Airborne and ground-based measurements of aerosol optical depth of freshly emitted anthropogenic plumes in the Athabasca Oil Sands region

Konstantin Baibakov[1*], Samuel LeBlanc[2,3], Keyvan Ranjbar[4], Norman T. O'Neill[4], Mengistu Wolde[1], Jens Redemann[5], Kristina Pistone[2,3], Shao-Meng Li[6,7], John Liggio[6], Katherine Hayden[6], Tak W. Chan[8], Michael J. Wheeler[6], Leonid Nichman[1], Connor Flynn[5] and Roy Johnson[3].

[1]Flight Research Laboratory, National Research Council Canada, Ottawa, Canada
[2]Bay Area Environmental Research Institute, Moffett Field, CA, USA
[3]NASA Ames Research Center, Moffett Field, CA, USA
[4]University of Sherbrooke, Sherbrooke, Canada
[5]School of Meteorology, University of Oklahoma, Norman, OK, USA
[6]Air Quality Process Research Section, Environment and Climate Change Canada, Toronto, ON, Canada
[7]College of Environmental Science and Engineering, Peking University, Beijing, 100871 China
[8]Climate Chemistry Measurement and Research, Climate Research Division, Environment and Climate Change Canada, ON, Canada
* now at Canadian Space Agency, Saint-Hubert, Canada

Correspondence to: Konstantin Baibakov (Konstantin.Baibakov@~~nrc-cnrc.gc~~canada.ca)

Field Code Changed

Abstract.

In this work we report the airborne aerosol optical depth (AOD) from measurements within freshly-emitted anthropogenic plumes arising from mining and processing operations in the Athabasca Oil Sands Region (AOSR) in the context of ground-based AERONET climatological daily averaged AODs at Fort McMurray (Alberta, Canada). During two flights on June 9 and June 18, 2018, the NASA airborne 4STAR (Spectrometers for Sky-Scanning, Sun-Tracking Atmospheric Research) sunphotometer registered high fine-mode (FM, $< 1\mu m$) in-plume AODs of up to 0.4 and 0.9, respectively, in the vicinity of the plume source (< 20 km). Particle composition shows that the plumes were associated with elevated concentrations of sulphates and ammonium. These high AODs significantly exceed climatological averages for June and were not captured by the nearby AERONET instrument (mean daily AODs of 0.10 ± 0.01 and 0.07 ± 0.02, maximum AOD of 0.12) due possibly to horizontal inhomogeneity of the plumes, plume dilution, and winds which in certain cases were carrying the plume away from the ground-based instrument. The average 4STAR out-of-plume (background) AODs deviated only marginally from AERONET daily-averaged values. While 4STAR AOD peaks were generally well correlated in time with peaks in the in-situ measured particle concentrations, we show that differences in particle size are the dominant factor in determining the 4STAR derived AOD. During the two flights of June 24 and July 5, 2018 when plumes likely travelled distances of 60 km or more, the average 4STAR FM AOD increased by 0.01-0.02 over ~50 km of downwind particle evolution which was supported by the increases in layer AODs calculated from the in-situ extinction measurements. Based on these observations as well as the increases in organic mass, we attribute the observed AOD increase, at least in part, to secondary organic aerosol formation.

The in-plume and out-of-plume AODs for this second pair of flights, in contrast to clear differences in in-situ optical and other measurements, were practically indistinguishable and compared favorably to AERONET within 0.01-0.02 AOD. This means that AERONET was generally successful in capturing the background AODs, but missed some of the spatially constrained high-AOD plumes with sources as close as 30-50 km, which is important to note since the AERONET measurements are generally thought to be representative of the regional AOD loading. The fact that industrial plumes can be associated with significantly higher AODs in the vicinity of the emission sources than previously reported from AERONET can potentially have an effect on estimating the AOSR radiative impact.

