2016-10-01

Assessing spatial and temporal variability of acid-extractable organics in oil sands process-affected waters

Richard A. Frank

Craig B. Milestone

Steve J. Rowland *Faculty of Science and Engineering*

John V. Headley

Richard J. Kavanagh

et al. *See next page for additional authors*

Title: Assessing spatial and temporal variability of acid-extractable organics in oil sands process-affected waters

Authors:

Richard A. Frank[a], Craig B. Milestone[a(b#)], Steve J. Rowland[c], John V. Headley[d], Richard J. Kavanagh[e], Sabine K. Lengger[c], Alan G. Scarlett[c], Charles E. West[c(f#)], Kerry M. Peru[d], and L. Mark Hewitt[a]

[a]Water Science and Technology Directorate, Environment Canada, 867 Lakeshore Road, Burlington, ON, Canada L7S 1A1

[b]Faculty of Applied Science and Technology, Sheridan College, 7899 McLaughlin Road Brampton, ON, Canada L6Y 5H9

[c]Petroleum and Environmental Geochemistry Group, Biogeochemistry Research Centre, University of Plymouth, Drake Circus, 5 Plymouth PL4 8AA, UK

[d]Water Science and Technology Directorate, Environment Canada, 11 Innovation Boulevard, Saskatoon, SK, Canada, S7N3H5

[e]Canadian Natural Resources Ltd., Calgary, AB

[f]EXPEC Advanced Research Center, Saudi Aramco, Dhahran 31311, Saudi Arabia

[#]**Present Address**

[*]**Corresponding Author:** Richard A. Frank

WSTD, Environment Canada

867 Lakeshore Road, Burlington, ON, Canada L7S 1A1

Phone: (905)336-4470 / Fax: (905)336-6430

Email: richard.frank@canada.ca

Total Word Count (Text + small figures/tables): 6305 + 600 (2 tables) + 1500 (5 figures) = 8305

Disclaimer: The views in this paper are only held by the authors and are not representative of the official policy of the authors' individual affiliations.

Abstract

The acid-extractable organic compounds (AEOs), including naphthenic acids (NAs), present within oil sands process-affected water (OSPW) receive great attention due to their known toxicity. While recent progress in advanced separation and analytical methodologies for AEOs has improved our understanding of the composition of these mixtures, little is known regarding any variability (i.e., spatial, temporal) inherent within, or between, tailings ponds. In this study, 5 samples were collected from the same location of one tailings pond over a 2-week period. In addition, 5 samples were collected simultaneously from different locations within a tailings pond from a different mine site, as well as its associated recycling pond. In both cases, the AEOs were analyzed using SFS, ESI-MS, HRMS, GC×GC-ToF/MS, and GC- & LC-QToF/MS (GC analyses following conversion to methyl esters). Principal component analysis of HRMS data was able to distinguish the ponds from each other, while data from GC×GC-ToF/MS, and LC- and GC-QToF/MS were used to differentiate samples from within the temporal and spatial sample sets, with the greater variability associated with the latter. Spatial differences could be attributed to pond dynamics, including differences in inputs of tailings and surface run-off. Application of novel chemometric data analyses of unknown compounds detected by LC- and GC-QToF/MS allowed further differentiation of samples both within and between data sets, providing an innovative approach for future fingerprinting studies.

Keywords

OSPW; acid extractable organics; chemical profiling; GC/LC-QToF/MS; GC×GC-ToF/MS; HRMS

1. Introduction

The oil sands region of northern Alberta, Canada contains an estimated 167.1 billion barrels of recoverable bitumen using current extraction technologies (Energy Resources Conservation Board, 2014). Surface mining uses a caustic hot water extraction for the recovery of bitumen from oil sand (FTFC, 1995), which results in a process-affected water consisting of an alkaline (pH 8-9) and saline aqueous suspension of sand, silt, clay, and residual bitumen (Zubot, 2010; Energy Resources Conservation Board, 2013). In situ mining practices are projected to recover approximately 80% of remaining bitumen reserves, however surface mining operations are still in development and in 2014 comprised 45.1% of total bitumen production (Energy Resources Conservation Board, 2015).

In an effort to maximize water recycling, oil sands process-affected water (OSPW) is stored in containments (FTFC, 1995) with a current estimated 840 million m^3 of OSPW in settling basins on industrial leases (Energy Resources Conservation Board, 2010). Numerous investigations into the toxicity of OSPW and tailings wastes have determined their toxicity towards aquatic (Nero et al., 2006; Kavanagh et al., 2011; Anderson et al., 2012; Gagné et al., 2012; He et al., 2012; van den Heuvel et al., 2012; Leclair et al., 2013; Wiseman et al., 2013) and terrestrial (Rogers et al., 2002; Gentes et al., 2006; Garcia-Garcia et al., 2011) organisms. Research continues to focus on identifying treatment methods to allow the reintroduction of OSPW into the surrounding environment, including ozonation (Anderson et al., 2012; He et al., 2012; Wiseman et al., 2013) and bioremediation processes (Del Rio et al., 2006; Han et al., 2008; Ahad and Pakdel, 2013; McKenzie et al., 2014), as well as work in support of the development of end pit lakes (Cumulative Environment Management Association (CEMA), 2012) and reclaimed wetlands (Oil Sands Wetlands Working Group - Alberta Environment, 2000).

In addition to the growing body of research investigating the reclamation of OSPW, recent attention has been directed to the potential for migration of OSPW from containment systems and entering the Athabasca River watershed. Previous studies had suggested the potential for OSPW migrations from tailings ponds into groundwater systems (Hunter, 2001; MacKinnon et al., 2005; Ferguson et al., 2009; Oiffer et al., 2009; Ross et al., 2012), however, these investigations were unable to differentiate the chemical profiles of OSPW from those of natural bitumen-influenced waters. Advances in high resolution mass spectrometry (HRMS) have indicated that differences in ion speciation may be used to differentiate sources of bitumen-influenced waters (Headley et al., 2009; Headley et al., 2011). Meanwhile, analyses of methyl ester derivatives of acids in OSPW by GC×GC-ToF/MS have indicated the structural diversity of the mixtures of acid extractable organics (AEOs; (Hao et al., 2005)), and with appropriate synthetic compounds have allowed the identification of tri- and pentacyclic diamondoid and monoaromatic acids (Rowland et al., 2011; Rowland et al., 2011; Rowland et al., 2011; Rowland et al., 2012), as well as over 30 bicyclic acids (Wilde and Rowland, 2015). A recent study (Frank et al., 2014) utilized both HRMS and GC×GC-MS methods to better distinguish the AEO distributions in groundwater influenced by OSPW from groundwater exposed to natural bitumen deposits, building upon earlier studies (Hunter, 2001; Ross et al., 2012).

One of the difficulties faced with the analysis of such complex mixtures of organic chemicals as OSPW AEOs is the vast array of structurally similar compounds present, most of which remain unknown. Recent advances in analytical techniques and bioinformatics data analysis have enabled untargeted profiling applications in various '-omics' fields to study differences in complex mass spectral datasets, where only those unknown compounds shown to indicate significant differences between samples are targeted for further study (Vaclavik et al.,

2011; Gao et al., 2012). Applications of this approach to complex environmental mixtures may offer the potential to conduct cause and effect evaluations or forensic source assignments, and while to date, applications in the environmental sector are relatively rare, their need is recognized (Bletsou et al., 2015).

While there has been an increased research focus in support of OSPW reclamation and monitoring efforts in recent years (Headley et al., 2013), little has been published regarding the homogeneity (or otherwise) of organic compound distributions in OSPW. Therefore, it is not known how representative analyses of individual grab samples from tailings ponds might be. Given the large size (i.e., 10 km^2) of some ponds (Han et al., 2009), spatial differences in OSPW chemistry are possible and likely will be dependent on factors including pond age, depth, proximity to tailings discharge points (solid and liquid), and orientation to predominating wind patterns. Similarly, fluctuations in the chemical compositions of OSPW entering individual ponds might also be expected, due to temporal variations in production and differences in bitumen compositions related to ore locations, for example. Variations in OSPW AEO compositions are of particular interest due to the inherent toxicity of some organic acids at current concentrations (i.e., 20-120 mg/L) reported in tailings ponds (Holowenko et al., 2002), and changes may occur with age which could affect the potential for mobility via groundwater transport.

In a companion study to this current investigation, distributions of specific, known, adamantane acids, including tricyclic acids and diacids, were investigated to assess short-term temporal and pond-scale spatial variability in AEO compositions (Lengger et al., 2015). In our current investigation, this companion study was greatly expanded upon by investigating a much wider range of known and unknown chemicals, utilizing a suite of methodologies to analyze the

same five OSPW samples collected over a 2-week period from the same location within a tailings pond at one mining operation (Industry A), as well as five samples at a second mine (Industry B) collected at one time from four different locations within a tailings pond and an associated recycle pond. Here, AEO mixtures were analyzed using HRMS, GC×GC-ToF/MS, and synchronous fluorescence spectroscopy (SFS), all of which have previously been shown to provide complementary data on AEO mixture compositions (Frank et al., 2014). This approach was extended further by bioinformatics analyses of profiles obtained by negative ion electrospray ionization liquid chromatography-quadrupole time-of-flight mass spectrometry (ESI-LC-QToF/MS) and positive ion electron impact gas chromatography-quadrupole time-of-flight mass spectrometry (GC-QToF/MS).

2. Methods

2.1 Sample Collection

Samples for this study were collected in 2011 from two mining operations north of Fort McMurray, AB, using pre-cleaned 1-L amber glass vessels provided by Environment Canada. Industry A provided five samples from the same location within a containment receiving fresh OSPW at the time of collection over a 2-week interval (Sample dates: November (07, 10, 14, 17, 21), 2011). Industry B provided four samples from four different locations within an OSPW containment (NE, NW, SE, SW), as well as one sample from an associated recycle pond (total = 5 samples; Sample date: September 22, 2011). Figure 1 depicts the sampling locations for Industries A and B and associated total AEO concentrations. Following sample collection, bottles were placed with ice into coolers and shipped by refrigerated transport to Burlington, ON where they were immediately placed into cold storage (4˚C) pending extraction.

2.2 Synchronous Fluorescence Spectroscopy (SFS)

OSPW samples were filtered using disk filters (PES, 25 mm GD/X, O.2 µm pore size; GE Healthcare UK Ltd., Buckinghamshire, UK). Analysis by SFS was performed using a Perkin-Elmer Luminescence spectrometer LS50B and data collection was controlled by FL Winlab 3 software (Perkin-Elmer, Norwalk, CT) as described previously (Kavanagh et al., 2009; Frank et al., 2014). Spectral maxima for SFS oil sands AEO profiles are at 272, 307 and 324 nm, using AEO mixtures extracted from active OSPW containments.

