	-	+	-	+	-	+	+	+	-	-	-	+	-	-	-	-	-	-
	D3	-	-	+	+	+	+	-	-	-	-	+	+	+	+	-	+	-	+	+	+	-	+
	D7	-	-	+	+	+	-	-	+	-	-	+	+	+	+	-	+	-	-	-	-	-	+
	D10	+	+	+	+	+	+	-	-	-	-	+	+	+	+	-	+	-	+	-	+	+	+
	D14	+	+	+	+	+	+	-	+	-	-	+	+	+	+	+	+	-	+	-	+	-	+
B	NE	+	+	+	+	-	+	-	+	-	-	+	+	+	+	+	-	+	+	+	+	+	+
	NW	-	-	-	-	-	-	-	+	-	-	+	+	+	+	+	-	+	+	+	+	+	+
	Rec	-	-	-	-	-	-	-	+	-	-	+	+	+	+	+	-	+	+	+	+	+	+
	SE	-	-	-	-	-	-	-	+	-	-	+	+	+	+	+	-	+	+	+	+	+	+
	SW	-	-	-	-	-	-	-	+	-	-	+	+	+	+	+	-	+	+	+	+	+	+

* Compounds VI and VIII might have been present, but could not be unambiguously identified due to co-elution of similar isomers. Compounds V and VII were thus not used in the statistical analysis.

Table 2. fC_n for both industries. Mean and standard deviations are shown.

	Industry A	**Industry B**	t_{exp}	**p-value**
fC_{11}	0.0184 ± 0.0012	0.0328 ± 0.0103	6.03	<0.00001
fC_{12}	0.0724 ± 0.0022	0.1117 ± 0.0142	11.9	<0.0000000001
fC_{13}	0.1760 ± 0.0046	0.2085 ± 0.0182	7.48	<0.0000001
fC_{14}	0.2809 ± 0.0046	0.2691 ± 0.0098	4.64	<0.0001
fC_{15}	0.2466 ± 0.0040	0.2183 ± 0.0212	5.67	<0.0001
fC_{16}	0.1425 ± 0.0051	0.1161 ± 0.0154	6.99	<0.000001
fC_{17}	0.0630 ± 0.0035	0.0433 ± 0.0068	10.90	<0.0000000001

Table 3. Linear correlation coefficients (R^2) and Pearson correlation coefficients (PCC) between fC_n and TIC, of all samples, and associated p-values.

	R^2	p-value	PCC	p-value
fC_{11}	0.36	<0.0001	0.62	<0.0001
fC_{12}	0.30	<0.001	0.57	<0.001
fC_{13}	0.30	<0.001	0.57	<0.001
fC_{14}	0.18	<0.01	0.45	<0.01
fC_{15}	0.40	<0.0001	0.65	<0.0001
fC_{16}	0.21	<0.01	0.48	<0.01
fC_{17}	0.14	<0. 1	0.40	<0. 1

Figure 1.

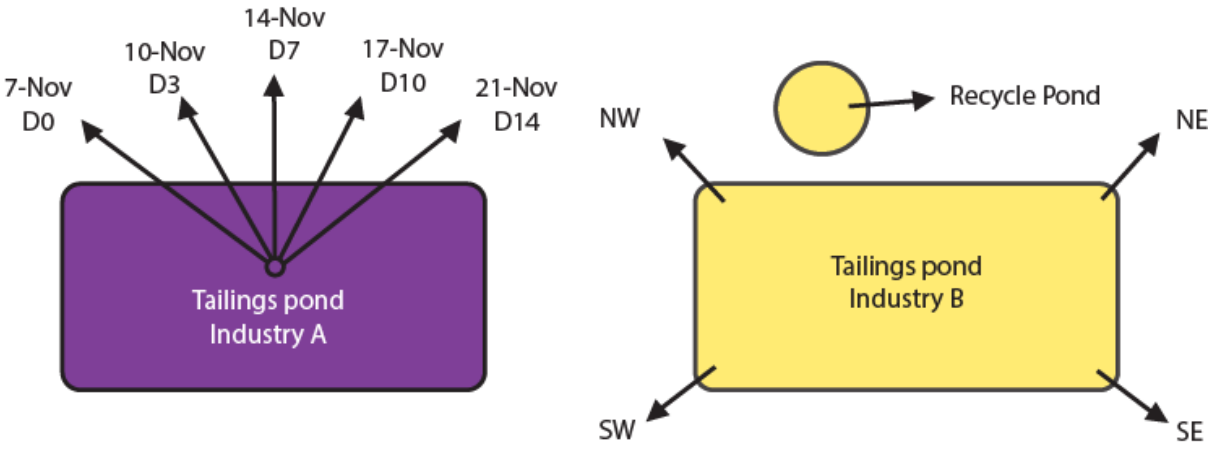

7-Nov
D0
10-Nov
D3
14-Nov
D7
17-Nov
D10
21-Nov
D14
Tailings pond
Industry A
Recycle Pond
NW
NE
Tailings pond
Industry B
SW
SE

Figure 2.

