



Characterization of organic composition in snow and surface waters in the Athabasca Oil Sands Region, using ultrahigh resolution Fourier transform mass spectrometry



Y. Yi^{a,c,*}, S.J. Birks^{b,c}, S. Cho^d, J.J. Gibson^{a,c}

^a Alberta Innovates-Technology Futures, Victoria, British Columbia, Canada

^b Alberta Innovates-Technology Futures, Calgary, Alberta, Canada

^c Department of Geography, University of Victoria, Victoria, British Columbia, Canada

^d Clean Energy Branch, Policy Division, Alberta Environment and Sustainable Resource Development, Edmonton, Alberta, Canada

HIGHLIGHTS

- Near-field snow samples were compositionally different from far-field snow
- The composition of Athabasca river upstream demonstrated similarity to far-field snow
- No evidence that Athabasca River become more similar to near-field snow
- Compound classes potentially indicating sources in AOSR are discussed

ARTICLE INFO

Article history:

Received 19 September 2014

Received in revised form 31 January 2015

Accepted 5 February 2015

Available online 6 March 2015

Editor: Kevin V. Thomas

Keywords:

Snow

Surface water

Dissolved organic composition

ESI-FTICR MS

Athabasca Oil Sands Region

ABSTRACT

This study was conducted to characterize the composition of dissolved organic compounds present in snow and surface waters in the Athabasca Oil Sands Region (AOSR) with the goal of identifying whether atmospherically-derived organic compounds present in snow are a significant contributor to the compounds detected in surface waters (i.e., rivers and lakes). We used electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICR MS) to characterize the dissolved organic compound compositions of snow and surface water samples. The organic profiles obtained for the snow samples show compositional differences between samples from near-field sites (<5 km from oil sands activities) and those from more distant locations (i.e., far-field sites). There are also significant compositional differences between samples collected in near-field sites and surface water samples in the AOSR. The composition of dissolved organic compounds at the upstream Athabasca River site (i.e., Athabasca River at Athabasca) is found to be different from samples obtained from downstream sites in the vicinity of oil sands operations (i.e., Athabasca River at Fort McMurray and Athabasca River at Firebag confluence). The upstream Athabasca River sites tended to share some compositional similarities with far-field snow deposition, while the downstream Athabasca River sites are more similar to local lakes and tributaries. This contrast likely indicates the relative role of regional snowmelt contributions to the Athabasca River vs inputs from local catchments in the reach downstream of Fort McMurray.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Understanding the impacts of human activities on aquatic ecosystems in the Athabasca Oil Sands Region (AOSR) is required to ensure sustainable development in northeastern Alberta and to assess the economic benefit of oil sands development to Alberta and Canada. Atmospheric deposition, including wet and dry deposition of organic acids and nitrogen and sulfur-containing compounds derived from bitumen

refining operations and fleet vehicles, can reach local aquatic ecosystems directly or may be entrained and transported through water cycling processes (Timoney and Lee, 2009; Hall et al., 2012; Kurek et al., 2013). Interaction with bitumen deposits, groundwater and other sources of organic compounds within the catchment may also occur, which gives rise to the potential for a complex pathway and transport of organic compounds. By comparing the organic compounds present in snow (i.e., representative of direct atmospheric deposition) with those in surface water samples (i.e., rivers and lakes), we gain insight into the dissolved organic composition in surface waters, which captures a different proportion of input from atmospheric, local catchment and groundwater sources of these waters at various sites, and can develop a

* Corresponding author at: Alberta Innovates-Technology Futures, Vancouver Island Technology Park, 3–4476 Markham St., Victoria, BC V8Z 7X8, Canada.

E-mail address: yiyi@uvic.ca (Y. Yi).

better understanding of the contribution of atmospherically-derived organic compounds to surface waters across the region.

Snow surveys provide a means of collecting an integrated sample of atmospherically-derived organic compounds deposited during the winter season (Kelly et al., 2010) and can be used to quantitatively evaluate the spatial distribution of atmospheric deposition across broad regions (Cho et al., 2014). In northern Alberta the snowmelt that occurs during the spring freshet is the most important hydrological event of the year, and results in peak flow in the Athabasca River and its tributaries along with transport and movement of particulate and dissolved substances within watersheds (Woo and Thorne, 2003). Here we use the composition of dissolved organic compounds present in snow to represent atmospherically-derived deposition in the region.

During the past several years, polycyclic aromatic hydrocarbons (PAHs) have been the focus of many investigations in the AOSR (Cho et al., 2014; Hall et al., 2012; Jautzy et al., 2013; Studabaker et al., 2012; Kelly et al., 2009; Kurek et al., 2013; Wang et al., 2014), because of the potentially adverse effects these compounds can have on terrestrial and aquatic organisms and human health (Colavecchia et al., 2004; Martin et al., 2005, and Vehniäinen et al., 2003). Elevated concentrations of polycyclic aromatic hydrocarbons (PAHs) have been identified in river water and sediments in the AOSR. Airborne pathways have been implicated as the main mechanism for transport of these contaminants from the AOSR industrial development into the aquatic ecosystem (Kurek et al., 2013; Timoney and Lee, 2011). However, other studies have suggested that transport by the Athabasca River, a waterborne pathway, is also important for bringing PAHs and other contaminants to the Peace Athabasca Delta located approximately 200 km downstream of major oil sands development (Hall et al., 2012; Wiklund et al., 2012). Although detailed measurements and documentation of PAH speciation in snow from the region are an important aspect of monitoring atmospheric deposition in AOSR (Wang et al., 2014; Cho et al., 2014), we are not aware of any other attempts to broadly characterize the full spectrum of organic compounds present in atmospheric deposition in the AOSR. The contribution of atmospherically-derived compounds to the overall organic load present in the Athabasca River, tributaries and lakes is still an important knowledge gap and improved understanding of multiple sources and transport pathways for organic compounds in the region is required.

New analytical techniques, like ultrahigh resolution mass spectrometry, have greatly improved our ability to detect and characterize polar organic compounds in environmental samples (Headley et al., 2011, 2013; Mazzoleni et al., 2010). Electrospray ionization (ESI) coupled with Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) provides more comprehensive characterization of the dissolved organic compounds due to its extremely high resolution ($m/\Delta m_{50\%} > 500,000$, in which m is ion mass and $\Delta m_{50\%}$ is mass spectral peak full width at half-maximum height) and mass accuracy (< 1 ppm), allowing for an improved ability to identify different organic species, particularly at very low concentrations such as those found in surface waters. FTICR MS has been used to characterize the organic compounds present in oil sands process affected waters (OSPW), and samples from lakes, rivers and groundwaters (Barrow et al., 2010; Grewer et al., 2010; Gibson et al., 2011; Headley et al., 2011; Ross et al., 2012; Headley et al., 2013; Nyakas et al., 2013; Frank et al., 2014), however ESI-FTICR MS has not been previously used to profile atmospheric deposition (e.g., snow) which often believed to have only trace concentrations of dissolved organic compounds.

The objectives of this study were to conduct a comparative assessment of the dissolved polar organic compounds present in snow and surface water samples in the AOSR, and to investigate if similarities and/or differences in the composition of snow and surface water samples can be used to identify and evaluate potential sources and pathways of dissolved organic compounds.

2. Methodology

2.1. Collection of snow and surface water samples

This study uses 79 surface water samples (i.e., 73 Athabasca River and tributary samples, and 6 lake samples) and 7 snow samples, collected by Alberta Environment and Sustainable Resource Development (AESRD) during various snow and surface water monitoring programs in 2012. The study area (Fig. 1), broadly covering the region downstream to Ft. McMurray in the Athabasca watershed, which is also rapidly developed for oil sands development, has been described previously by Cho et al. (2014).

Snow samples were collected in March 2012 (Fig. 1), using a prairie corer (7 cm in diameter and 111.75 cm in length). The snow from the corer was emptied into Teflon bags, and subsequently melted in a dark room at a temperature of 22 °C. Detailed snow sampling procedure is described in Cho et al. (2014). A geographical centre (GC) of oil sands operations (Fig. 1, black triangle, 57° 1.5'N, 111° 33.0'W) is used for geographic reference (Cho et al., 2014). The snow sampling locations included 3 far-field locations (i.e., NW-S8, E-S7, NW-S2), located distantly from the major oil sands activity area (61 km, 38 km and 30 km from GC, respectively) and 4 near-field locations (<5 km from GC) along the Athabasca and Steepbank Rivers (i.e., KAR-4, KAR-6a, KAR-16 and KST-2), located much closer to the centre of oil sands activities (2 km to 4 km from GC).

Six lake samples were collected by AESRD personnel during surveys conducted using a float-equipped helicopter in August 2012 (Fig. 1). Sampling was conducted using AESRD water quality sampling protocols (AENV, 2006) designed to minimize contamination. The samples were taken from about 10 cm below the water surface at the centre of the lake. Lake water temporally integrates inputs from surface runoff, groundwater, and atmospheric deposition over each catchment and should reflect a mixture of compounds from all of these sources. The lakes sampled were fairly small and shallow (maximum depth less than 4 m, except lake BM2), with residence times between 4 months and 7 years (Bennett et al., 2008). The lake sampling was scheduled for late summer, when thermal stratification would have broken down, and the lakes would be expected to be well-mixed so that the sampling would be most representative of average lake conditions. The only exception to this is lake BM2 (also known as Namur Lake), which is much larger and deeper (maximum depth 27.4 m) than the other lakes. The sample obtained from BM2 should be considered representative of the epilimnion of the lake, (approximately upper 12 m, RAMP, 2000).