2.3 AEO extraction

Samples from the spatial and temporal studies were removed from cold storage, stirred vigorously for 10 minutes and subsampled (100 mL) for AEO extraction. Subsamples were allowed to attain room temperature, filtered through 0.2µm disk filter cartridges (PES, 25 mm GD/X, O.2 µm pore size; GE Healthcare UK Ltd., Buckinghamshire, UK), acidified to pH 2 with HCl and extracted using SPE (200 mg Isolute ENV+, Biotage Charlotte NC, USA), as described previously (Frank et al., 2014), using acetonitrile as eluent. Acetonitrile eluents were evaporated with N_2 just to dryness and then reconstituted in 1.5mL of acetonitrile. Each extract was then divided equally for analysis by GC×GC-ToF/MS and LC/GC-QToF/MS, as well as archiving. For GC×GC-ToF/MS, extracts were methylated using BF_3/methanol, as described previously (Frank et al., 2014). For LC-QToF/MS, extracts were evaporated and reconstituted in 100% MeOH. For GC-QToF/MS, extracts were solvent exchanged into DCM and methylated with freshly prepared excess diazomethane in diethyl ether for 24h. Methylated extracts were solvent exchanged to toluene for profiling. For HRMS speciation, analyses, and low resolution ESI-MS,

a separate extraction was performed on 10 mL initial OSPW volumes, as described above. Sample extracts were solvent exchanged after SPE from acetonitrile to 50/50 acetonitrile/water with 0.1% NH_4OH.

2.4 Infusion Negative Ion Electrospray Ionization Mass Spectrometry

Concentrations of NAs were determined using low resolution negative ion ESI-MS. In brief, samples (5.0 uL) were introduced into the eluent stream (200 µL min^{-1}, 50:50 CH_3CN:H_2O containing 0.1% NH_4OH) using a Waters 2695 advanced separation system (Milford, MA). Mass spectrometry analysis was conducted using a Quattro Ultima mass spectrometer (Micromass, UK). MS conditions were as follows: source temperature 90°C, desolvation temperature 220°C, cone voltage setting 62 V, capillary voltage setting 2.63 kV, cone gas N_2 147 L h^{-1}, desolvation gas N_2 474 L h^{-1}. The low and high mass resolutions settings were set at 14.0 (arbitrary units) and ion energy was 1.7 eV. Entrance voltage was 96 V, collision energy 13 eV and exit voltage 56 V. The multiplier was set at 410 V. Data was acquired using full scan MS (m/z 100–550). MassLynx V.4.1 software was utilized for all instrumental control and data acquisition and processing.

High resolution (120000 resolving power) negative ion ESI-MS (LTQ Orbitrap Velos, ThermoFisher Scientific) was used to characterize the components within the complex mixtures. In brief, the mass spectrometer was operated in full scan at an m/z scan range of 100 to 600 and the following conditions: sheath gas flow rate, 25 (arbitrary units); spray voltage, 2.90 kV; auxiliary gas flow rate, 5 (arbitrary units); S lens radiofrequency level, 67%; heater temperature, 50°C; and capillary temperature, 275°C. Infusion solvent used was 50:50 acetonitrile:water containing 0.1% ammonium hydroxide at a flow rate of 200 mL min^{-1}. The mass accuracy was

less than 2 ppm error for all mass assignments. Software used for molecular analysis was Xcalibur Ver 2.1 (ThermoFisher Scientific) and Composer Ver 1.0.2 (Sierra Analytics). Principal component analysis (PCA; Mass Frontier 6.0 (Thermo Fisher Scientific and HighChem Ltd)), was utilized to investigate patterns and clustering of data with respect to spatial, temporal, and industrial variances. PCA calculations are based on differences in abundance and presence or absence of compounds detected from high resolution full scan data (m/z 100-600) acquisitions.

2.5 GC×GC-ToF/MS

Comprehensive multidimensional gas chromatography time-of-flight mass spectrometry (GCxGC-ToF/MS) analyses were conducted using an Agilent 7890A gas chromatograph (Agilent Technologies, Wilmington, DE) fitted with a Zoex ZX2 GCxGC cryogenic modulator (Houston, TX, USA) interfaced with an Almsco BenchTOFdx™ time-of-flight mass spectrometer (mass resolution 1000 at mass 1000; Almsco International, Llantrisant, Wales, UK). The first-dimension column was a HP5-MS 30 m x 0.25 mm x 0.20 µm (Agilent), and the second-dimension column was a 50% phenyl polysilphenylene siloxane 3 m x 0.1 mm x 0.1 µm BPX50 (SGE, Melbourne, Australia). Helium was used as carrier gas and the flow was kept constant at 2.0 mL min^{-1}. Samples (1µL) were injected at 300°C splitless. The oven was programmed from 80°C (hold for 1 min), then heated to 340°C at 2°C min^{-1}. The secondary oven was offset by 10°C. The modulation period was 8s. The MS transfer line temperature was 290°C and ion source 280°C. Data processing was conducted using GC Image™ v2.3 (Zoex, Houston, TX, USA).

2.6 LC-QToF/MS

Analyses by LC-QToF/MS (Agilent QToF 6520, Agilent 1200 series LC system, Agilent Technologies, Santa Clara, CA, USA) were performed using a Poroshell 120 EC-C18 column (Agilent Technologies; 3.0 x 50 mm, particle size 2.7 µm), temperature controlled at 40°C. A binary solvent system of 100% water (Solvent A) and methanol (Solvent B), both containing 0.1% formic acid, were used in gradient programme with a constant flow of 0.4 mL min^{-1}. The initial mobile phase was 5% B for 2 minutes, followed by a linear gradient to 95% B over 20 minutes, held for 10 minutes, before returning to equilibrate at initial conditions (16 min). A sandwich injection of sample (2 µL) and isotopically-labelled internal standards (d_{19}-decanoic acid (mass 191.38); d_9-9-anthracene carboxylic acid (mass 231.30); C/D/N Isotopes, Pointe-Claire, QC, Canada) (1 µL), with a needle rinse for 10 seconds, was used for each injection. Each sample was injected in triplicate, as technical replicates, in a randomised order. A QC sample, pooled from each of the samples, was run every 5 injections, followed by a solvent blank to assess any carryover.

The QToF/MS was run using an ESI source in negative ion mode. Parameters for the mass spectrometer were as follows: VCAP 3000, nebuliser pressure 35 psig; drying gas 10 L min^{-1}; gas temperature 350°C; skimmer voltage 65V; fragmentor voltage 130V; all other conditions were set by the data system autotune. Acquisition of accurate mass spectra was obtained at 3 scan sec^{-1} in the range of m/z 100-1050. Continuous internal calibration was performed using internal reference ions for constant alignment (Agilent reference mixture, purine (m/z 119.0362), HP-0921 (m/z 966.0007).

2.7 GC-QToF-MS

Analyses by GC-QToF/MS (Agilent 7200 QToF/MS and 6890 GC) were performed using an electron impact source in positive ion mode. A HP-5ms column (Agilent; 30m x 250µm x 0.25 µm) was used, with a gas (N_2) flow of 1.2 mL min^{-1}, using a splitless injection at 50°C for 0.5 min, then 750°C min^{-1} to 320°C for 5 min, using a sandwich injection comprised of 1 µL sample and 1 µL internal standard (d_{19}-decanoic acid (methyl ester); C/D/N Isotopes). Other parameters were as follows: source 275°C; thermal aux 2 280°C; fixed emission current 35 µA; m/z scan range 50-500; 5 spectra sec^{-1}; continuous internal reference ions (m/z 68.9947, 365.9895). The oven-temperature program was as follows: initial equilibration time of 3 min, temperature ramped by 50°C min^{-1} for 1 min, then ramped at 15°C min^{-1} to 120°C, then 5°C min^{-1} to 300°C, held for 2 min, then 10°C min^{-1} to 320°C and held for a final time of 5 min.

2.8 GC/LC-QToF data analysis

Data were acquired using MassHunter Workstation Software -Data Acquisition (LC/MS Version B5.00; GC/MS Version B6.00, Agilent), with data and statistical (ANOVA) analyses performed using MassHunter – Qualitative Analysis (Qual, Version 6.00, Agilent) and Mass Profiler Professional (MPP, Version 12.6.1, Agilent).

Data files from LC-QToF/MS were processed using a recursive analysis workflow to generate a list of potential entities. Raw data files were first processed in Qual using the Molecular Feature Extractor (MFE) "Small Molecules" algorithm and the following parameters that were not default: peak height >600 counts, negative ion species of –H, peak spacing tolerance of 0.0025 m/z ± 7.0 ppm, charge state limit ≤ 3, a quality score of 80.00, and with all features requiring two or more ions (i.e. deprotonated molecular ion with at least one ^{13}C isotope). The resulting molecular feature, or entity, lists from each file were then imported into

MPP for mass and retention time alignment using the software algorithm and an RT window of 0.3 min. Following alignment, a "pooled" entity list was exported for recursion, where in Qual, the data files were again opened, and this time the "Find by Ion" algorithm was used to target the list of ions derived from the pooled entity list (which reduces MFE false negatives). The generated molecular feature list for each sample file was then reimported into MPP for further re-alignment (mass and retention time, using tolerances from above), normalisation (percentile, 0.75), and filtering (>5000 counts, ≥ 2 ions). Entities were further filtered based on flags (present in at least 2 injections) and frequency (present in all triplicate injections in at least one sample). Averaged samples were then blank-subtracted, and principal component analysis (PCA) plots generated for unsupervised pattern recognition.

GC-QToF/MS molecular feature lists were generated in Qual by the "Find by Chromatogram Deconvolution" algorithm with an RT window size factor 100, SNR threshold 2.00, extraction window 0.3 left and 0.7 right, sharpness threshold 25, absolute height 500 counts and a relative height of 2% of the largest peak. Molecular feature lists were exported directly into MPP, and processed as above for LC-QToF/MS data.

3. Results

3.1 SFS

Analysis by SFS of the collected filtered OSPW samples revealed the triple peak maxima at 272, 307 and 324 nm (Figure S1) reported previously for the AEO fraction of OSPW (Kavanagh et al., 2009; Rowland et al., 2011). The spatial variability sample set from Industry B contained two additional maxima (359 and 373 nm) not present in the Industry A samples. There were no noticeable differences between samples within each OSPW source.

3.2 Infusion –ESI (low res and HRMS)

Total AEO concentrations were determined by direct injection ESI-MS and are provided in Figure 1. Ion speciation plots (Figure 2) generated by HRMS data did not reveal any noticeable trends between samples originating from the same source. The temporal data set from Industry A was more consistent across the different ion species than the spatial data set for Industry B, however, a PCA plot of the HRMS data (Figure 3) revealed distinct separation of the two OSPW sources, as well as some degree of separation within the sample sets. Consistent with previous HRMS analyses of OSPW, O_2 species were more abundant than O_4 species (Barrow et al., 2015).

3.3 GC×GC-ToF/MS

In addition to the known tricyclic adamantane acids and diacids identified previously in these samples (Lengger et al., 2015), analysis by GC×GC-ToF/MS also revealed (Figure 4) the unknown, but tentatively assigned, "Family A" (*m/z* 145; 7 isomers) and "Family B" (*m/z* 237 & 310 doublet) aromatic acid methyl esters identified previously in several other OSPW samples (Rowland et al., 2011; Frank et al., 2014). In respect of the distributions of these unknown Family A/B compounds, the GC×GC-ToF/MS profiles of the temporal samples from Industry A were all very similar to one another, indicating little change over the course of the two week period from which the samples were collected. Three of the spatial samples from Industry B, which differed only in the location within the pond from which they were taken, were also similar to one another in this respect, but noticeable exceptions included the distributions exhibited by the NE sample extract, which had much lower peak volumes for all of the Family

A/B isomers relative to the other samples (Figure S2). Similarly, the NW sample also had lower peak volumes for most of the Family A/B isomers, however the differences were less pronounced than the NE sample.