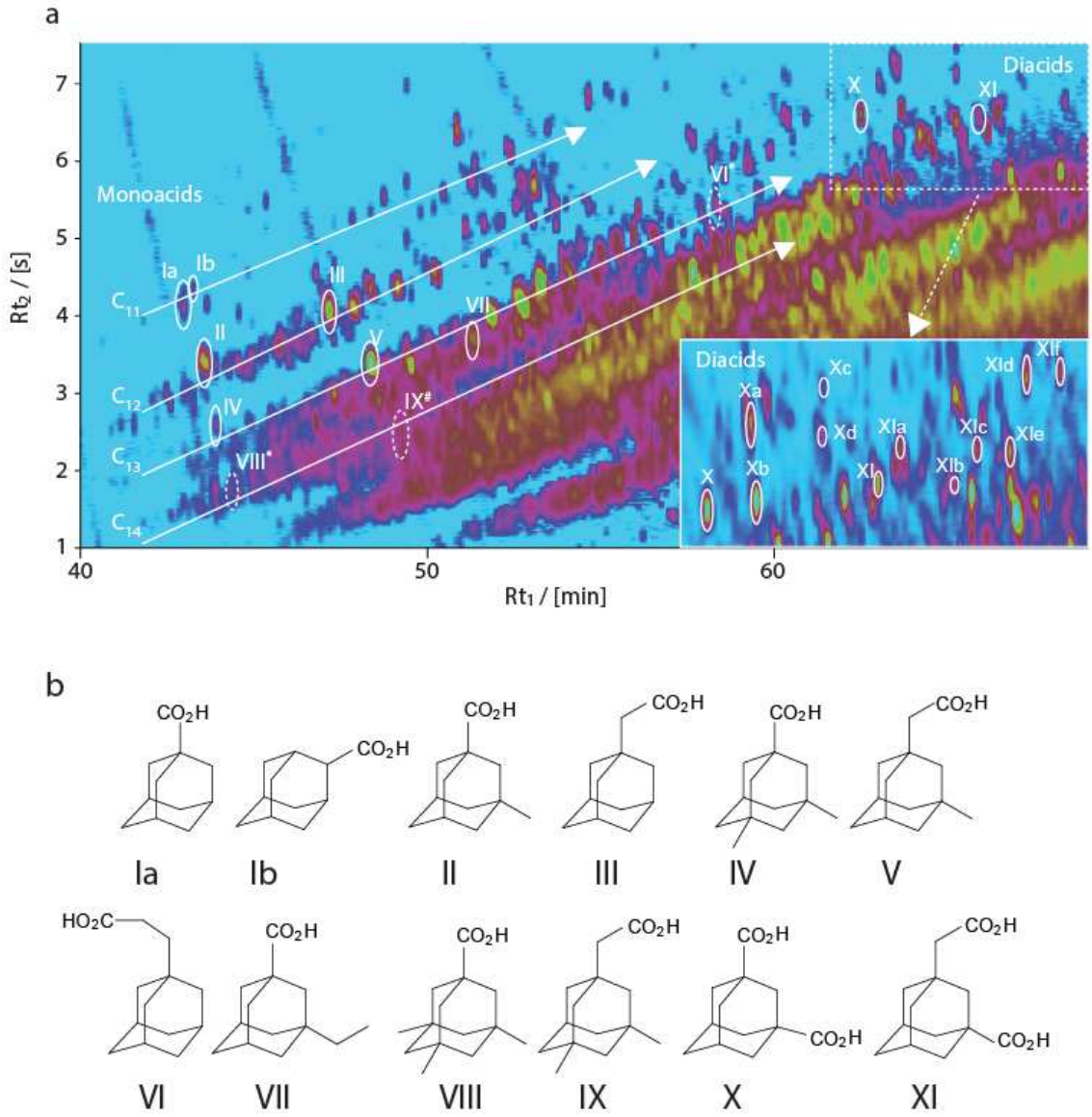

a
Monoacids
Diacids
Diacids
Ia
Ib
II
III
IV
V
VI*
VII
VIII*
IX#
X
XI
Xa
Xb
Xc
Xd
XIa
XIb
XIc
XId
XIe
XIf
C11
C12
C13
C14
Rt2 / [s]
Rt1 / [min]
b
Ia
Ib
II
III
IV
V
VI
VII
VIII
IX
X
XI

Figure 3.