There were a total of 23 river sampling sites, 17 that were regularly sampled and 6 additional sites where samples were only taken periodically (Fig. 1 and Table 1). Sampling sites included locations along the Athabasca River as well as several tributaries that flow through the oil sands development area. The Athabasca River samples included one site upstream of all oil sands development (Athabasca River at Athabasca, Atha-3), a site at Fort McMurray upstream of most of the mining activity (Athabasca River at Fort McMurray, Atha-2) and three sites downstream of development (Athabasca River upstream of Firebag River, Atha-1; at Old Fort, Old Fort; and at Devil's Elbow, Dev). The tributary sampling sites included locations near the mouth of the Firebag (Fir-1), Ells (Ells-1), Muskeg (Mus-3), MacKay (Mac), Steepbank (Ste-1) and Clearwater Rivers (Cle-1) as well as some upstream sites on these tributaries. The 6 river sites that were sampled less frequently were Firebag River near winter road (Fir-2), Steepbank River 5 km upstream Athabasca River (Ste-2), Clearwater River 2 km d/s Christina River (Cle-2), Ells River at Ft. McKay Pump House (Ells-2), Athabasca River at Devil's Elbow (Dev), and Athabasca River at (Old Fort). AESRD attempted to collect monthly samples from the 17 regular river sampling locations between March and September 2012. However, logistical difficulties and resource limitations prevented monthly sample analysis at all sampling sites. Unfortunately, river samples were not

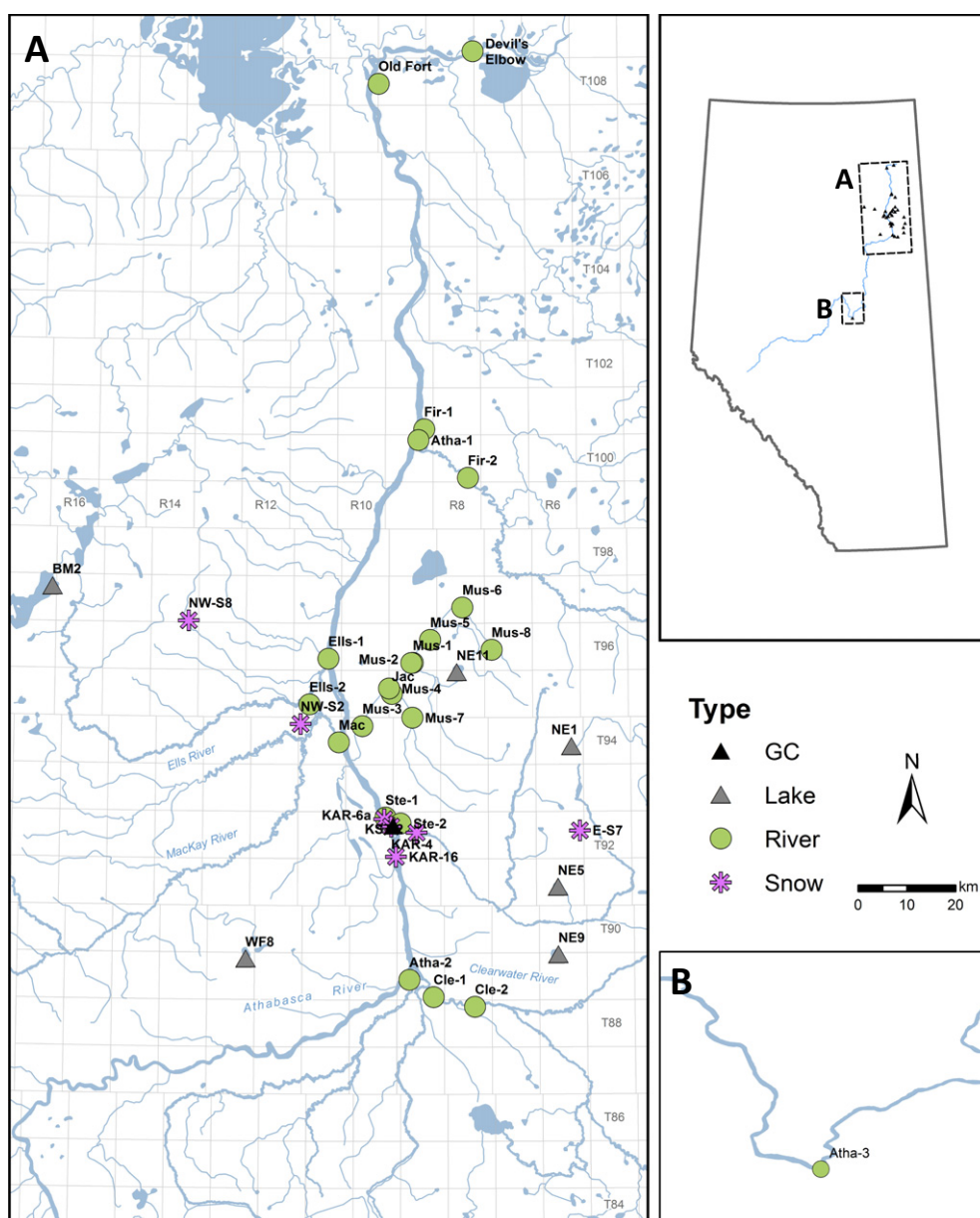


Fig. 1. Sampling locations of snow, river and lake samples in 2012.

available for organic profiling for April and May, and the March samples did not include any from the Athabasca River main stem. To best capture the temporal variability in organic compounds present in the dynamic river environment, all of the available samples from the river sites were analyzed. The monthly river samples from March through September are expected to represent conditions before, during and after the spring freshet and thereby capture varying proportions of snowmelt, local catchment runoff and baseflow groundwater inputs over the open water season.

2.2. ESI-FTICR MS analysis

Surface water samples and melted snow were pretreated with liquid–liquid extraction (LLE), followed by analysis on ultrahigh resolution FTICR MS (Gibson et al., 2011). LLEs were performed using EA-DCM (10:2,v/v). 3-mL of water sample was adjusted to pH 2 with formic acid, and extracted with 1 mL of EA-DCM. After a 30-s vortex mixing and subsequent clarification by centrifugation, 0.5-mL of the top organic layer was carefully transferred and dried under nitrogen. The dried

residue of LLE extracts was then reconstituted in 0.5 mL of a mixed acetonitrile–water–25% ammonium hydroxide (60:39.2:0.8, v/v) solution containing 2 μ L of the “ES Tuning Mix” (used for internal mass calibration) solution to be ready for direct infusion. ~18 μ L of prepared sample solution was finally directly infused into a 12-Tesla Apex-Qe hybrid quadrupole FTICR MS (Bruker Daltonics, Billerica, USA) at a flow rate of 3 μ L/min. The instrument was operated in ESI (–) mode with a detection range of mass-to-charge ratio (m/z) of 150–1100. The ESI (–) mode for ionization is used for detection and analysis of the acidic components of dissolved organic compounds (e.g., naphthenic acids, NAs) in complex environmental samples (Headley et al., 2013; Wang and Kasperski, 2010; Barrow et al., 2010). An example of the ESI-FTICR mass spectrum acquired in sample analysis is shown in Supplementary information Figure S1. We focused our interpretation on the ESI (–) results, because of the larger datasets available in the literature for this mode that can be used for discussion. The mass spectra were recorded with an accumulation of 200 scans in broadband acquisition and a data acquisition size of 1 MB per second.

Table 1
Summary of number of detected compounds and homologues in snow, river and lake samples.