3.4 LC & GC-QToF-MS

For LC & GC-QToF/MS analyses, data for each OSPW sample was processed to obtain individual unknown entities, defined here by the combination of molecular weight, retention time, and abundance, as described below. To be considered as an "entity" for the temporal and spatial OSPW comparisons, several filtering criteria were applied in the data analysis workflow, summarized in Table 1. Following filtering and blank subtraction, a total of 2732 and 174 entities were obtained for LC- and GC-QToF/MS, respectively. Due to the relative abundance of detected entities in the LC-QToF/MS data set, this was selected for more in depth analyses. In order to assess temporal, spatial, and industrial variation, unsupervised PCA was utilised on filtered entities (from Table 1) to investigate any patterns or clustering of the data. These PCA calculations include relative changes in abundance, as well as assessing the presence/absence of compounds in each sample. Firstly, these plots (averaged values from triplicate injections) show that samples taken from the two industries can be separated into two distinct clusters for both LC- and GC-QToF/MS data (Figure 5). This separation (Figure S3) is primarily along the X-axis for both LC (PC-1, 28.5%) and GC (PC-1, 51.3%). The temporal series of samples from Industry A are clustered, with the majority of the separation occurring along the Y-Axis. The spatial series of samples from Industry B, while remaining distinct from Industry A, are less clustered. While the samples obtained from the SE and SW corners, and the recycle pond, plot

together, the samples taken from the NW and NE corners are distinct and are very different from each other, indicating that some entities/features are chemically distinct in each of these samples.

Results in Table 2 present the number of total entities as well as those that are unique to each sample in space or time for each industry. Within the temporal sample set (Industry A) the total entities detected at each site ranged from the lowest 758 (14-Nov) to highest of 1084 (07-Nov), with unique entities per sample ranging from 52 (6.3%, 21-Nov) to 240 (23.3%, 10-Nov). For the spatial sample set (Industry B), both the NE (315 unique entities) and NW (305 unique entities) samples showed that close to 30% of the detected entities at those sites were unique to each site, while for the remaining three samples (SW, SE and Recycle pond), unique entities were much lower, ranging from 9.7-12.5%. The total entities for the NE and NW samples (1035 and 980, respectively) were also higher relative to the SW (752), SE (825) and Recycle pond (911) samples.

When considering simple presence/absence of entities for each industry, of the 2732 total entities generated by LC-ToF/MS (Table 1), 790 (28.9%) were unique to Industry A (temporal), 849 (31.1%) were unique to Industry B (spatial), while 1093 (40.0%) entities were present in at least one sample from both ponds (Figure S4). Further analysis of the LC-QToF/MS data revealed that of the 2732 entities detected in the temporal and spatial data sets, 1138 of these were considered statistically different (includes presence/absence and relative abundance changes with $p<0.05$) between the two industries (Figure S5; statistics based on the five samples from each industry).

4. Discussion

This investigation set out to broadly characterize the variabilities in profiles of unknown AEOs in samples collected from each of two tailings ponds, separated by time and space. While the dynamics impacting the transformation and fate of all of the mixture components of bitumen-influenced waters are of great interest and are deserving of investigation to support monitoring and remediation initiatives, the objective of this current study was to assess the possibility of chemical variability within OSPW containments as single samples are routinely collected and assumed as representative of the entire tailings pond. Furthermore, this investigation sought to identify a method, or complement of methods, capable of differentiating highly complex mixtures within OSPW. Oil sands process-affected water is a saline aqueous suspension of sand, silt, clay, and residual bitumen (Zubot, 2010; Energy Resources Conservation Board, 2013), containing complex mixtures of neutral organics (e.g. polycyclic aromatic hydrocarbons (PAHs) and dibenzothiophenes), more polar organics (e.g. carboxylic 'naphthenic' acids), and inorganics including metals and salts (Madill et al., 2001; Kavanagh et al., 2011; van den Heuvel et al., 2012; Brown and Ulrich, 2015). Tailings ponds are visually variable due to non-uniform distributions of residual bitumen at the surface, and would be expected to be inconsistent in chemical composition due to the differences in ores being processed, however analytical methodologies capable of adequately assessing potential chemical differences within tailings ponds have been lacking. This investigation used a suite of analytical techniques to investigate the chemical variability present within two tailings ponds, with separate assessments of temporal and spatial variability, expanding greatly upon a companion paper (Lengger et al., 2015) that assessed the variability of adamantine acids within the same sample set. The temporal assessment sample set from Industry A encompassed a modest range of 2 weeks, with the spatial assessment sample set from Industry B including samples from the four geographical corners of

the tailings pond, as well as a sample from a separate, associated, recycle pond. The samples in the spatial data set are of interest due to the fact that the Industry B tailings pond receives inputs from three different sources. There are two types of pipelines that deposit tailings to the tailings pond: floatation tailing lines and naphtha recovery unit (NRU) tailings lines. The floatation lines primarily discharge a mixture of water, sand, and trace amounts of bitumen, and also deposit sand onto the beaches along the perimeter of the pond. Conversely, the NRU tailing lines discharge a mixture containing greater amounts of bitumen, solvent, and water, and relatively less sand, depositing on the north and east sides of the tailings pond. Finally, tributaries of the Tar River flow into the west side of the tailings pond, providing a third input source. While the majority of the water that would typically flow in these tributaries has been diverted to a fisheries compensation lake, large amounts of surface water runoff likely enter the northwest and west side of the tailings pond.

Analysis by HRMS was able to distinguish differences between the two different industries, consistent with similar source apportionment investigations (Headley et al., 2013 and references herein). While there were no evident trends within either the temporal or spatial sample sets, nor were there any trends in total NAs, a PCA plot of the HRMS data (Figure 3) demonstrated a clear separation of the two OSPW sources, with more distinct separation within the spatial samples. This demonstrates the applicability of HRMS to differentiate tailings ponds, expanding upon previous work using O_2:O_4 speciation ratios (Barrow et al., 2015). SFS analysis was not capable of identifying differences within either sample set, but a slight difference between the two industries was observed (Figure S1). The additional maxima detected by SFS in the Industry B samples (359 and 373 nm) have not been reported previously (Kavanagh et al., 2009; Rowland et al., 2011; Frank et al., 2014; Moustafa et al., 2014), suggesting that the

aromatic acids associated with these maxima could be among the 849 entities unique to Industry B (Figure S4) and potentially useful for source identification.

GC×GC-ToF/MS analyses of the Industry A and B OSPW samples using previously reported mono-aromatic acid "Families" (Rowland et al., 2011; Frank et al., 2014) revealed distinct differences between individual samples. While all Family A and B isomers were present in all of the investigated samples, consistent with previous GC×GC-ToF/MS analyses of other OSPW sources (Frank et al., 2014), there were distinct differences in relative abundances. Family A and B profiles from sites within the Industry B pond were similar, with the most notable exception of the NE sample. The NE sample from Industry B is considerably lower in terms of both base peak intensity and volumes for the investigated Family A and B isomers (Figure 4), and given the knowledge that fresh tailings release was occurring in this area at the time of sampling, suggests that the increased relative abundance of these mono-aromatic acids may be degradation products from other unknown acids. Also worthy of note, although the separate Recycle Pond sample from Industry B appeared to be similar in Family A and B profiles to the NW, SW, and SE tailings pond samples from the same industry, it was analyzed at about 50% of the concentration of other Industry B samples due to a larger proportion of the Recycle Pond sample not successfully methylating to produce methyl esters. Therefore, after taking into account differences in the concentrations of samples analyzed (Industry A was ~20% > than Industry B), with the exception of the NE and Recycle Pond samples from Industry B, there were few differences between Industries A and B detected when investigating the Family A and B mono-aromatic acids detectable by GC×GC-ToF/MS. However, although the presence/absence of the Family A and B isomers did not appear to differ in the temporal samples (Figure 4), determination by GC×GC-ToF/MS of known tricyclic diamondoid acids and diacids as methyl

esters, did reveal some differences between the Industry A temporal samples and also showed statistically significant differences between the NE and other spatial samples within the same Industry B pond samples (Lengger et al., 2015).

While some compounds in the complex mixtures that comprise AEOs in OSPW bitumen influenced waters have been confirmed (Bowman et al., 2014; Wilde et al., 2015), the vast majority remain unknown. By using the untargeted approaches applied in this study, it was possible to qualitatively differentiate these unknowns, and to also evaluate differences in relative abundances by utilizing chemometric data analysis software. The application of this chemometric platform, originally designed for proteomic and metabolomics studies, to the present assessment of tailings pond variability is novel, as we are unaware of any previous application in not only this field, but environmental applications in general. Furthermore, this approach is likely best-suited to further elucidate the complexity of the soluble organic mixtures within bitumen-influenced waters, a result that is vital to monitoring and remediation initiatives in the oil sands region.

In this study, both GC- and LC-QToF/MS techniques were employed, with the LC-method coupled to a recursive data workflow. In brief, if an entity was observed in any sample, that entity was subsequently searched for again in all other samples. Relative to data generated by LC-QToF/MS, GC-QToF/MS profiles produced a lower number of entities, which we believe is due to the complexity of the mixtures (i.e. inability to chromatographically resolve components) and the extensive fragmentation inherent in EI ionization, meaning that subsequent deconvolution of individual entities becomes difficult. Future work will evaluate a softer (i.e. chemical) ionization technique, and together with recent software updates that enable a recursive workflow, will help simplify the GC deconvolution process and potentially retrieve more

entities. Despite this, the GC-QToF/MS data still allowed differentiation between the samples from the two industries, although did not provide any further insight into the temporal and spatial studies. It should be noted that this is a truly untargeted approach that considers all of the generated information, not just data for target compounds, which could partially explain why more scatter was observed for the data in the PCA plots.

When considering the profiles generated by negative ESI-LC-QToF/MS, PCA analysis also showed that the samples from the two industries were differentiated. The data for samples from Industry A (samples taken over a two week period) and compared using PCA analysis, show that while there were differences between samples, no real trend was evident. The results obtained for the samples from Industry B, which differed only in their position in the pond from which they were taken, indicated that the samples from the SE and SW corners, and from the recycle pond, were more closely related in chemical composition, relative to the samples from the NE and NW corners, that were different both from each other, and from the other three samples. Since there are no known discharges from either floatation tailings lines or NRU tailings lines in the SE and SW corners, less "fresh" tailings can be expected in these areas and consequently, the materials being deposited into the tailings pond are perhaps likely to be more thoroughly mixed (evenly distributed) in these areas. Water from the SW corner of the tailings pond is conveyed to the recycle water pond for re-use in oil sands extraction process, which might also help to explain the similarity between these three samples.