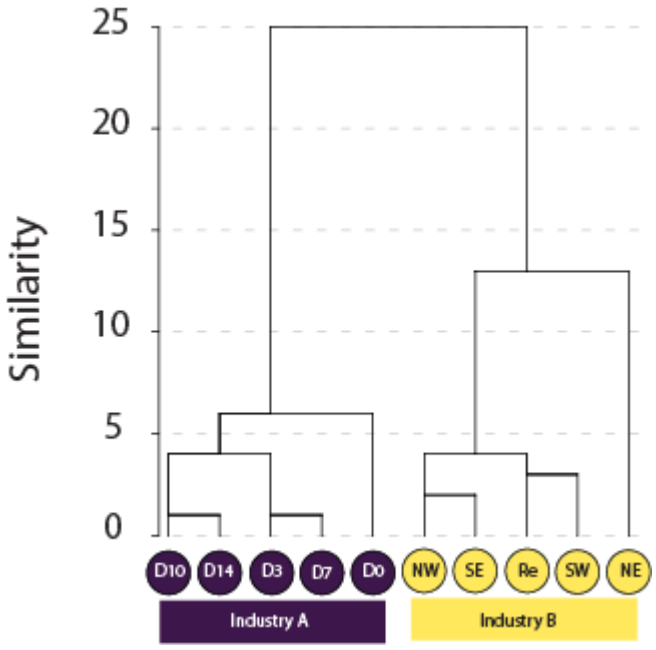

Similarity
25
20
15
10
5
0
D10
D14
D3
D7
D0
NW
SE
Re
SW
NE
Industry A
Industry B

Figure 4.