Site code	Site description	Mar-2012		Jun-2012		Jul-2012		Aug-2012		Sep-2012	
		All peaks	Homo	All peaks	Homo	All peaks	Homo	All peaks	Homo	All peaks	Homo
<i>Snow</i>											
NW-S8	Far-field snow (61 km from GC)	1594	627								
NW-S2	Far-field snow (38 km from GC)	1941	717								
E-S7	Far-field snow (30 km from GC)	1678	625								
Kar-6a	Near-field snow (4 km from GC)	2074	734								
KST-2	Near-field snow (2 km from GC)	2318	816								
Kar-4	Near-field snow (2 km from GC)	2016	714								
Kar-16	Near-field snow (4 km from GC)	2074	737								
<i>Tributaries and river</i>											
Fir-1	Firebag River near mouth			2864	889	2949	829	3187	987	3358	952
Mus-1	Muskeg creek at confluence	3026	840	3342	976	3197	877			3479	948
Mus-2	Muskeg River above Muskeg Creek	2660	789	3103	873	2703	777	3503	1007	3350	947
Atha-1	Athabasca River u/s Firebags			2831	927	2725	801	3056	949	3302	992
Mus-3	Muskeg River at Gauge	2933	818	3358	959	3302	862	3420	969		
Jac	Jackpine Creek above Muskeg River	2780	785	3334	963	3345	856			3426	936
Mus-4	Muskeg River u/s Jackpine Creek	3028	855	3441	1015	3275	870			3463	931
Mus-5	Muskeg River u/s Stanley Creek	2771	821	3505	987	3351	879			3412	957
Mus-6	Muskeg River 11 km u/s Stanley Creek	2713	784	3353	1003	3277	864			3437	976
Atha-2	Athabasca river u/s Ft. McMurray			2728	910	2782	815	2872	934	2864	929
Atha-3	Athabasca River at Athabasca			2921	952	2168	713	2790	917	2569	886
Mac	MacKay River at HWY 63			3399	1009	3287	848	3412	938	3387	959
Ells-1	Ells River at Mouth			3115	962	3007	841	3189	953	3036	954
Ste-1	Steepbank River at Mouth			3327	1017	3344	868	3500	979	3285	941
Mus-7	Muskeg river at Jackpine Creek			3394	993	3225	860			3218	908
Mus-8	Muskeg River 27.5 km u/s Stanley Creek			3339	991	3281	896			3223	934
Cle-1	Clearwater River 2 km u/s water way			3181	980	3039	858	3203	965	3203	965
Fir-2	Firebag River near Winter Road	2393	762								
Ste-2	Steepbank River 5 km u/s Athabasca River	2763	797								
Cle-2	Clearwater River 2 km d/s Christina River	2512	806								
Ells-2	Ells River at Ft. Mckay Pumphouse	2883	820								
Dev	Athabasca River at Devil's Elbow					2969	865				
Old Fort	Athabasca River at Old Fort									2746	852
<i>LAKES</i>											
NE1	Lake							3206	818		
NE5	Lake							3286	884		
NE9	Lake							2880	806		
NE11	Lake							3131	851		
BM2	Lake							2652	789		
WF8	Lake							2897	819		

2.3. Data processing and statistical analysis

The data acquired from ESI-FTICR MS include response intensities for thousands of polar compounds identified by their mass to charge ratio (m/z). Mass spectra were batch-processed using a custom VBA script, which has been described within the instrument vendor's data processing software suite, DataAnalysis® (Han et al., 2008; Nyakas et al., 2013). Basically, the raw mass spectra acquired in the negative ion mode (ESI⁻) were internally calibrated with the reference masses of the standard "ES Tuning Mix" solution. The selection of the reference masses covered the mass range of data acquisition to make sure all the ions on the spectra were correctly calibrated. Monoisotopic peaks corresponding to the isotopic distribution patterns were determined and those with signal-to-noise ratios ≥ 10 were picked to generate two-dimensional (m/z versus intensity) data matrices. The spectra obtained from the field blanks were also processed, and the peaks detected in the field blanks were subtracted from the data matrices regardless of peak intensity. The resulting individual mass lists were further processed with another custom software program (LabView®, National Instruments, Austin, TX, <http://www.ni.com/labview/>) for peak alignment, which generates multi-sample datasets combining allowable masses within 2 ppm. Details of the customized software are described in Han et al., 2008. The final dataset consisted of the molecular mass and the corresponding intensities for all of the detected compounds.

The results from ESI-FTICR MS analyses are challenging to visually present and compare because of the large number of dissolved polar compounds detected in each sample. Kendrick Mass Defect (KMD) plots have been used to graphically illustrate the distribution of the thousands of compounds detected in individual samples (Marshall and Rodgers, 2004, 2008). KMD plots also have the benefit of making homologous compounds easy to identify. Homologues are characterized by the same number of heteroatoms (nitrogen, oxygen and sulfur), rings and double bonds, but differ from each other by repeating CH_2 units. Homologues have the same KMD value, and can be readily identified in KMD plot as points that plot along a horizontal line. Identification of homologues is important as it improves our capability to infer molecular formulae for the high mass compounds (Hughey et al., 2001; Marshall and Rodgers, 2004, 2008). The relative contributions (RC), which are calculated as the intensity of an individual peak divided by the total intensity of all detected peaks, were used to characterize the composition. It is important to note that the RC values cannot be treated as accurate measures of absolute concentration. This is because different compounds may have different ionization efficiencies in such complex mixtures (Barrow et al., 2010). However, RC generally reflects composition changes in a complex mixture (Hughey et al., 2007). The RC values are color coded in the KMD plots to profile the composition of the organic compounds; and they are also used in principle component analysis (PCA) to statistically compare compositional differences among samples.

PCA is used as an exploratory tool to examine and illustrate similarities and differences in the homologues present in the snow and surface water samples. PCA is a multivariate statistical technique that transforms and extracts meaningful information from large datasets such as profiling results from FTICR MS (Hur et al., 2010; Sleighter et al., 2010). PCA identifies a few linear combination of variables that are capable of explaining the overall variance in the dataset. As a result, the first few principal components can be used to represent the dataset without losing significant information, which substantially reduce the data dimensions and helps evaluate the similarity and differences among samples. In this study, we present score plots of the first principal component (PC1) vs the second principal component (PC2) to provide a statistical overview. In a score plot the closer two samples plot, the more similar are their dissolved organic compositions. Samples which are far apart in the score plot, are more likely to be different in compositions. The origin in the score plot represents the average of the entire dataset. The closer to the origin in the PCA score plot, the less distinctive is the sample compared to the average composition of the dataset. PCA was carried out using SIMCA-P+ (V12.0, Umetrics AB Umeå, Sweden).

The ESI-FTICR MS results from individual samples are also compared by grouping detected homologues into the major compound classes based on their heteroatoms. Molecular formulae of the detected homologues were computed based upon the accurately measured monoisotopic masses with a custom algorithm written in the Matlab R2011 (MathWorks Inc. Natick, MA). Heuristic filtering based on seven golden rules was applied in the algorithm to infer rationale molecular formulae along with identification of homologue series (Kind and Fiehn, 2007; Koch et al., 2007). For instance, the H/C and O/C ratios were limited to 0.2 to 3.1 and 0 to 1.2 respectively, while the N/C and S/C ratios were restricted in the range of 0 to 1.3 and 0 to 0.8 (Kind and Fiehn, 2007). Homologues were categorized into sixty major heteroatom classes based upon their elemental compositions, including O_n , NO_n , N_2O_n , SO_n , and S_2O_n (where $n = 1$ to 12 in all cases), consistent with the classification scheme used by Headley et al. (2011). By grouping thousands of compounds into classes and summarizing total relative contributions (RC) for each class, the profiling results are transformed into a matrix of compound classes and total relative contributions. The chemical compositions of organic profiling were then compared via bar plots. The procedure for data presentation, compositional sorting, and graphical imaging was programmed and standardized using Matlab.

3. Results and discussion

3.1. Detected compounds

Thousands of polar compounds are detected via ESI-FTICR MS analyses. The number of detected compounds and homologues is summarized in Table 1. The 7 snow samples (collected in March 2012) included 3 far-field site samples and 4 near-field site samples. As a group, the snow samples had the lowest number of peaks (average 1,956) with slightly more peaks detected in the near-field sites than in the far-field sites (Table 1). On average, 710 homologues were detected in snow samples with an overall range of 625 to 816. Just over 30% of detected peaks in the snow samples are homologues.

The total number of detected polar organic compounds in the Athabasca River and tributary samples ranged between 2393 and 3028 in March 2012 (~2700 compounds on average), but increased to between 2569 and 3479 in September 2012 (~3200 compounds on average, Table 1). A general trend of increasing numbers of dissolved compounds at individual sampling sites from March to September is evident, with some exceptions in June and July when the numbers of peaks were most variable from station to station. The March river samples were taken under-ice, before the spring freshet when streamflow is sustained by baseflow. The samples taken later in the season would include baseflow as well as catchment runoff. The increase in the number of

peaks likely reflects the more diverse sources of organic compounds in open water conditions. Approximately 30% of detected organic compounds were identified as homologues. The number of homologues appears closely correlated with the total number of compounds detected. On average, 904 homologues were detected in the Athabasca River and tributary samples, ranging between 713 and 1,017. It is also interesting to note that tributary samples generally have more detected compounds (and homologues) than samples from the main stem of the Athabasca River.

The ESI-FTICR MS results for the 6 lakes sampled in August 2012 (Table 1) had an average of 3009 compounds which was similar to the number detected in rivers but more than in the snow. More homologues were identified in lake water (828) compared to snow (710). Overall, the percentage of homologues in the lake samples is comparable to the percentage of homologues in river and lake water (~30%).

3.2. Statistical analysis

PCA was used to provide a statistical overview of the entire dataset (including near- and far- field snow, Athabasca River, tributaries, and lakes). The analysis illustrates differences in the polar organic composition of different water types in the AOSR (Fig. 2). The PC1 explains 52.1% of total variation of the dataset while the PC2 explains 10.6% of the variance. The result shows a clear separation between snow and surface water samples in the score plot, as well as clustering within these categories associated with geographic locations.

As illustrated in Fig. 2, surface water samples plot in the area near the origin with variability along the PC2 whereas the snow samples are compositionally distinct along both the PC1 and PC2 axes, indicating compositional differences between the snow and surface waters. The near-field (Kar-6a, Kar-16, Kar-4 and KST-2) and far-field (NW-S8, E-S7 and NW-S2) snow samples plot in two separate clusters based upon the PC2 (Fig. 2). Although only 7 snow samples were included in this study, this geographical clustering of far- and near-field samples is distinct suggesting that dissolved organic compounds in atmospheric deposition within the vicinity of major oil sands activities differ compositionally from background conditions represented by far-field samples.