The PCA analysis of the data generated by HRMS (Figure 3) showed that the samples from the two industries were differentiated in a manner similar to that of the LC-QToF/MS (Figure 5). For example, no pattern was evident for samples from Industry A (temporal), however, the results obtained for the spatially derived samples from Industry B indicated that the

SE and NW corners, and the recycle pond, were more closely related relative to the NE and SW corners, which were clearly separated. These results are consistent with the patterns described above for LC-QToF/MS (Figure 5), with the exception of the relative positions of the SW and NW corners, which may reflect differences in mass resolution, instrument configuration and/or data filtering of the two methods.

When profiled by LC-QToF/MS, samples from both the NE and NW corners contained a greater number of chemical entities and also a greater percentage of unique entities (26.7% and 27.4%, respectively) when compared to the other three sites (8.9-11.4%). This differentiation of the sample from the NE corner via these data was also consistent with the conclusions derived from the GCxGC-ToF/MS for the methyl esters of the monoaromatic acids . The further distinction of the NW corner sample, when profiled using LC-QToF data, is perhaps not surprising. Using LC, different classes of compounds (i.e., more polar compounds), may be more easily detected relative to GC methods. The aforementioned differences for the NE and NW spatial samples might be explained by the fact that the NE sample was collected closer to the discharge from the floatation tailings line, while the NW sample was collected near the discharge from the NRU tailings line and near the surface water runoff from the west side of the tailings pond. Other potential explanations for the observed differences could include differential zonal biological activity or sedimentation/adsorption of non-polar compounds at the NE corner.

The present investigation demonstrated that the ponds sampled from each of the two industries were distinguishable using data from either or both of the GC- and LC-based techniques employed. GC×GC-ToF/MS identification of known compounds showed previously that this was possible in a statistically defensible manner (Lengger et al., 2015). A total of 1138

compounds were considered significantly statistically different ($p<0.05$) herein between the OSPW AEO extracts from the two industries based on the data obtained by LC-QToF/MS, with approximately 1600 compounds common to both industries (Figure S5). Simple presence/absence, and presence in at least one sample, plotted in a Venn Diagram (Figure S4) also indicated that 1093 compounds were present in AEO fractions of OSPW from both industries, with 1639 compounds that were unique to one or other of the samples from Industries A and B.

In summary, all of the methods employed in this investigation were capable of identifying differences in the chemical profiles of samples collected from two different industrial oil sands tailings ponds. Chemical analyses of tentatively assigned mono-aromatic acids by GC×GC-ToF/MS, and analyses of unknown entities by GC- and LC-QToF/MS, were further capable of identifying differences between samples collected over a two-week interval from the same location as well as differences between samples collected at the same time from spatially different locations of a different pond, with the greatest differences being observed for the spatial sample set. The novel application of untargeted analyses by GC- and LC-QToF/MS that utilized -omics software to identify entities, both common and unique, throughout the sample set shows great promise for chemical source attribution initiatives, and when coupled with HRMS techniques, could prove vital for monitoring programs in the oil sands region. Further, this study highlights the presence of a yet untapped potential for collaborative approaches between untargeted analysis techniques and studies employing compound identification. This study also indicates a potential critical deficiency in research characterizing OSPW, as many studies likely collect water from a single location within a tailings pond for chemical and toxicological assessments, and/or extractions for targeted mixtures of compounds. Future research assessing

the extent of OSPW chemical variability within, and between, tailings ponds is needed, with greater temporal and spatial repetition, as well as incorporating discrete vertical sampling to assess chemical profile differences at different depths due to settling.

Disclaimer

The values of total NAs or AEF (Figure 1) measured should be considered semi-quantitative due to: a) a lack of certified AEF standards or alternatives and b) limited method validation and uncertainty estimates. The method and analytical parameters used in the present method can be found in SOPs available from the corresponding author. It should be understood that the present semi-quantitative method is a step in progression to a final quantitation goal.

Acknowledgements

This work was supported primarily by Environment Canada internal resources. The remainder was supported by an Advanced Investigators Grant (No. 228149) awarded to S.J. Rowland for project OUTREACH, by the European Research Council. Special thanks to Canadian Natural Resources Ltd. and Syncrude Canada Ltd. for sample collection.

Table 1. Workflow and associated entities for filtering data for LC & GC-QToF-MS profiling.

	LC-QToF/MS	**GC-QToF/MS**
Total Aligned Entities	5180	797
Present in at least 2/33 injections	4739	592
Present in all of at least 1 set of technical replicates	3651	330
Following ANOVA $p \geq 0.05$	3012	252
Number in Blank common to samples	280	78
Total Entities After Filtering and Blank Subtraction	**2732**	**174**

Table 2. Entities exclusive to each sample from Industry A or Industry B ponds for ESI(-ve)-LC-QToF. Entities are based on filtering in Table 1, with Venn diagrams used to determine exclusive entities.

	Temporal (Industry A)				**Spatial (Industry B)**		
	Number of entities unique to sample day	**Entities present on at least one other day**	**Total Entities present at site**		**Number of entities unique to sample site**	**Entities present in at least one other site**	**Total Entities present at site**
7th Nov	151 (13.9 %)	933 (86.1 %)	1084	NE	315 (30.4 %)	720 (69.6 %)	1035
10th Nov	240 (23.3 %)	791 (76.7 %)	1031	NW	305 (31.1 %)	675 (68.9 %)	980
14th Nov	88 (11.6 %)	670 (88.4 %)	758	SW	73 (9.7 %)	679 (90.3 %)	752
17th Nov	157 (16.8 %)	775 (83.2 %)	932	SE	83 (10.1 %)	742 (89.9 %)	825
21st Nov	52 (6.3 %)	772 (93.7 %)	824	Recycle	114 (12.5 %)	797 (87.5 %)	911

Figure Captions

Figure 1 Map depicting sampling locations for Industry A and B tailings ponds. Table inset provides sample contents of total naphthenic acids within the acid extractable fractions (AEF, mg/L).

Figure 2 HRMS speciation profiles for samples of Industry A (temporal) and B (spatial) tailings ponds. Ions classified in figures represent 90-93% of the total detected.

Figure 3 PCA scores plots of negative-ion ESI-HRMS for Industry A (temporal, red ellipse) and Industry B (spatial, green ellipse) samples. Ellipses were drawn for visualization purposes and do not represent statistically significant groupings, blue ellipse depicts samples clustering closely within the spatial study.

Figure 4 GC×GC-ToF/MS peak volumes for Family A (m/z 145) and Family B (m/z 237, 310) mono-aromatic acids (as methyl esters) normalized to sample volumes collected from Industry A (temporal) and B (spatial) tailings ponds.

Figure 5 PCA scores plots of negative-ion ESI-LC-QToF/MS (left) and positive-ion EI-GC-QToF/MS (right) for Industry A (temporal, red ellipses) and Industry B (spatial, green ellipses) samples. Ellipses were drawn for visualization purposes and do not represent statistically significant groupings, blue ellipse depicts samples clustering closely within the spatial study.

Figure S1 Synchronous fluorescence spectra of the samples of Industry A (temporal, in red) and B (spatial, in blue) tailings ponds.

Figure S2 3D view of m/z 145 mass chromatogram of Family A compounds, illustrating differences noted in NE corner of the pond from Industry B, relative to the NW corner.

Figure S3 PCA scores for each component of the temporal and spatial datasets.

Figure S4 Venn diagram for total entities both exclusive and shared between Industry A and Industry B detected by negative-ion ESI-LC-QToF/MS.

Figure S5 Volcano plot for fold changes between Industry A and Industry B (t-test, $p < 0.05$, Fold change >2.0). Each dot represents one entity (total=2732) Red dots are statistically different between the two industries (1138 entities).

References

Ahad, J. M. E. and H. Pakdel (2013). "Direct evaluation of in situ biodegradation in Athabasca oil sands tailings ponds using natural abundance radiocarbon." Environmental Science and Technology **47**: 10214-10222.

Anderson, J. C., S. B. Wiseman, A. Moustafa, M. Gamal El-Din, K. Liber and J. P. Giesy (2012). "Effects of exposure to oil sands process-affected water from experimental reclamation ponds on *Chironomus dilutus*." Water Research **46**(6): 1662-1672.

Anderson, J. C., S. B. Wiseman, N. Wang, A. Moustafa, L. Perez-Estrada, M. Gamal El-Din, J. W. Martin, K. Liber and J. P. Giesy (2012). "Effectiveness of ozonation treatment in eliminating toxicity of oil sands process-affected water to *Chironomus dilutus*." Environmental Science and Technology **46**: 486-493.

Barrow, M. P., K. M. Peru, B. Fahlman, L. M. Hewitt, R. A. Frank and J. V. Headley (2015). "Beyond naphthenic acids: environmental screening of water from the Athabasca region using APPI-FTICR mass spectrometry." Journal of the American Society of Mass Spectrometry **26**(9): 1508-1521.

Bletsou, A. A., J. Jeon, J. Hollender, E. Archontaki and N. S. Thomaidis (2015). "Targeted and non-targeted liquid chromatography-mass spectrometric workflows for identification of transformation products of emerging pollutants in the aquatic environment." Trends in Analytical Chemistry **66**: 32-44.

Bowman, D. T., G. F. Slater, L. A. Warren and B. E. McCarry (2014). "Identification of individual thiophene-, indane-, tetralin-, cyclohexane-, and adamantane-type carboxylic acids in composite tailings pore water from Alberta oil sands." Rapid Communications in Mass Spectrometry **28**: 2075-2083.

Brown, L. D. and A. C. Ulrich (2015). "Oil sands naphthenic acids: A review of properties, measurement, and treatment." Chemosphere **127**: 276-290.

Cumulative Environment Management Association (CEMA) (2012). End Pit Lakes Guidance Document**:** 436.

Del Rio, L. F., A. K. M. Hadwin, L. J. Pinto, M. D. MacKinnon and M. M. Moore (2006). "Degradation of naphthenic acids by sediment micro-organisms." Journal of Applied Microbiology **101**: 1049-1061.

Energy Resources Conservation Board (2010). News Release - ERCB approves Fort Hills and Syncrude Tailings Pond plans with conditions.
Energy Resources Conservation Board (2013). ST98-2013 Alberta's Energy Reserves 2012 and Supply/Demand Outlook 2013-2022**:** 282.
Energy Resources Conservation Board (2014). ST98-2014 Alberta's Energy Reserves 2013 and Supply/Demand Outlook 2014-2023**:** 289.
Energy Resources Conservation Board (2015). ST98-2015 Alberta's Energy Reserves 2014 and Supply/Demand Outlook 2015-2024**:** 299.
Ferguson, G. P., D. L. Rudolph and J. F. Barker (2009). "Hydrodynamics of a large oil sand tailings impoundment and related environmental implications." Canadian Geotechnological Journal **46**: 1446-1460.
Frank, R. A., J. W. Roy, G. Bickerton, S. J. Rowland, J. V. Headley, A. G. Scarlett, C. E. West, K. M. Peru, F. M. Conly and L. M. Hewitt (2014). "Profiling oil sands mixtures from industrial developments and natural groundwaters for source identification." Environmental Science and Technology **48**(5): 2660-2670.
FTFC (1995). Volume I: Clark hot water extraction fine tailings. In: Advances in oil sands tailings research. Alberta Department of Energy, Oil sands and research division, Edmonton, AB, Canada, FTFC (Fine Tailings Fundamentals Consortium).
FTFC (1995). Volume II: Fine tails and process water reclamation. In: Advances in oil sands tailings research. Alberta Department of Energy, Oil sands and research division, Edmonton, AB, Canada, FTFC (Fine Tailings Fundamentals Consortium).
Gagné, F., M. Douville, C. André, T. Debenest, A. Talbot, J. Sherry, L. M. Hewitt, R. A. Frank, M. E. McMaster, J. Parrott and G. Bickerton (2012). "Differential changes in gene expression in rainbow trout hepatocytes exposed to extracts of oil sands process-affected water and the Athabasca River." Comparative Biochemistry and Physiology C Comparative Pharmacology **155**: 551-559.
Gao, W., H. Yang, L.-W. Qi, E.-H. Liu, M.-T. Ren, Y.-T. Yan, J. Chen and P. Li (2012). "Unbiased metabolite profiling by liquid chromatography–quadrupole time-of-flight mass spectrometry and multivariate data analysis for herbal authentication: Classification of seven Lonicera species flower buds." Journal of Chromatography A **1245**: 109-116.