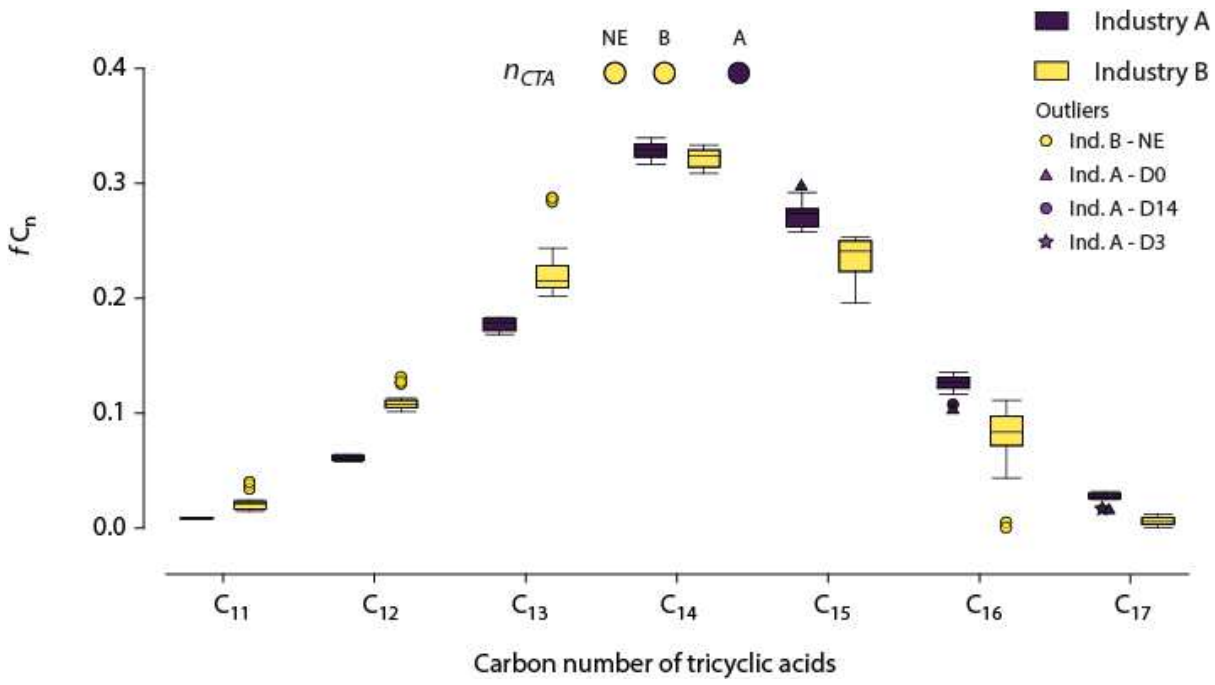

Industry A
Industry B
Outliers
Ind. B - NE
Ind. A - D0
Ind. A - D14
Ind. A - D3
NE
B
A
n_{CTA}
$f C_n$
0.4
0.3
0.2
0.1
0.0
C_{11}
C_{12}
C_{13}
C_{14}
C_{15}
C_{16}
C_{17}
Carbon number of tricyclic acids

Figure 5.