Among surface water samples (i.e., rivers and lakes) there are also interesting patterns related to sampling location (Fig. 2). The Athabasca River samples are distinct from the tributary samples along the PC1, and show a systematic distribution from upstream to downstream sites. Athabasca River water collected from the upstream site (i.e., Athabasca River at Athabasca; Atha-3) are more similar to far-field snow samples, but become progressively more similar to tributary and lake samples as the river flows through the development area (i.e., Athabasca River at Ft. McMurray; Atha-2) to more distal downstream sites (i.e., Athabasca River at Firebag River; Atha-1) (Fig. 2). The observation that the upstream Athabasca river samples trend towards the composition of far-field snow suggests that upstream of Fort McMurray the dissolved organic compounds present in the Athabasca river are more likely influenced by catchment runoff at a regional scale. As the Athabasca River flows downstream across the AOSR, the composition of dissolved organic compounds become more similar to the average composition of the entire dataset, and suggesting similar sources as those in local lakes and tributaries. There is no evidence of the dissolved organic compounds in the Athabasca River becoming more similar to the near-field snow samples as the river flows through the development area.

The dissolved organic compounds in lake waters are similar to those in tributary samples with broad range along the PC2. In Fig. 2, both lake and tributaries generally plot close to the origin point, which suggest that the dissolved organic compositions in the lakes and tributaries are not dramatically different from the average organic composition in the entire dataset. The compositional similarity between lakes and tributaries suggests they inherited their organic signature from a common water source or runoff processes. The headwater lakes sampled for this study tend to accumulate catchment runoff over the open water

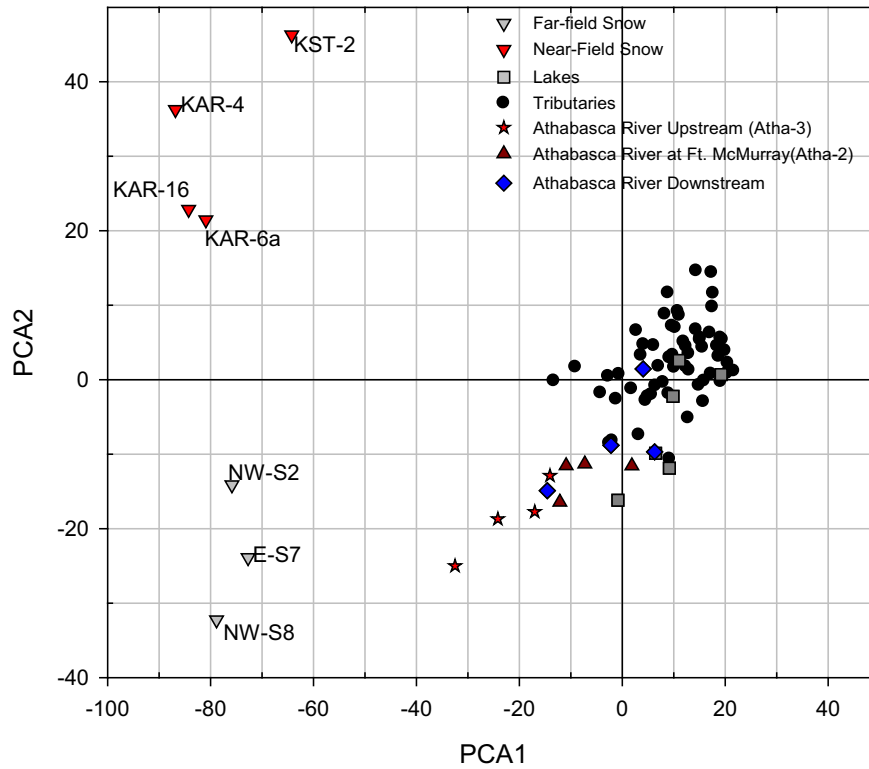


Fig. 2. PCA analysis of homologous composition from the entire dataset (including Athabasca River, tributaries, lakes and snow samples). The Athabasca River downstream samples include samples from the Athabasca River at the confluence with the Firebag River (Atha-1), Devils Elbow and Old Fort.

season, whereas tributaries reflect a more dynamic temporal signal, with the spring freshet resulting in an annual flushing of the catchment. Despite these different time-scales for runoff processes in the shallow

lakes and tributaries included in this study, the similarity in their organic signatures are consistent with having originated from a common source, such as shallow catchment runoff.

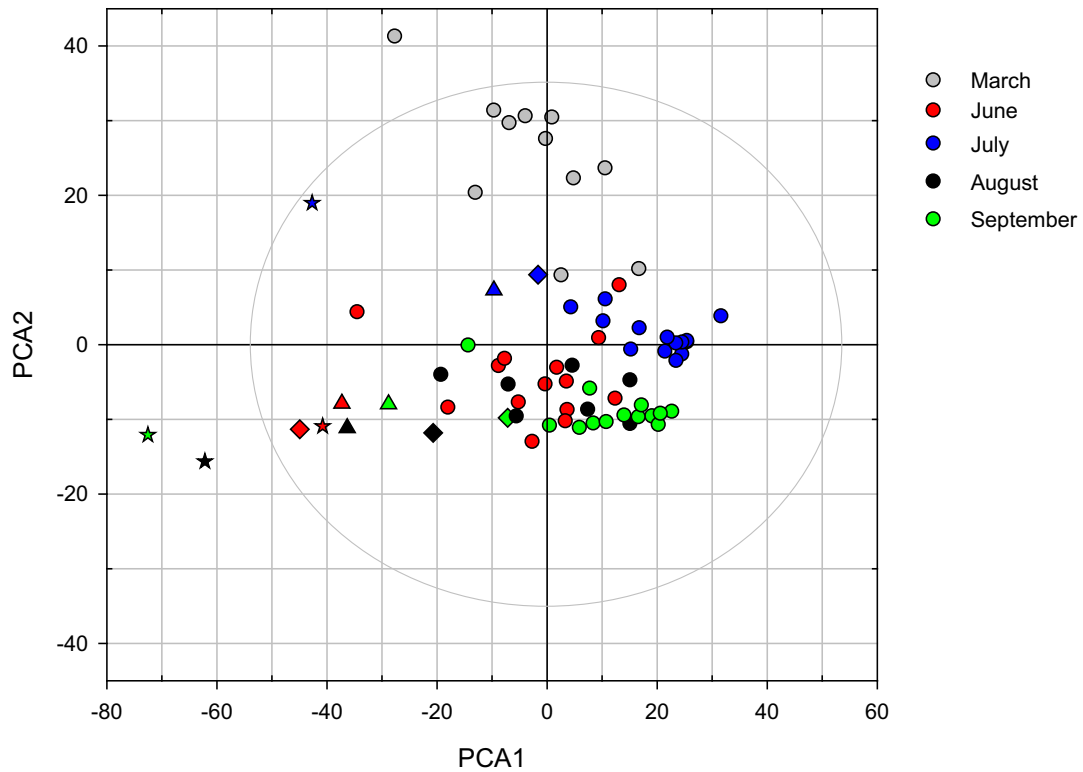


Fig. 3. PCA analysis of homologous composition in river samples only (including the Athabasca River and its tributaries). The symbol colors indicate the sampling month (shown in legend). Note, different symbols (which is consistent with symbols applied in Fig. 2, ★ Athabasca River upstream (Atha-3), ▲ Athabasca at Ft. McMurray (Atha-2) and ◆ Athabasca River downstream) are used to distinguish samples from Athabasca River main stem.

The PCA of all of the monthly river data show temporal variations in the organic compounds (Fig. 3). The first axis (PC1) explains 37.1% of variability, while the second axis (PC2) explains 13.4% of variability. The most obvious temporal distinction in the dataset is that samples from March 2012 are separated from other samples along the PC2 (Fig. 3). Closer examination shows that within the March data, there is significant spatial variation depending on sampling site (Supplementary information, Figure S2). The composition of dissolved organic compounds in July 2012 also appears different from other months. This indicates that the organic compositions in river baseflow during the winter, when the discharge is mainly sustained by groundwater, are different from organic compositions in the spring and summer, when the river is more hydrologically active due to snowmelt inputs and other catchment runoff contributions (Gibson and Prowse, 2002; St Amour et al., 2005). It is also interesting to note that March is also the month when fewer number of compounds and homologues are detected in river samples, compared to samples from the same location but different months (Table 1). Only tributary samples were available in March, so it is unknown if the organic composition of the Athabasca River is similar to the tributary compositions at that time or not. However, the distinct composition of dissolved organic compounds in March is evident in all of the tributary samples (Firebag, Muskeg, Jackpine, Ells, and Steepbank Rivers). Previous characterization of the organic compounds present in rivers in the region have been more far spatially and temporally limited (e.g., Grewer et al., 2010, 5 rivers, each sampled once at one sampling sites; Headley et al., 2011, 1 river, sampled once at 5 sampling sites). Even though the dataset used here would have benefitted from more samples from the under-ice period, particularly for the Athabasca River, it includes a more comprehensive characterization of dissolved organic compounds in rivers (9 rivers, 23 sampling sites, each with 1 to 5 sampling events) than has previously been available, and has provided new insights into the temporal and spatial variability in surface waters in the region.