Garcia-Garcia, E., J. Q. Ge, A. Oladiran, B. Montgomery, M. Gamal El-Din, L. Perez-Estrada, J. L. Stafford, J. W. Martin and M. Belosevic (2011). "Ozone treatment ameliorates oil sands process water toxicity to the mammalian immune system." Water Research **45**: 5849-5857.

Gentes, M.-L., C. Waldner, Z. Papp and J. E. G. Smits (2006). "Effects of oil sands tailings compounds and harsh weather on mortality rates, growth and detoxification efforts in nestling tree swallows (*Tachycineta bicolor*)." Environmental Pollution **142**: 24-33.

Han, X., M. D. MacKinnon and J. W. Martin (2009). "Estimating the in situ biodegradation of naphthenic acids in oil sands process waters by HPLC/HRMS." Chemosphere **76**: 63-70.

Han, X., A. C. Scott, P. M. Fedorak, M. Bataineh and J. W. Martin (2008). "Influence of molecular structure on the biodegradability of naphthenic acids." Environmental Science and Technology **42**: 1290-1295.

Hao, C., J. V. Headley, K. M. Peru, R. A. Frank, P. Yang and K. R. Solomon (2005). "Characterization and pattern recognition of oil-sand naphthenic acids using comprehensive two-dimensional gas chromatography/ time-of-flight mass spectrometry." Journal of Chromatography A **1067**: 277-284.

He, Y., S. Patterson, N. Wang, M. Hecker, J. W. Martin, M. Gamal El-Din, J. P. Giesy and S. B. Wiseman (2012). "Toxicity of untreated and ozone-treated oil sands process-affected water (OSPW) to early life stages of the fathead minnow (*Pimephales promelas*)." Water Research **46**: 6359-6368.

Headley, J. V., M. P. Barrow, K. M. Peru, B. Fahlman, R. A. Frank, G. Bickerton, M. E. McMaster, J. Parrott and L. M. Hewitt (2011). "Preliminary fingerprinting of Athabasca oil sands polar organics in environmental samples using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry." Rapid Communications in Mass Spectrometry **25**: 1899-1909.

Headley, J. V., K. M. Peru, S. A. Armstrong, X. Han, J. W. Martin, M. M. Mapolelo, D. F. Smith, R. P. Rogers and A. G. Marshall (2009). "Aquatic plant-derived changes in oil sands naphthenic acid signatures determined by low-, high-, and ultrahigh-resolution mass spectrometry." Rapid Communications in Mass Spectrometry **23**(4): 515-522.

Headley, J. V., K. M. Peru, M. H. Mohamed, R. A. Frank, J. W. Martin, R. R. O. Hazewinkle, D. Humphries, N. P. Gurprasad, L. M. Hewitt, D. C. G. Muir, D. Lindeman, R. Strub, R. F. Young, D. M. Grewer, R. M. Whittal, P. M. Fedorak, D. A. Birkholz, R. Hindle, R. Reisdorph, X. Wang,

K. L. Kasperski, C. Hamilton, M. Woudneh, G. Wang, B. Loescher, A. Farwell, D. G. Dixon, M. S. Ross, A. Dos Santos Pereira, E. King, M. P. Barrow, B. Fahlman, J. Bailey, D. W. McMartin, C. H. Borchers, C. H. Ryan, N. S. Toor, H. M. Gillis, L. Zuin, G. Bickerton, M. E. McMaster, E. Sverko, D. Shang, L. D. Wilson and F. J. Wrona (2013). "Chemical fingerprinting of oil sands naphthenic acids in environmental samples - A review of analytical methods." Journal of Environmental Science and Health A **48**: 1145-1163.

Holowenko, F. M., M. D. MacKinnon and P. M. Fedorak (2002). "Characterization of naphthenic acids in oil sands wastewaters by gas chromatography-mass spectrometry." Water Research **36**: 2843-2855.

Hunter, G. P. (2001). Investigation of groundwater flow within an oil sands tailings impoundment and environmental implications. MSc, University of Waterloo.

Kavanagh, R., B. K. Burnison, R. A. Frank, K. R. Solomon and G. Van Der Kraak (2009). "Detecting oil sands process-affected waters in the Alberta oil sands region using synchronous fluorescence spectroscopy." Chemosphere **76**: 120-126.

Kavanagh, R. J., 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 (2011). "Fathead minnow (*Pimephales promelas*) reproduction is impaired in aged oil sands process-affected waters." Aquatic Toxicology **101**: 214-220.

Leclair, L. A., G. Z. MacDonald, L. J. Phalen, B. Köllner, N. S. Hogan and M. R. van den Heuvel (2013). "The immunological effects of oil sands surface waters and naphthenic acids on rainbow trout (*Oncorhynchus mykiss*)." Aquatic Toxicology **142-143**: 185-194.

Lengger, S. K., A. G. Scarlett, C. E. West, R. A. Frank, L. M. Hewitt, C. M. Milestone and S. J. Rowland (2015). "Use of the distributions of adamantane acids to profile short-term temporal and small-scale spatial variations in the composition of oil sands process-affected waters." Environmental Science: Processes and Impacts **17**(8): 1415-1423.

MacKinnon, M. D., G. Kampala, B. Marsh, P. M. Fedorak and S. Guigard (2005). Indicators for assessing transport of oil sands process-affected waters. Bringing groundwater quality research to the watershed (Proceedings of GQ2004 International Conference of Groundwater Quality). N. R. Thomson, IAHS Press. **297:** 71-80.

Madill, R. E. A., M. T. Orzechowski, G. Chen, B. G. Brownlee and N. J. Bunce (2001). "Preliminary risk assessment of the wet landscape option for reclamation of oil sands mine tailings: bioassays with mature fine tailing pore water." Environmental Toxicology **16**: 197-208.

McKenzie, N., S. Yue, X. Liu, B. A. Ramsay and J. A. Ramsay (2014). "Biodegradation of naphthenic acids in oil sands process waters in an immobilized soil/sediment bioreactor." Chemosphere **109**: 164-172.

Moustafa, A. M. A., E.-S. Kim, A. Alpatova, N. Sun, S. Smith, S. Kang and M. Gamal El-Din (2014). "Impact of polymeric membrane filtration of oil sands process water on organic compounds quantification." Water Science and Technology **70.5**: 771-779.

Nero, V., A. Farwell, L. E. J. Lee, T. Van Meer, M. D. MacKinnon and D. G. Dixon (2006). "The effects of salinity on naphthenic acid toxicity to yellow perch: Gill and liver histopathology." Ecotoxicology and Environmental Safety **65**: 252-264.

Oiffer, A. A. L., J. F. Barker, F. M. Gervais, K. U. Mayer, C. J. Ptacek and D. L. Rudolph (2009). "A detailed field-based evaluation of naphthenic acid mobility in groundwater." Journal of Contaminant Hydrology **108**: 89-106.

Oil Sands Wetlands Working Group - Alberta Environment (2000). Guideline for Wetland Establishment on Reclaimed Oil Sands Leases - Report #ESD/LM/00-1; Pub No. T/517.

Rogers, V. V., M. Wickstrom, K. Liber and M. D. MacKinnon (2002). "Acute and subchronic mammalian toxicity of naphthenic acids from oil sands tailings." Toxicological Sciences **66**: 347-355.

Ross, M. S., A. dos Santos Pereira, J. Fennell, M. Davies, J. Johnson, L. Sliva and J. W. Martin (2012). "Quantitative and qualitative analysis of naphthenic acids in natural waters surrounding the Canadian oil sands industry." Environmental Science and Technology **46**: 12796-12805.

Rowland, S. J., A. G. Scarlett, D. Jones, C. E. West and R. A. Frank (2011). "Diamonds in the rough: Identification of individual naphthenic acids in oil sands process water." Environmental Science and Technology **45**: 3154-3159.

Rowland, S. J., C. E. West, D. Jones, A. G. Scarlett, R. A. Frank and L. M. Hewitt (2011). "Steroidal aromatic 'naphthenic acids' in oil sands process-affected water: Structural comparisons with environmental estrogens." Environmental Science and Technology **45**: 9806-9815.

Rowland, S. J., C. E. West, A. G. Scarlett, C. Ho and D. Jones (2012). "Differentiation of two industrial oil sands process-affected waters by two-dimensional gas chromatography/mass

spectrometry of diamondoid acid profiles." Rapid Communications in Mass Spectrometry **26**: 572-576.

Rowland, S. J., C. E. West, A. G. Scarlett, D. Jones and R. A. Frank (2011). "Identification of individual tetra- and pentacyclic naphthenic acids in oil sands process water by comprehensive two-dimensional gas chromatography-mass spectrometry." Rapid Communications in Mass Spectrometry **25**: 1198-1204.

Vaclavik, L., O. Lacina, J. Hajslova and J. Zweigenbaum (2011). "The use of high performance liquid chromatography–quadrupole time-of-flight mass spectrometry coupled to advanced data mining and chemometric tools for discrimination and classification of red wines according to their variety." Analytica Chimica Acta **685**: 45-51.

van den Heuvel, M. R., N. S. Hogan, S. Roloson and G. J. Van Der Kraak (2012). "Reproductive development of yellow perch (*Perca flavescens*) exposed to oil sands-affected waters." Environmental Toxicology and Chemistry **31**(3): 654-662.

Wilde, M. J. and S. J. Rowland (2015). "Structural Identification of Petroleum Acids by Conversion to Hydrocarbons and Multidimensional Gas Chromatography-Mass Spectrometry." Analytical Chemistry ***(in press)***.

Wilde, M. J., C. E. West, A. G. Scarlett, D. Jones, R. A. Frank, L. M. Hewitt and S. J. Rowland (2015). "Bicyclic naphthenic acids in oil sands process water: Identification by comprehensive multidimensional gas chromatography-mass spectrometry." Journal of Chromatography A **1378**: 74-87.