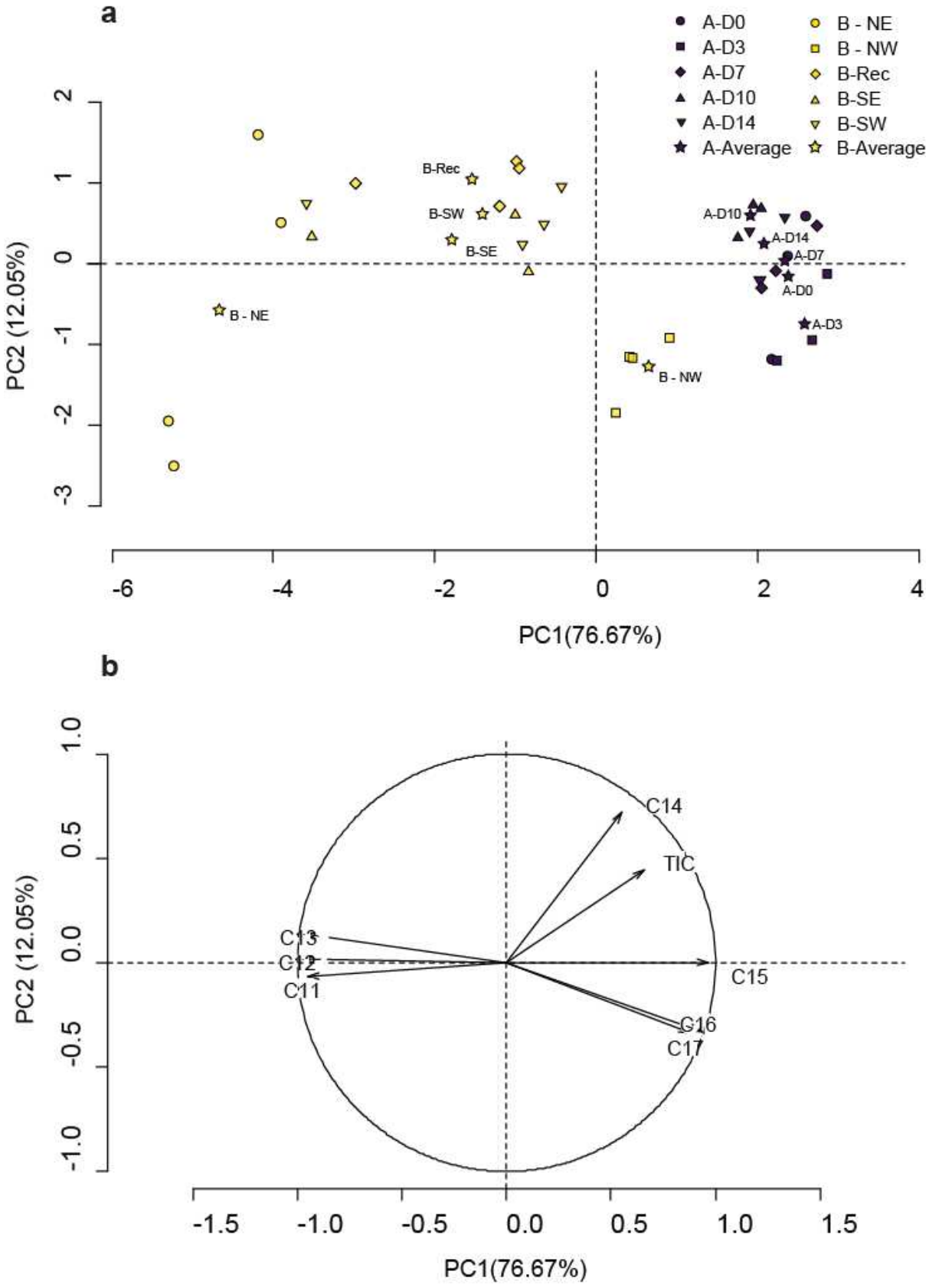
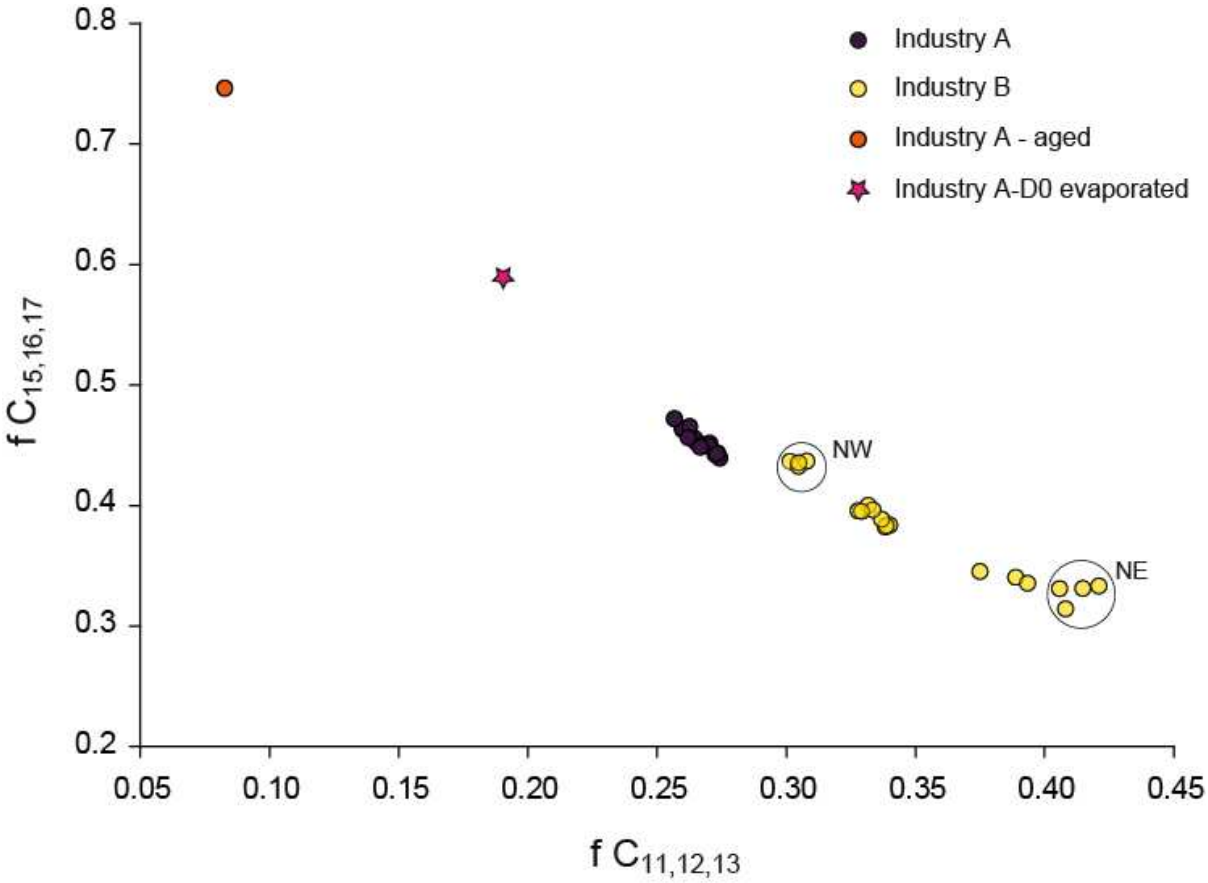

a
A-D0
A-D3
A-D7
A-D10
A-D14
A-Average
B - NE
B - NW
B-Rec
B-SE
B-SW
B-Average
PC2 (12.05%)
PC1(76.67%)
b
C14
TIC
C13
C12
C11
C15
C16
C17
PC2 (12.05%)
PC1(76.67%)

Figure 6.


Industry A
Industry B
Industry A - aged
Industry A-D0 evaporated
NW
NE
f C15,16,17
f C11,12,13
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45