More detailed illustration of the temporal and spatial variation in the dissolved organic composition of river water is presented in the Supplementary information Figure S2 by separating the data presented in Fig. 3 into individual months so that the spatial and temporal variations are discernible. The distinct labelling of dissolved organic compounds from upstream to downstream reaches of the Athabasca is still evident, particularly during July, August and September (Supplementary information Figure S2 d–f). During the spring freshet (i.e., June, Supplementary information Figure S2c), the differences in organic compounds between the different sampling locations are reduced, which may be due to the overwhelming effect of snowmelt contributing a common organic signal across all of the sampling locations. The temporal variations in the organic compounds detected at different sampling locations indicate large temporal variations in the sources of organic compounds in rivers in the AOSR.

3.3. Organic profiles

KMD plots are used to present all homologues detected in an individual sample and visually illustrate the compositional information by highlighting the relative abundance of compounds (as indicated by colors). Of the 86 KMD plots produced, six are included in Fig. 4 to provide a view of compositional differences that represent the groupings and clustering in Figs. 2 and 3. The representative samples selected include a far-field snow sample (NW-S8, Fig. 4a), a near-field snow sample (KST-2, Fig. 4b), a tributary sample (Muskeg River at Gauge, August 2012, Fig. 4c), a lake water sample (NE1, August 2012, Fig. 4d), a downstream Athabasca River sample (Athabasca River at Confluence with Firebag, August 2012, Fig. 4e), and an upstream Athabasca River sample (Athabasca River at Athabasca, August 2012, Fig. 4f)

The KMD plots for the far- (Fig. 4a) and near-field (Fig. 4b) snow samples show clear differences from surface waters (Fig. 4c–f). Relatively abundant compounds, as indicated visually by dark-red colors,

appeared irregularly in the far-field snow (Fig. 4a), unlike the clear bulls-eye patterns illustrated in lake and river samples (Fig. 4c–f). Meanwhile, the bulls-eye pattern in the near-field snow (Fig. 4b) is different in shape from the pattern in lake and river samples (Fig. 4c–f). Most of the relatively abundant species in snow (Fig. 4a and b) are within a mass range of 200 to 400 Da (x-axis), compared to a broader mass range of 200 to 600 Da in surface water samples (Fig. 4c–f). The dominant species in snow are also characterized by KMD values of usually <200 (y-axis), suggesting less hydrogen content in the molecular structure.

The Muskeg River (Fig. 4c), lake NE-1 (Fig. 4d) and the Athabasca River (including downstream and upstream sites, Fig. 4e & f) share many similarities in mass distribution of homologues. Most of the organic compounds detected in these samples are in the mass ranges between 200 and 600 Da. The most relatively abundant compounds in these samples form a bulls-eye pattern constrained by the KMD values between 150 and 300. As the molecular mass increases from 200 to 400 Da, the KMD values of these relative abundant compounds also tend to increase from 150 to 300, suggesting the increase in molecular size is likely due to increase in carbon chain (or rings) rather than increase in heteroatoms (especially oxygen and sulfur). It is very interesting to note that the bulls-eye pattern extends beyond KMD values of 300 in the tributary and lake samples (Fig. 4c & d) including compounds with molecular mass of 400 to 500 Da. This extension appears to be weak in Athabasca River downstream site (Fig. 4e), but absent in the Athabasca River upstream site (Fig. 4f). The broader bulls-eye pattern in the tributaries and lakes than in Athabasca River samples is consistent with more peaks in tributaries and lakes, when samples collected in August were compared (Table 1).

Visual comparison of the KMD plots can be used to identify blocks of compounds that may be responsible for the clustering of sample types in the PCA (Fig. 2). For example, the organic compounds with a molecular mass range of 300 to 500 Da and with KMD values between 50 and 150 are present in snow samples (Fig. 4a & b) with relatively high RC (indicated by dark red color); but they are only present with relatively low abundance (indicated by the blue color in the KMD plots) in the Athabasca River samples (Fig. 4e & f); and are generally absent in tributary and lake samples (Fig. 4c & d). Lake and tributary samples differ from snow and the Athabasca River samples, because of the absence of compounds in this block of homologues. Organic compounds with mass range of 300 to 400 Da and KMD values between 300 and 400, may be distinctive in the overall differences between snow and surface waters. These compounds are generally absent in far-field snow (Fig. 4a); show very limited appearance in Athabasca River upstream samples (Fig. 4f), but are present with higher RCs in lakes, tributaries and the downstream Athabasca samples (Fig. 4c, d & e). Overall, the visual comparison of KMD plots from individual samples appears to be consistent with the PCA. The distinction between snow (Fig. 4a&b) and surface water (Fig. 4c–f) is apparent. Lake waters (Fig. 4d) are similar to tributaries (4c), but there are subtle differences in the organic compounds present in the main stem of the Athabasca River (Fig. 4e & f), that differentiate this water from other surface water samples. These general observations on the major differences in organic compounds present in near- and far-field snow samples, rivers and lakes, can be used to guide future investigations that focus on identifying the individual compounds that are present or absent in the different types of waters.

3.4. Chemical compositions by compound class

The results of the organic profiling can also be examined by classifying the polar compounds based upon inferred molecular formula and presenting the relative contribution (RC) of each compound class. Compound classes are categorized based on a classification scheme similar to Headley et al. (2011). Using the same representative samples selected for Fig. 4, chemical compositions provide an additional perspective on the differences in the snow and surface water samples. Fig. 5 presents the results from the compound classes detected in most samples,

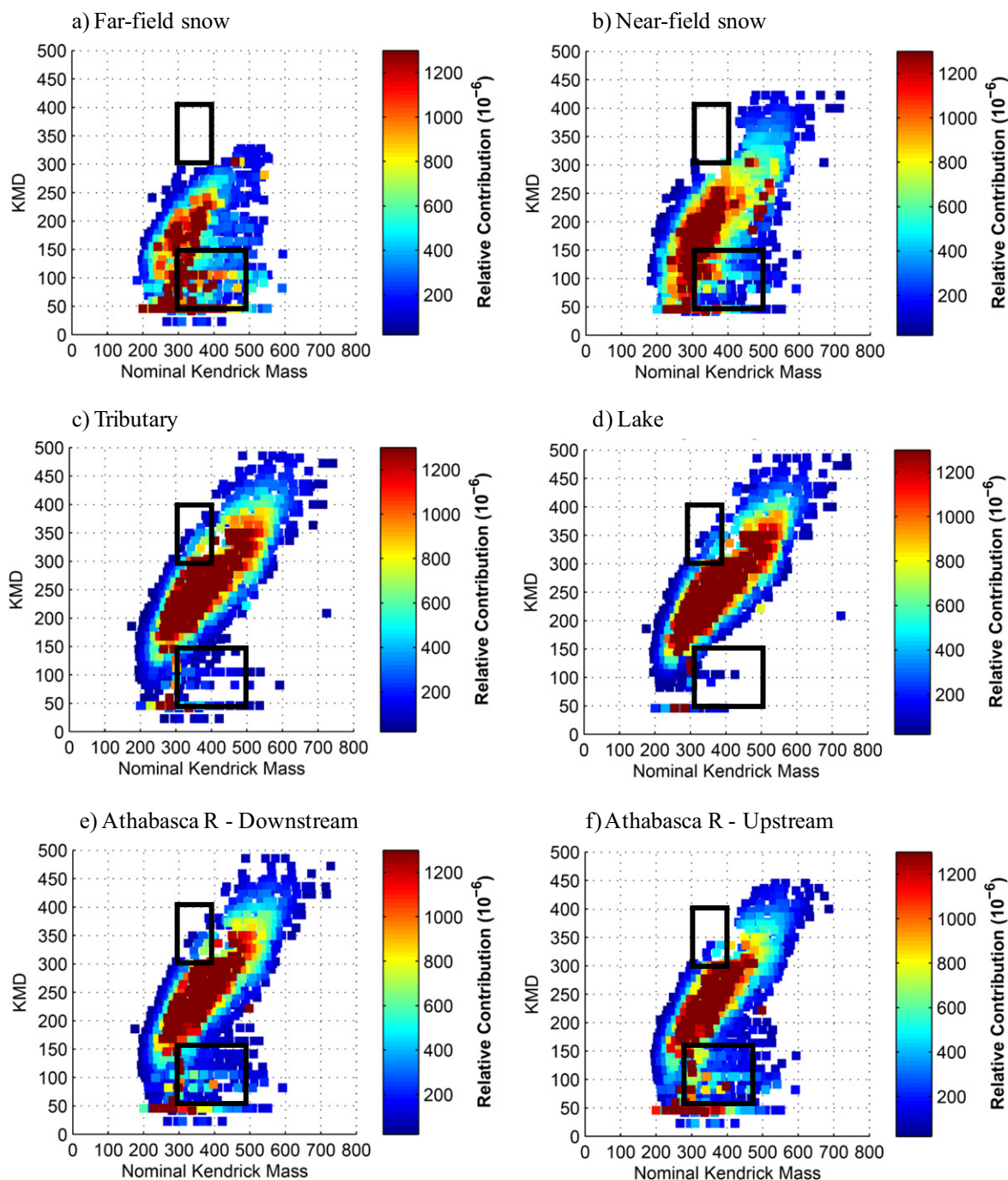


Fig. 4. Representative Kendrick plots of the dissolved organic compounds (homologues) detected via ESI-FTICR MS (negative modes): a) far-field snow (NW-S8); b) near-field snow (KST-2); c) tributary (Muskeg river at Gauge, Mus-3); d) lake (NE-1); e) Athabasca River downstream site (Athabasca River at Confluence with Firebag River, Atha-1); f) Athabasca River upstream site (Athabasca River at Athabasca, Atha-3). August samples from Athabasca River and its tributaries were chosen; lakes were sampled in August; snow samples were collected in March. Boxes highlight the block of compounds with specific KMD and nominal mass ranges discussed in the text.

but the full spectrum of speciation and their corresponding RC are included in the Supplementary Information Figure S3.