Wiseman, S. B., Y. He, M. Gamal El-Din, J. W. Martin, P. D. Jones, M. Hecker and J. P. Giesy (2013). "Transcriptional responses of male fathead minnows exposed to oil sands process-affected water." Comparative Biochemistry and Physiology C Comparative Pharmacology **157**: 227-235.

Zubot, W. A. (2010). Removal of Naphthenic Acids from Oil Sands Process Water using Petroleum Coke, MSc Thesis in Environmental Engineering, University of Alberta, Dept of Civil and Environmental Engineering**:** 169 pages.

Figures and Supplementary Information

Title: Assessing spatial and temporal variability of acid-extractable organics in oil sands process-affected waters

Authors:

Richard A. Frank[a], Craig B. Milestone[a(b#)], Steve J. Rowland[c], John V. Headley[d], Richard J. Kavanagh[e], Sabine K. Lengger[c], Alan G. Scarlett[c], Charles E. West[c(f#)], Kerry M. Peru[d], and L. Mark Hewitt[a]

[a]Water Science and Technology Directorate, Environment Canada, 867 Lakeshore Road, Burlington, ON, Canada L7S 1A1

[b]Faculty of Applied Science and Technology, Sheridan College, 7899 McLaughlin Road Brampton, ON, Canada L6Y 5H9

[c]Petroleum and Environmental Geochemistry Group, Biogeochemistry Research Centre, University of Plymouth, Drake Circus, 5 Plymouth PL4 8AA, UK

[d]Water Science and Technology Directorate, Environment Canada, 11 Innovation Boulevard, Saskatoon, SK, Canada, S7N3H5

[e]Canadian Natural Resources Ltd., Calgary, AB

[f]EXPEC Advanced Research Center, Saudi Aramco, Dhahran 31311, Saudi Arabia

[#]**Present Address**

[*]**Corresponding Author:** Richard A. Frank

WSTD, Environment Canada

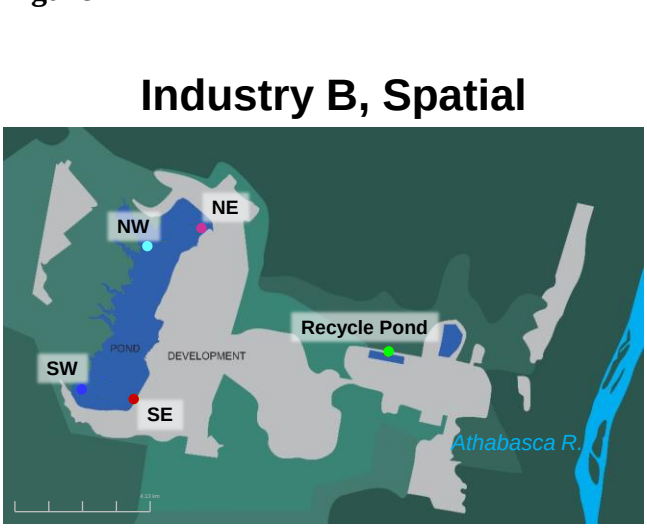
867 Lakeshore Road, Burlington, ON, Canada L7S 1A1

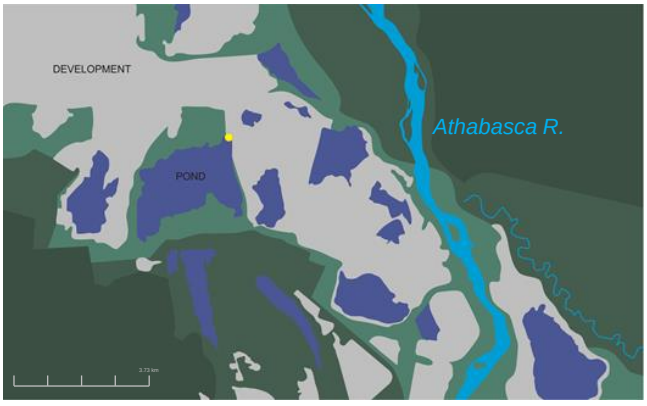
Phone: (905)336-4470 / Fax: (905)336-6430

Email: richard.frank@canada.ca

Manuscript Figures (Figure captions included with manuscript text)

Figure 1





Sample Description	[AEF] mg/L
NW	35.3
NE	25.5
SE	32.2
SW	24.4
Recycle Pond	38.4
07 Nov	41.8
10 Nov	37.6
14 Nov	43.3
17 Nov	41.2
21 Nov	51.5

Figure 2

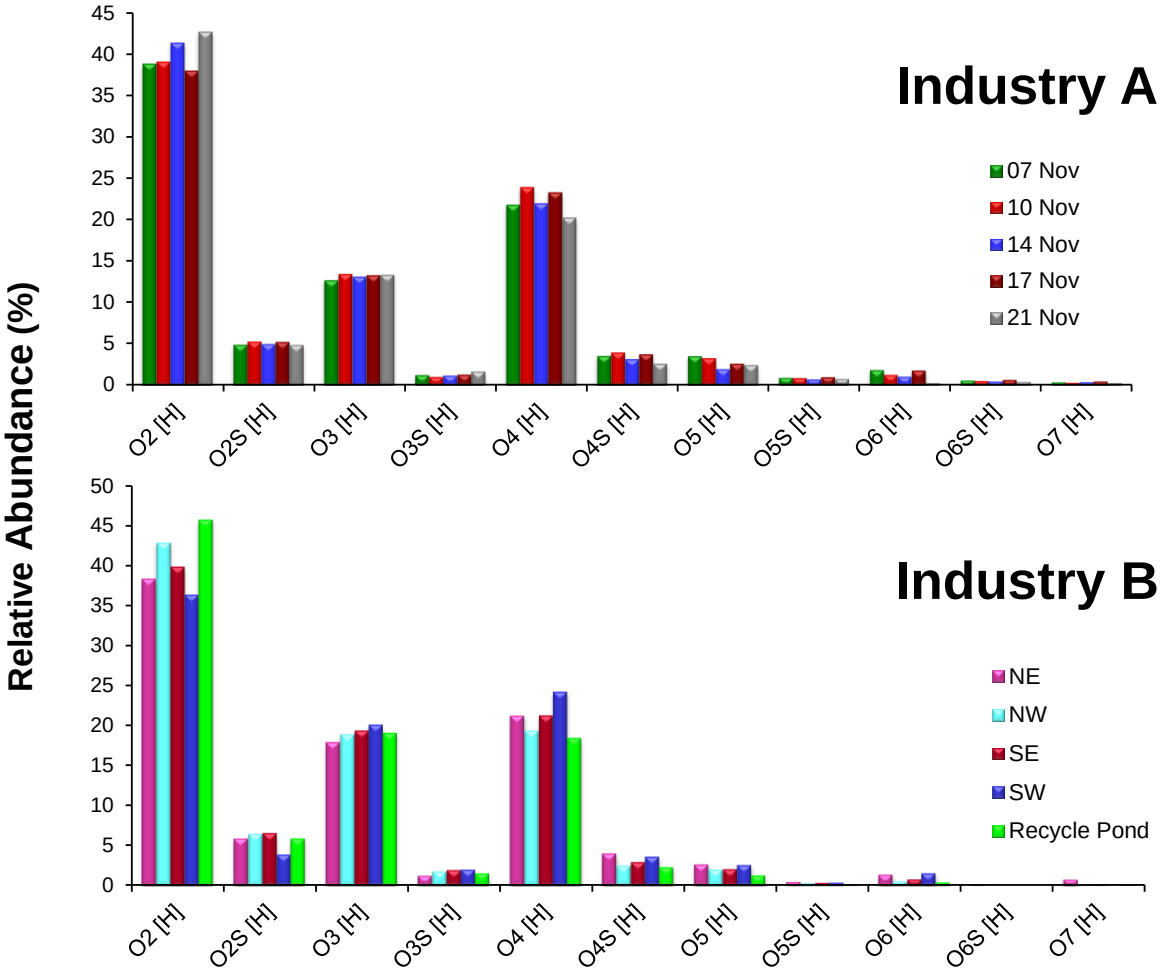

Relative Abundance (%)
Industry A
07 Nov
10 Nov
14 Nov
17 Nov
21 Nov
O2 [H]
O2S [H]
O3 [H]
O3S [H]
O4 [H]
O4S [H]
O5 [H]
O5S [H]
O6 [H]
O6S [H]
O7 [H]
Industry B
NE
NW
SE
SW
Recycle Pond