The main difference between the snow (Fig. 5a–b) and the surface water samples (Fig. 5c–f) is the dominance of O_n ($n < 4$) compound classes, particularly the O_2 compound classes in snow samples. For compound classes O_n ($n \geq 4$), a pattern of progressive decrease in RC is apparent in snow samples. In contrast, surface waters typically demonstrate a bell-shaped distribution in the O_4 to O_{12} classes with RC peaks in O_6 to O_8 compound classes. There are large differences in the relative contribution of O_2 compounds between the Athabasca River samples and tributaries and lakes. In the Athabasca River sample,

the relative contribution of O_2 compounds (Fig. 5e & f) are comparable (or greater) to the peaks of O_4 to O_{12} , which suggest some degree of similarities to snow samples (Fig. 5a & b). In lakes and tributaries, O_2 compounds are considerably less in RC.

The O_2 class is commonly regarded as “classic” naphthenic acids. The term “oxy-naphthenic acids”, describing compounds formed after mild oxidation of the classic naphthenic acids, were introduced by Lee (1940). Previous studies have also reported finding mono- and di-oxide naphthenic acids (i.e., O_3 and O_4 classes), as well as O_n compounds with more oxygen atoms, in extracts from OSPW (Batatineh et al., 2006; Han et al., 2009; Grewer et al., 2010). The profiling results in this study

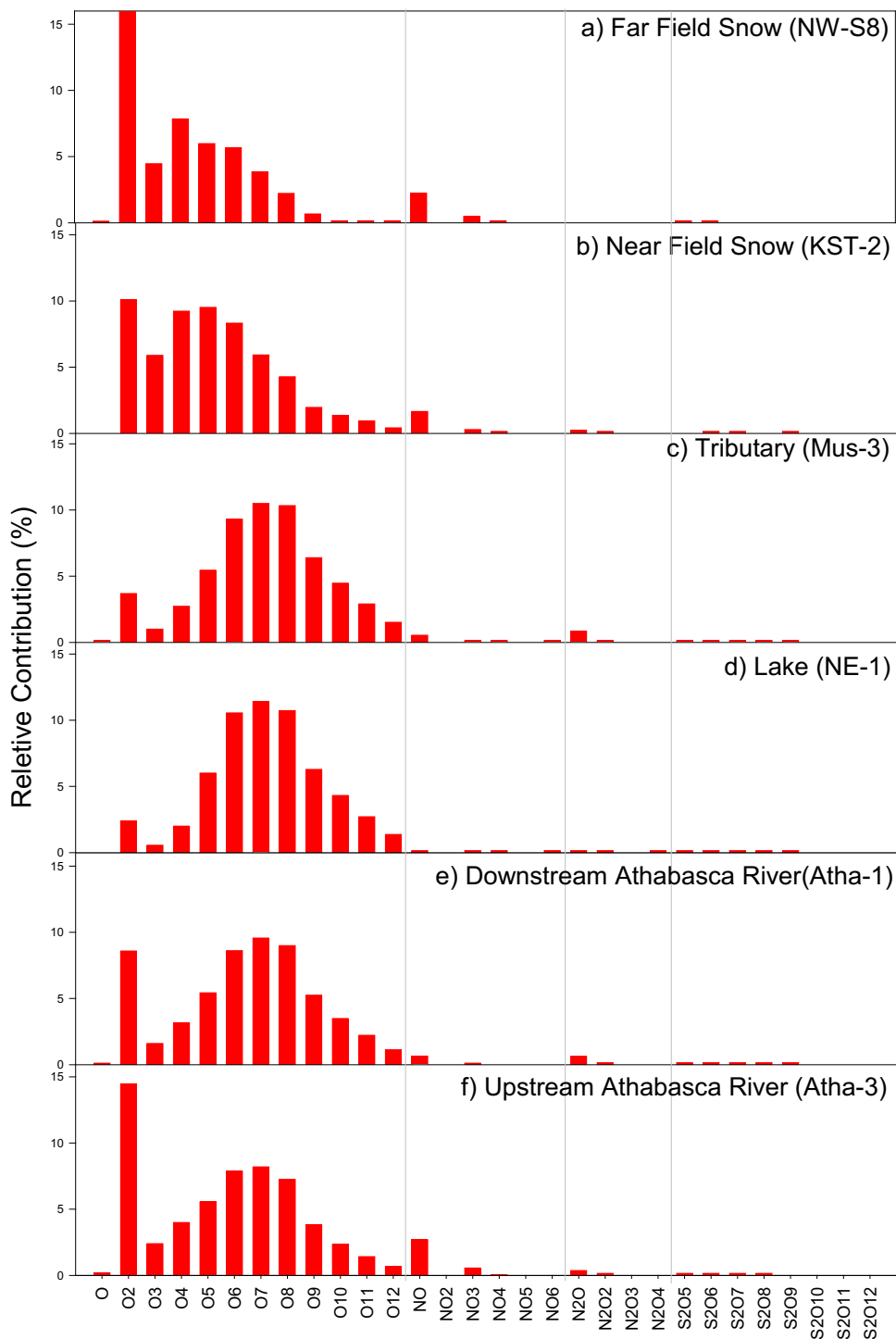


Fig. 5. Relative contribution (%) of the compound classes detected for representative samples: a) far-field snow (NW-S8); b) near-field snow (KST-2); c) tributary (Muskeg river at Gauge, Mus-3); d) lake (NE-1); e) Athabasca River downstream site (Athabasca River at Confluence with Firebag River, Atha-1); f) Athabasca River upstream site (Athabasca River at Athabasca, Atha-3). August samples from Athabasca River and its tributaries were chosen; lakes were sampled in August; snow samples were collected in March.

reveal a broad spectrum of oxy-naphthenic acids (e.g., O_{3-12}), that warrant closer examination, particularly in their contribution relative to classic naphthenic acids (O_2).

Frank et al. (2014) proposed the use of $O_2:O_4$ ratios as a potential indicator of oil sands processed affected water (OSPW). They suggested that the high $O_2:O_4$ ratios (1.21 to 1.69) in OSPW samples, compared to those found in far-field groundwater samples (0.40–0.57) could be used in combination with other parameters as indicators of impact from OSPW. The $O_2:O_4$ for the 7 snow samples analyzed in this study

were between 1.09 and 1.68 in near-field snow, and between 2.26 to 3.48 in far-field snow, both of which are within or higher than the range of OSPW identified by Frank et al. (2014). The range of $O_2:O_4$ for snow reported here provides additional support that $O_2:O_4$ ratios may not be a good indicator of OSPW impacts. As has been pointed out in previous compilations of $O_2:O_4$ ratios from various environmental samples in the AOSR (Yi et al., 2014, also shown in Supplementary Information Figure S4), there are wide ranges in the $O_2:O_4$ in different natural waters, and OSPW. Some of the variations in reported $O_2:O_4$ values

are likely related to different profiling and pretreatment techniques (Yi et al., 2014), that highlight the need for standardized analytical protocol. The range of $O_2:O_4$ ratios for surface water in this study are between 0.81 and 5.12 in the Athabasca River samples and between 0.76 and 2.60 in tributary samples, which is consistent with previous results for rivers in the region (Grewer et al., 2010; Yi et al., 2014). Clearly, the $O_2:O_4$ ratios are not a simple indicator of impact of OSPW. The overlap in the $O_2:O_4$ ratios for both near and far field snow samples with those reported by Frank et al. (2014) suggest that shifts towards higher $O_2:O_4$ ratios in shallow groundwater could also be attributed to increased snowmelt recharge, or greater interaction with surface waters (Supplementary information Figure S4). Increased efforts to develop comprehensive datasets of organic profiling of natural waters and anthropogenic sources of organic compounds in the region are needed to better understand the labelling and transport of organic compounds, so that suitable indicators of oil sands derived organic compounds (e.g., atmospherically derived organic compounds and OSPW) can be identified.

In addition to O_n compounds, there are organic nitrogen- and sulfur-containing compounds (e.g., NO_n , N_2O_n and S_2O_n) detected by ESI-FTICR MS in snow and surface water samples from the AOSR that may be useful for understanding the distribution and transport of organic compounds. Even though the RC values for these individual compound classes are usually less than 1%, they were not present in all samples and may be useful for distinguishing different sources of organic compounds. For instance, S_2O_n ($n = 5-9$) and N_2O_n ($n = 1-4$) compounds are ubiquitous in surface water (Fig. 5c-f), but these were absent or present in very limited forms in far- (Fig. 5a) near-field snow (Fig. 5b). Far-field snow samples (Fig. 5a) only detected S_2O_n ($n = 5-6$) compounds but no N_2O_n ($n = 1-4$) compounds. In contrast, N_2O and N_2O_2 compounds are identified in near-field snow (Fig. 5b), while the appearance of S_2O_n compounds is limited to S_2O_n (6-7 and 9). A detailed listing of the S_2O_n ($n = 5-9$) and N_2O_n ($n = 1-4$) compounds identified in this study reveal lots of these compounds are in the molecular mass of 300 to 400 Da with KMD values between 300 and 400 Da, which fall within one of the aforementioned blocks of compounds suggested in the visual comparison of KMD plots (Fig. 4). Differences in N_2O_n or S_2O_n compounds in near- and far-field snow may have potential to differentiate between compounds present in snow near development, from those present in background snow.

Differences in the relative contribution of NO compound classes may account for some of the separation noted in the PCA (Fig. 2). Snow samples (Fig. 5a & b) contain higher relative contributions of NO compounds than the tributary, lake or downstream Athabasca River samples (Fig. 5c, d, e). The upstream Athabasca river (Fig. 5e) sample is the only surface water type with appreciable NO contributions, and this likely contributes to the similarity noted between upstream Athabasca River samples and snow (far-field in both PC1 and PC2; near-field in PC1) noted in Fig. 2. The molecular formulae inferred for the NO compound classes detected in this study range from ~150 Da (containing 9 carbon) to >450 Da (containing 30 carbons) in molecular weight. More importantly, these compounds are all characterized by relatively low KMD values (50 to 150), which is again consistent with our observations of blocks of compounds in the visual comparison of the KMD plots (Fig. 4).

The distribution of O_2 compounds in the upstream Athabasca River sample is also similar to the distribution observed in far-field snow, with significantly greater RC than other O_n compounds. According to PCA analysis (Fig. 2), the upstream Athabasca River sample is more similar to the cluster of far-field snow samples than the other surface water samples. It is not surprising that the dissolved organic compounds present in the upstream Athabasca River may have inherited part of their signature from atmospheric organics deposited in the region (i.e., represented by far-field snow). As a result, NO (characterized by low KMD values) and O_2 compounds would be the target compounds to explore the linkage between far-field snow and Athabasca River.

4. Summary and conclusion

This study has provided an improved characterization of dissolved organic compounds present in surface waters and snow in the AOSR. FTICR MS identified thousands of dissolved polar compounds in ESI (-) mode and our results suggest that the composition of thousands of organic compounds may have value in identifying different organic sources and their pathways to aquatic ecosystems.

There were significant compositional differences between the dissolved polar organic compounds present in snow and surface waters in the AOSR, suggesting that direct transfer of the dissolved organic compounds present in snowmelt to the Athabasca River, its tributaries and lakes is not likely the dominant source of organic compounds to these surface water bodies. Samples from the Athabasca River appear to have distinct compositions from near-field snow samples (representing atmospheric deposition from oil sands development). There is no evidence of the dissolved organic compounds in the Athabasca River becoming more similar to the near-field snow samples as the river flows through the development area. The compositional differences between snow, river and lake were centred around the relative contributions of oxygen containing compounds (i.e., O_n classes) as well as nitrogen (i.e., NO_n and N_2O_n classes) and sulfur containing (i.e., S_2O_n) compounds. Differences in O_n classes were not limited to O_2 (classic naphthenic acids) compounds, whereas compounds containing nitrogen and sulfur warrant more in-depth investigations to explore their potential usage in tracing and labelling source and pathway of organic compounds in the region.

The results of the ESI-FTICR MS analyses also identified compositional differences within the group of snow samples. The near-field samples differed from the far-field snow samples indicating spatial variations in the dissolved organic compounds in atmospheric deposition in the broad region. Distinctions between far-field and near-field snow also suggest that the dissolved organic composition in atmospheric deposition may carry the signal of industrial development and could present unique opportunities for labelling and tracing sources and pathways of organic compounds.

The PCA results show that samples from the upstream Athabasca River site are similar to far-field snow samples, consistent with the influence of regional atmospheric deposition via snowmelt runoff in the Athabasca river basin. As the Athabasca River flows through the oil sands development area, samples from downstream Athabasca River sites become progressively more similar to local tributaries and lakes, indicating increasing relative contributions from local catchment runoff. We propose that organic profiles at the upstream sites are likely linked more closely to snow and other sources at a regional scale, while signatures at downstream sites (i.e., Athabasca River at Fort McMurray; Atha-2 and at Firebag River confluences; Atha-1) are more likely to reflect input from tributaries and local sources of dissolved organic compounds. The similarities in the organic compounds present in lakes and tributaries suggest that it could be difficult to label and differentiate lake water from tributary discharge via profiling organic compositions.

The focus of this study was to use ultrahigh resolution analytical techniques to profile the full range of dissolved organic compounds with the overall goal of characterizing the compositional differences in snow and surface water. The results have identified some groups of homologues and compound classes that distinguish various types of waters in the AOSR that can be used to help focus future work to identify potential indicators of organic sources. In particular, there are two blocks of compounds that appear to contribute to differences/similarities between snow and surface water in the region. The first group, that potentially distinguishes between near- and far-field snow, and surface water, includes compounds with mass ranges within 300 to 400 Da and KMD values between 300 and 400, many of which fall within N_2O_n or S_2O_n class. The other group, that contributes to the similarity between Athabasca river and snow (particularly far-field snow), is characterized by mass ranges of 300 to 500 Da and KMD values between 50 and 150, which includes NO compound class.

The results of this study have identified some intriguing organic signatures as well as spatial and temporal trends in the homologues detected in ESI (–) mode in snow and surface waters that could be used to better understand the sources and transport of organic compounds in the region. However, identifying which of these compounds could be key indicators of the sources in the region would be greatly improved with an expanded database that includes samples of the various potential natural and anthropogenic end-members, characterized using the same analytical techniques and laboratory procedures.

Acknowledgment

This research was funded by the Alberta Innovates Technology Futures and the Oil Sands Research and Information Network (OSRIN), University of Alberta. The authors are grateful to the staffs of Alberta Environment and Sustainable Resource Development (AESRD) including Brian Jackson, Shelley Manchur, Jessica Di Maria, Mike Bilyk, Chris Ware and Roderick Hazewinkel for their help with fieldwork and coordinating the field component. The authors also wish to thank Don Jones (Alberta Innovates Technology Futures; AITF) for assisting with sample handling and database management; Preston McEachern for suggestions on snow study design and providing valuable inputs; and Drs. Christoph H. Borchers and Jun Han (Genome BC Proteomics Centre – University of Victoria) for the sample analysis. This manuscript also benefitted from comments made by Dr. Kusumakar Sharma (AESRD) and Emily Taylor (AITF).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.scitotenv.2015.02.018>.