Figure 3

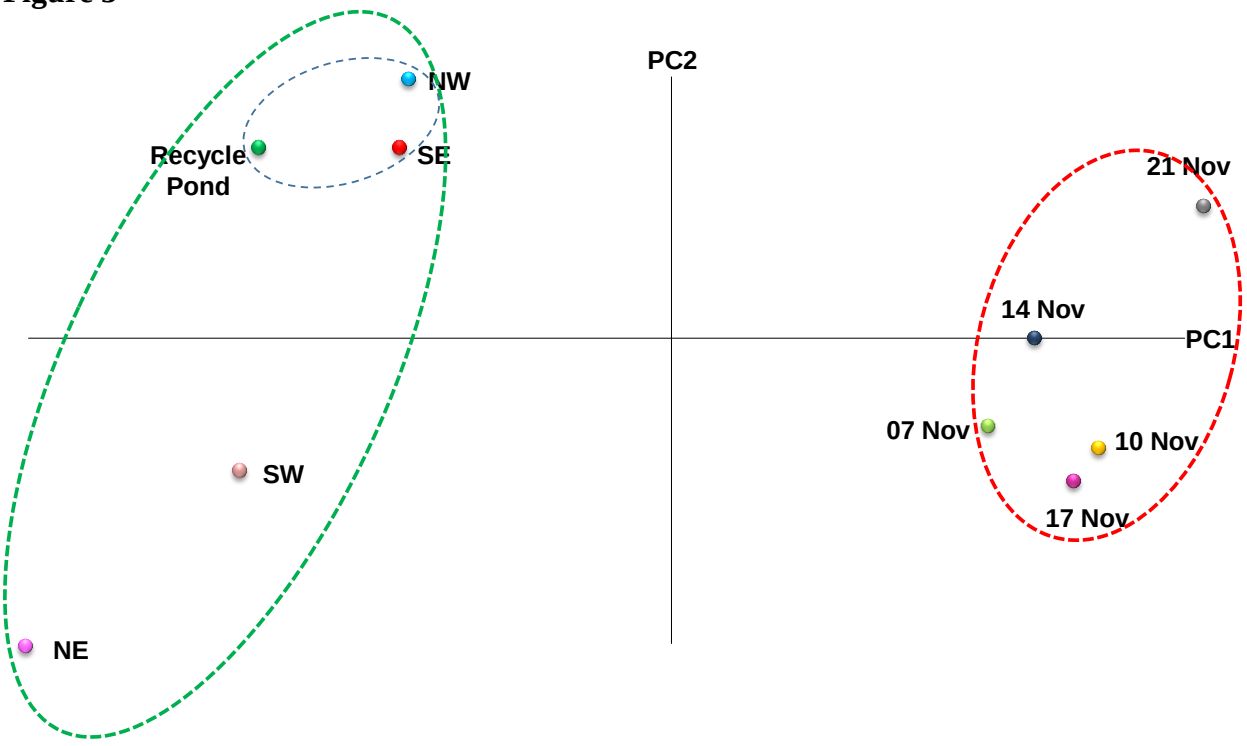

PC2
PC1
NW
SE
Recycle
Pond
SW
NE
21 Nov
14 Nov
07 Nov
10 Nov
17 Nov

Figure 4

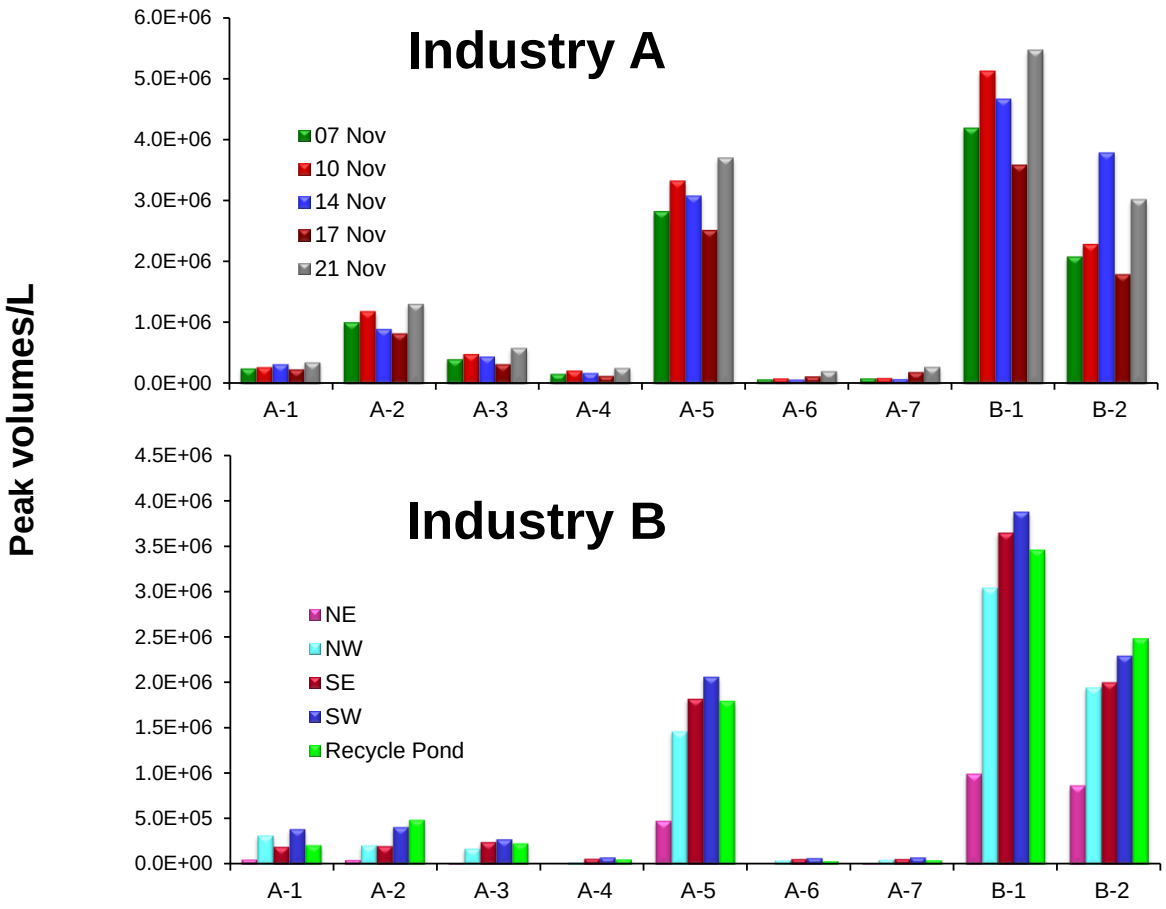
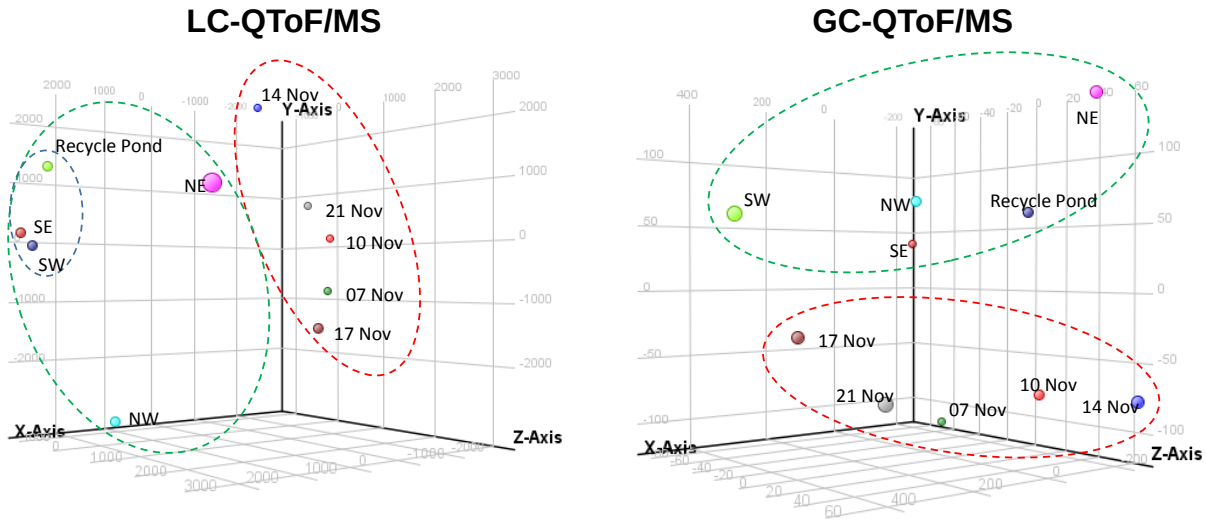

Peak volumes/L
Industry A
6.0E+06
5.0E+06
4.0E+06
3.0E+06
2.0E+06
1.0E+06
0.0E+00
07 Nov
10 Nov
14 Nov
17 Nov
21 Nov
A-1
A-2
A-3
A-4
A-5
A-6
A-7
B-1
B-2
Industry B
4.5E+06
4.0E+06
3.5E+06
3.0E+06
2.5E+06
2.0E+06
1.5E+06
1.0E+06
5.0E+05
0.0E+00
NE
NW
SE
SW
Recycle Pond
A-1
A-2
A-3
A-4
A-5
A-6
A-7
B-1
B-2

Figure 5


LC-QToF/MS
14 Nov
Y-Axis
Recycle Pond
NE
21 Nov
SE
10 Nov
SW
07 Nov
17 Nov
NW
X-Axis
Z-Axis
GC-QToF/MS
Y-Axis
NE
SW
NW
Recycle Pond
SE
17 Nov
10 Nov
21 Nov
07 Nov
14 Nov
X-Axis
Z-Axis

Supplemental Information (Figure captions included with manuscript text)