References

- AENV, 2006. Aquatic Ecosystems Field Sampling Protocols. AENV, Edmonton (AB) (Publication No: T/883. ISBN 0-7785-5079-6).
- Barrow, M.P., Witt, M., Headley, J.V., Peru, K.M., 2010. Athabasca Oil sands process water: characterization by atmospheric pressure photoionization and electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 82, 3727–3735.
- Bataineh, M., Scott, A.C., Fedorak, P.M., Martin, J.W., 2006. Capillary HPLC/QTOF-MS for characterizing complex naphthenic acids mixtures and their microbial transformation. *Anal. Chem.* 78, 8354–8361.
- Bennett, K.E., Gibson, J.J., McEachern, P., 2008. Water yield estimates for critical loadings assessment: comparisons of gauging methods vs. an isotopic approach. *Can. J. Fish. Aquat. Sci.* 65, 83–99.
- Cho, S., Sharma, K., Brassard, B., Hazewinkel, R., 2014. Polycyclic aromatic hydrocarbon deposition in the snowpack of the Athabasca oil sands region of Alberta, Canada. *Water Air Soil Pollut.* 225, 1910. <http://dx.doi.org/10.1007/s11270-014-1910-4>.
- Colavecchia, M.V., Backus, S.M., Hodson, P.V., Parrott, J.L., 2004. Toxicity of oil sands to early life stages of fathead minnows (*Pimephales promelas*). *Environ. Toxicol. Chem.* 23, 1709–1718.
- Frank, R.A., Roy, J.W., Bickerton, G., Rowland, S.J., Headley, J.V., Scarlett, A.G., West, C.E., Peru, K.M., Parrott, J.L., Conly, F.M., Hewitt, L.M., 2014. Profiling oil sands mixtures from industrial developments and natural groundwaters for source identification. *Environ. Sci. Technol.* 48 (5), 2660–2670.
- Gibson, J.J., Prowse, T.D., 2002. Stable isotopes in river ice: identifying primary over-winter streamflow signals and their hydrological significance. *Hydrol. Process.* 16, 873–890.
- Gibson, J.J., Birks, S.J., Moncur, M., Yi, Y., Tattrie, K., Jasechko, S., et al., 2011. Isotopic and geochemical tracers for fingerprinting process-affected waters in the oil sands industry: a pilot study. Oil Sands Research and Information Network. University of Alberta, Edmonton, Alberta, p. 109.
- Grewer, D.M., Young, R.F., Whittal, R.M., Fedorak, P.M., 2010. Naphthenic acids and other acid-extractables in water samples from Alberta: what is being measured? *Sci. Total Environ.* 408, 5997–6010.
- Hall, R.L., Wolfe, B.B., Wiklund, J.A., Edwards, T.W.D., Farwell, A.J., Dixon, D.G., 2012. Has Alberta oil sands development altered delivery of polycyclic aromatic compounds to the peace-Athabasca delta? *Plos One* 7.
- Han, J., Danell, R.M., Patel, J.R., Gumerov, D.R., Scarlett, C.O., Speir, J.P., et al., 2008. Towards high-throughput metabolomics using ultrahigh-field Fourier transform ion cyclotron resonance mass spectrometry. *Metabolomics* 4, 128–140.
- Han, X.M., MacKinnon, M.D., Martin, J.W., 2009. Estimating the in situ biodegradation of naphthenic acids in oil sands process waters by HPLC/HRMS. *Chemosphere* 76, 63–70.
- Headley, J.V., Barrow, M.P., Peru, K.M., Fahlan, B., Frank, R.A., Bickerton, G., et al., 2011. Preliminary fingerprinting of Athabasca oil sands polar organics in environmental samples using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.* 25, 1899–1909.
- Headley, J.V., Peru, K.M., Mohamed, M.H., Frank, R.A., Martin, J.W., Hazewinkel, R.R.O., et al., 2013. Chemical fingerprinting of naphthenic acids and oil sands process waters—a review of analytical methods for environmental samples. *J. Environ. Sci. Health A* 48, 1145–1163.
- Hughey, C.A., Hendrickson, C.L., Rodgers, R.P., Marshall, A.G., Qian, K.N., 2001. Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Anal. Chem.* 73, 4676–4681.
- Hughey, C.A., Galasso, S.A., Zumberge, J.E., 2007. Detailed compositional comparison of acidic NSO compounds in biodegraded reservoir and surface crude oils by negative ion electrospray Fourier transform ion cyclotron resonance mass spectrometry. *Fuel* 86, 758–768.
- Hur, M., Yeo, I., Park, E., Kim, Y.H., Yoo, J., Kim, E., No, M.H., Koh, J., Kim, S., 2010. Combination of statistical methods and Fourier transform ion cyclotron resonance mass spectrometry for more comprehensive, molecular-level interpretations of petroleum samples. *Anal. Chem.* 82 (1), 211–218.
- Jautzy, J., Ahad, J.M.E., Gobeil, C., Savard, M.M., 2013. Century-long source apportionment of PAHs in Athabasca oil sands region lakes using diagnostic ratios and compound-specific carbon isotope signatures. *Environ. Sci. Technol.* 47, 6155–6163.
- Kelly, E.N., Short, J.W., Schindler, D.W., Hodson, P.V., Ma, M.S., Kwan, A.K., et al., 2009. Oil sands development contributes polycyclic aromatic compounds to the Athabasca River and its tributaries. *Proc. Natl. Acad. Sci. U. S. A.* 106, 22346–22351.
- Kelly, E.N., Schindler, D.W., Hodson, P.V., Short, J.W., Radmanovich, R., Nielsen, C.C., 2010. Oil sands development contributes elements toxic at low concentrations to the Athabasca River and its tributaries. *Proc. Natl. Acad. Sci. U. S. A.* 107, 16178–16183.
- Kind, T., Fiehn, O., 2007. Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinform.* 8.
- Koch, B.P., Dittmar, T., Witt, M., Kattner, G., 2007. Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. *Anal. Chem.* 79, 1758–1763.
- Kurek, J., Kirk, J.L., Muir, D.C.G., Wang, X.W., Evans, M.S., Smol, J.P., 2013. Legacy of a half century of Athabasca oil sands development recorded by lake ecosystems. *Proc. Natl. Acad. Sci. U. S. A.* 110, 1761–1766.
- Lee, D.D., 1940. In: Office USP (Ed.), *Thermosetting Resin Reaction Product of Furfural With an Oxy-Naphthenic Acid*. Standard Oil Co., p. 4.
- Marshall, A.G., Rodgers, R.P., 2004. Petroleomics: the next grand challenge for chemical analysis. *Acc. Chem. Res.* 37, 53–59.
- Marshall, A.G., Rodgers, R.P., 2008. Petroleomics: chemistry of the underworld. *Proc. Natl. Acad. Sci. U. S. A.* 105, 18090–18095.
- Martin, F.L., Pearce, T.G., Hewer, A., Phillips, D.H., Semple, K.T., 2005. A biomarker model of sulthealth genotoxicity (DNA single-strand breaks and adducts) using the sentinel organism *Aporrectodea longa* in spiked soil. *Environ. Pollut.* 138, 307–315.
- Mazzoleni, L.R., Ehrmann, B.M., Shen, X.H., Marshall, A.G., Collett, J.L., 2010. Water-soluble atmospheric organic matter in fog: exact masses and chemical formula identification by ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry. *Environ. Sci. Technol.* 44, 3690–3697.
- Nyakas, A., Han, J., Peru, K.M., Headley, J.V., Borchers, C.H., 2013. Comprehensive analysis of oil sands processed water by direct-infusion Fourier-transform ion cyclotron resonance mass spectrometry with and without offline UHPLC sample prefractionation. *Environ. Sci. Technol.* 47, 4471–4479.
- RAMP (Regional Aquatics Monitoring Program), 2000. Oil Sands Regional Aquatic Monitoring (RAMP) 1999 Report. Golder Associated Ltd.
- Ross, M.S., Pereira, A.S., Fonnell, J., Davies, M., Johnson, J., Sliva, L., Martin, J.W., 2012. Quantitative and qualitative analysis of naphthenic acids in natural waters surrounding the Canadian oil sands industry. *Environ. Sci. Technol.* 46, 12796–12805.
- Sleighter, R.L., Lie, Z.F., Xue, J.H., Hatcher, P.G., 2010. Multivariate statistical approaches for the characterization of dissolved organic matter analyzed by ultrahigh resolution mass spectrometry. *Environ. Sci. Technol.* 44 (19), 7576–7582.
- St Amour, N.A., Gibson, J.J., Edwards, T.W.D., Prowse, T.D., Pietroniro, A., 2005. Isotopic time-series partitioning of streamflow components in wetland-dominated catchments, lower Liard River basin, Northwest Territories, Canada. *Hydrol. Process.* 19, 3357–3381.
- Studabaker, W.B., Krupa, S., Jayanty, R.K.M., 2012. Measurement of polynuclear aromatic hydrocarbons (PAHs) in epiphytic lichens for receptor modeling in the Athabasca oil sands region (AOSR): a pilot study. In: Percy, K.E. (Ed.), *Alberta Oil Sands: Energy, Industry and the Environment*. Elsevier, Oxford, pp. 391–425.
- Timoney, K.P., Lee, P., 2009. Does the Alberta tar sand industry pollute? The scientific evidence. *Open Conserv. Biol.* 3, 65–81.
- Timoney, K.P., Lee, P., 2011. Polycyclic aromatic hydrocarbons increase in Athabasca River Delta sediment: temporal trends and environmental correlates. *Environ. Sci. Technol.* 45 (10), 4278–4284.
- Vehniäinen, E.R., Häkkinen, J., Oikari, A., 2003. Photon-induced lethal and sublethal toxicity of retene, a polycyclic aromatic hydrocarbon derived from resin acid, to coregind larva. *Environ. Toxicol. Chem.* 22, 2995–3000.
- Wang, X.M., Kasperski, K.L., 2010. Analysis of naphthenic acids in aqueous solution using HPLC-MS/MS. *Anal. Methods* 2, 1715–1722.
- Wang, Z.D., Yang, C., Parrott, J.L., Frank, R.A., Yang, Z., Brown, C.E., Hollebone, B.P., Landriault, M., Fieldhouse, B., Liu, Y., Zhang, G., Hewitt, L.M., 2014. Forensic source differentiation of petrogenic, pyrogenic and biogenic hydrocarbons in Canadian oil sands environmental samples. *J. Hazard. Mater.* 166–177.
- Wiklund, J.A., Hall, R.L., Wolfe, B.B., Edwards, T.W.D., Farwell, A.J., Dixon, D.G., 2012. Has Alberta oil sands development increased far-field delivery of airborne contaminants to the peace-Athabasca Delta? *Sci. Total Environ.* 433, 379–382.
- Woo, M.K., Thorne, R., 2003. Streamflow in the Mackenzie Basin, Canada. *Arctic* 56, 328–340.
- Yi, Y., Gibson, J.J., Birks, S., Han, J., Borchers, C.H., 2014. Comments on “profiling oil sands mixtures from industrial developments and natural groundwaters for source identification”. *Environ. Sci. Technol.* 48, 11013–11014. <http://dx.doi.org/10.1021/es503498p>.