Figure S1

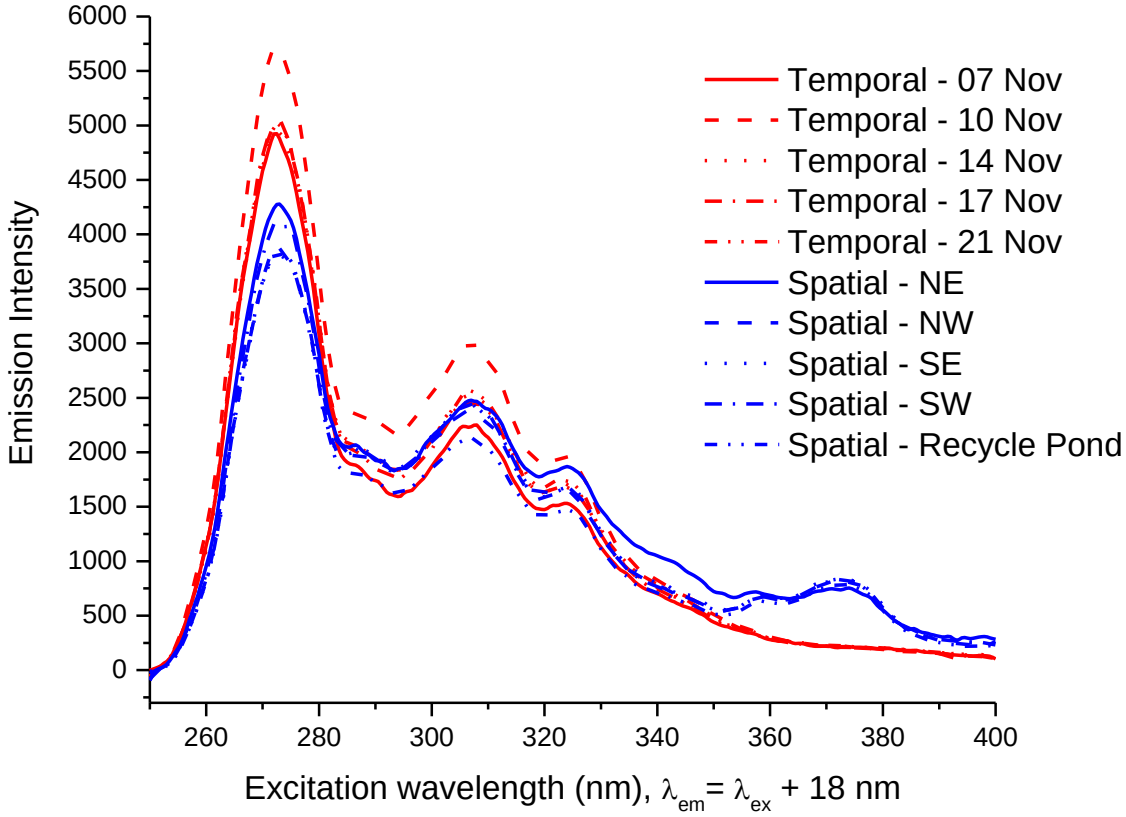


Figure S2

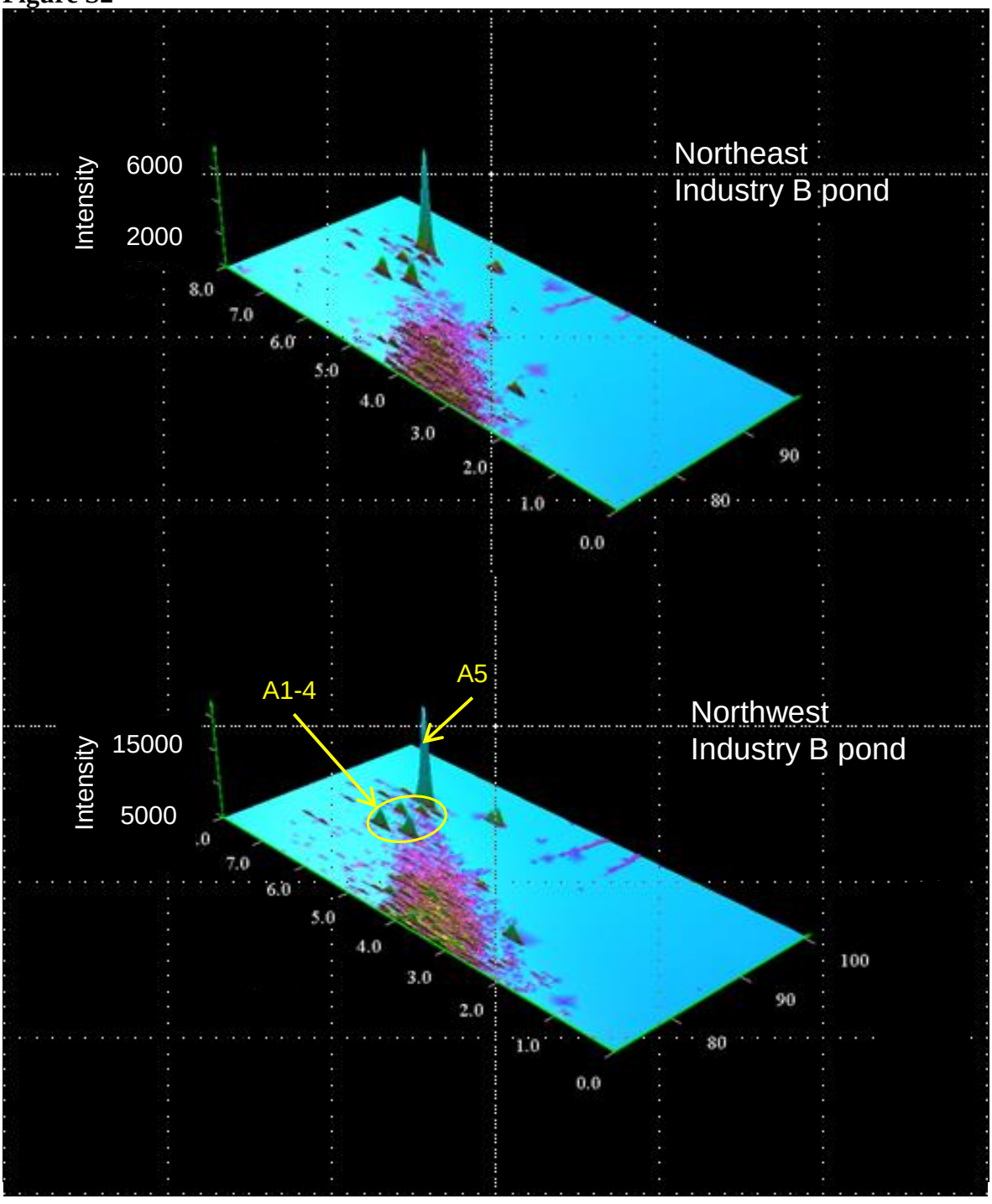
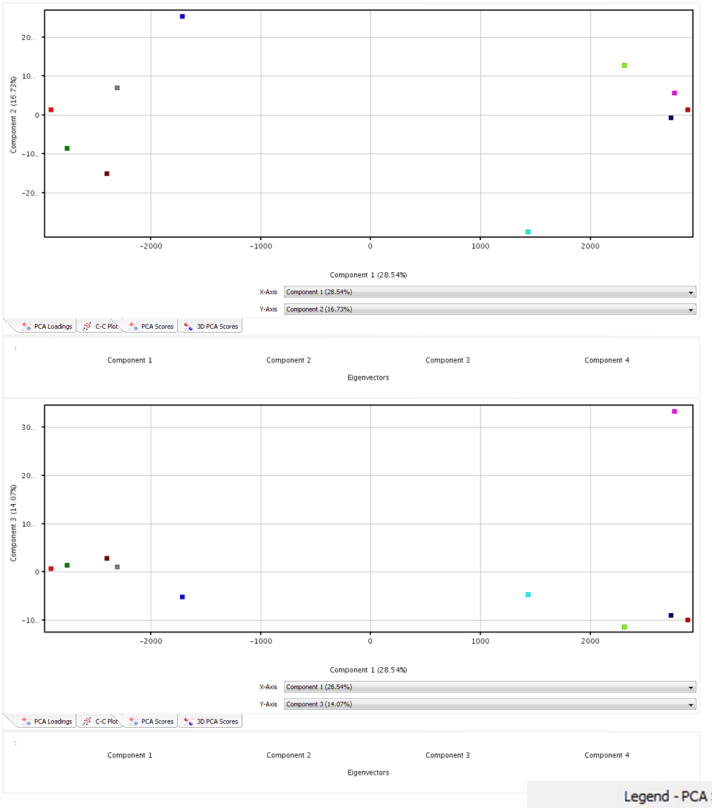

Intensity
6000
2000
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
90
80
Northeast
Industry B pond
A1-4
A5
Intensity
15000
5000
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
100
90
80
Northwest
Industry B pond

Figure S3

A – LC-QToF/MS



B – GC-QToF/MS

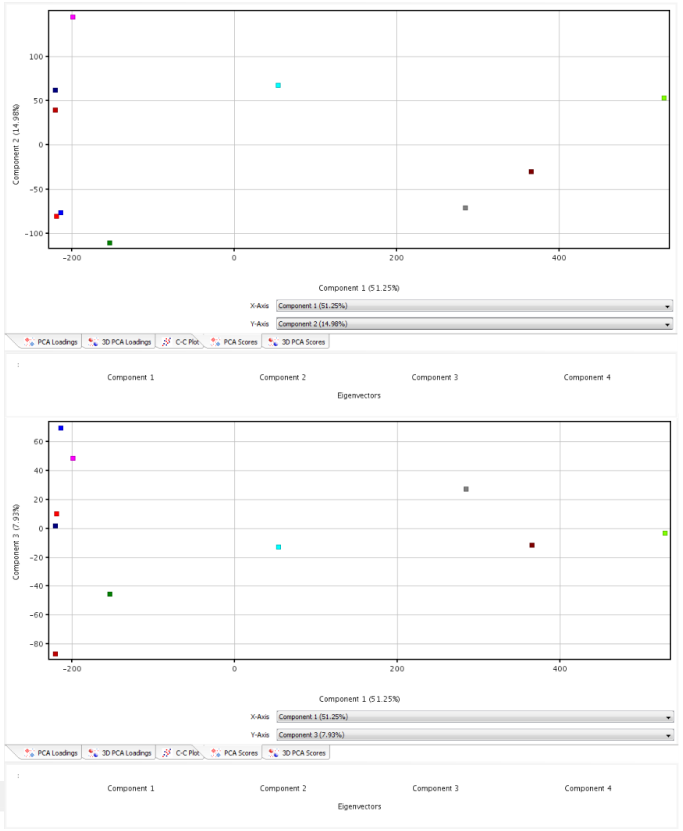


Figure S4

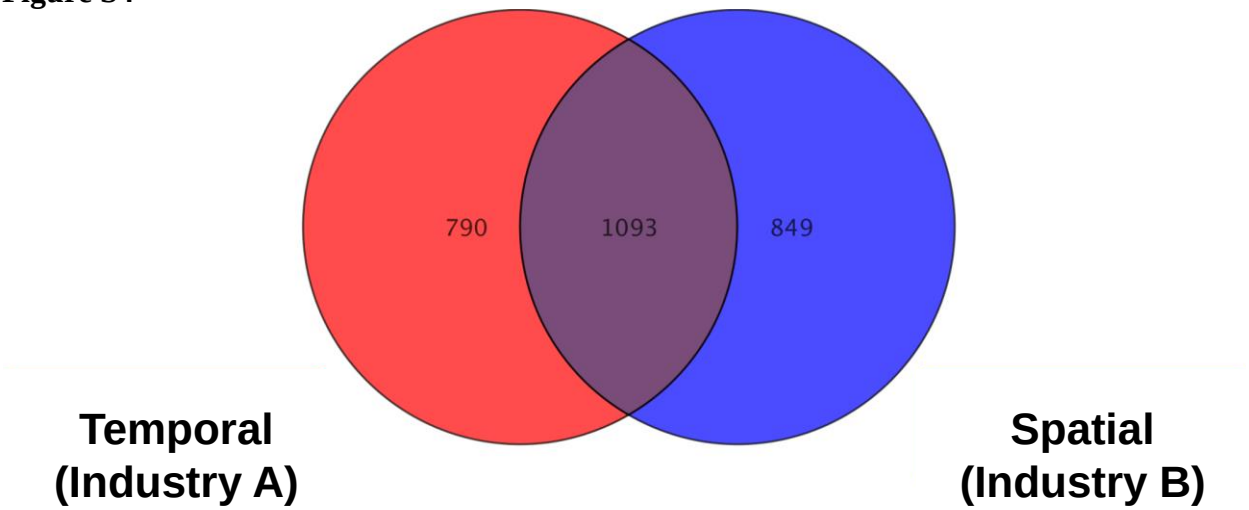
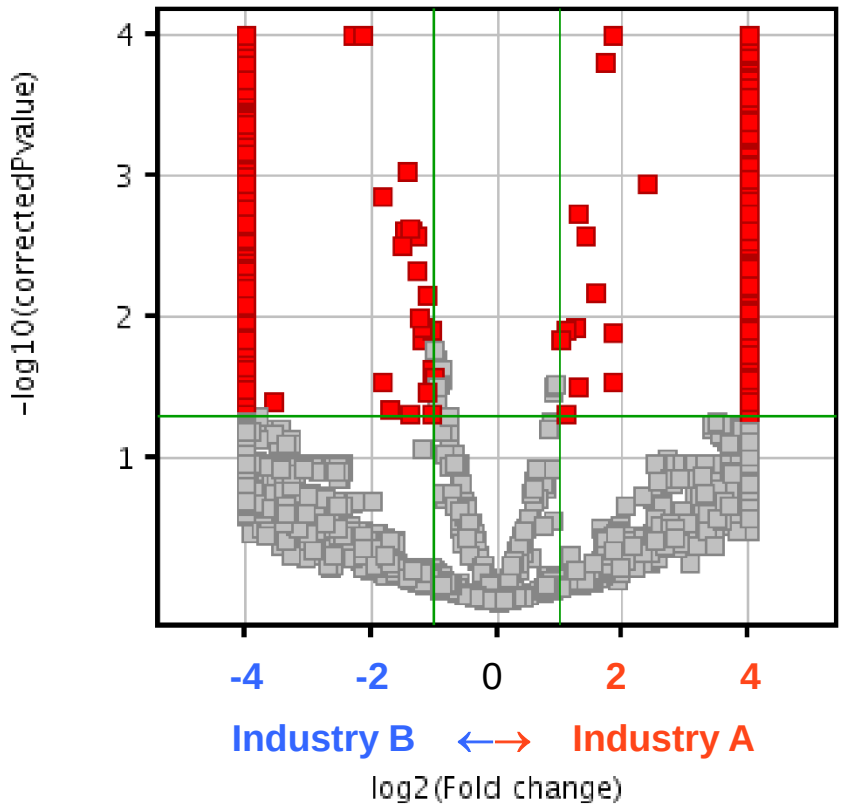


Figure S5


-log10(correctedPvalue)
4
3
2
1
-4
-2
0
2
4
Industry B ←→ Industry A
log2(Fold change)