1 Introduction

Atmospheric aerosols play an important role in the Earth's climate. Depending on their properties (such as size, composition, and concentration) and climatological factors (such as ambient radiative energy budget and surface albedo), aerosols can either cool the surface by reflecting sunlight back to space or have a net warming effect by directly absorbing this shortwave radiation (e.g. Kaufman et al., 2002, Anderson et al., 2003, Charlson et al., 2005). Aerosols can also affect cloud properties by modifying droplet size, cloud lifetime, and cloud burn-off rates (Ackerman et al., 2000, Wilcox, 2010, McComiskey and Feingold, 2012). Despite recent progress, aerosols, together with clouds, remain the greatest source of uncertainty in climate models (Boucher et al., 2013). This is in part due to the complexity of aerosol-generating processes (natural and anthropogenic), and aerosol transport mechanisms resulting in particles of various size, concentration, chemical composition, and mixing state that ultimately determine aerosol radiative impact. A recent work by Matus et al., 2019, based on the multi-sensor A-train observations, estimates that the global mean aerosol direct radiative effect is -2.40 W/m^2 with anthropogenic direct radiative forcing accounting for 21%.

The Athabasca Oil Sands Region (AOSR) in Alberta, Canada, is home to the third-largest oil deposit in the world (behind Venezuela and Saudi Arabia) with an estimated 164.1 billion barrels of recoverable crude oil (as of December 2018, see NRCAN, 2020). In addition to the emissions of anthropogenic pollutant and greenhouse gases (e.g., Charpentier and Bergerson, 2009, Bytnerowicz et al., 2010, Baray et al., 2018, Liggio et al., 2019) the AOSR generates a substantial amount of aerosols, either directly as sulfate, black carbon (BC), primary organic aerosol (POA) and dust (e.g., Landis et al. 2017, Cheng et al., 2018), or from secondary atmospheric processes. In fact, Liggio et al., 2016 showed that the AOSR is a large source of secondary organic aerosol (SOA, formed through gas-to-particle conversion) with formation rates comparable to major metropolitan areas. Understanding the radiative impact of anthropogenically produced aerosols in the AOSR through better characterization of their radiative properties is therefore of significant importance.

The estimation of aerosol radiative impact requires knowledge of key aerosol properties. One of the most important of these properties is the aerosol optical depth (AOD) (e.g., Kahn, 2012), defined as the total vertical column of aerosol extinction (the

combined capacity of aerosol particles to scatter and absorb light). The mid-visible AOD (typically reported at wavelengths between 440 and 550nm) can vary significantly from less than 0.1 for rural sites with background aerosols to more than 3 for aerosol plumes from intense forest fires, dust storms or volcanic eruptions. Model analysis suggests that 95% of global AOD at 440 nm, however, are below 0.4 (Andrews et al., 2017). AOD measured from or near ground level is commonly retrieved

5 using passive sunphotometry. A sunphotometer measures AOD in several spectral channels and can yield an estimation of particle abundance as well as aerosol size indicators (effective radius (r_{eff}) of submicron ("fine mode", FM) and supermicron ("coarse mode", CM) modes for example) from the spectral information (O'Neill et al., 2003). There are currently two Cimel CE-318 ground-based sunphotometers operating in the AOSR as part of the global AERONET sunphotometer network: one near Fort McMurray (56.8N, 111.5W, operating since 2005) and one near Fort McKay (57.2N, 111.6W, operating since 2013).

The analysis of local AOD climatologies including that of AOSR, serves as a basis for a global aerosol climatology and is essential for validating satellite retrievals (Toledano et al, 2007, Sioris et al, 2017b). Sioris et al, 2017a analyzed AOD data records for several Canadian AERONET stations, including Fort McMurray, in the context of ground-based particulate matter ($PM_{2.5}$) concentrations and presented annual climatologies for both the FM and CM AODs. They found that Fort McMurray

15 FM AODs (at 500 nm) are generally < 0.1 for Sep.-Apr. and < 0.2 for May-Aug. The summer-time AOD increase was attributed to forest fires plumes and no significant trends in the long-term AOD record were observed. In addition to ground-based measurements, Shinozuka et al., 2011 reported on AOD measurements acquired on board the NASA P-3 aircraft using the AATS-14 sunphotometer (AOD from 14 wavelength bands in the 354-2139 nm range) in the vicinity of Fort McMurray in June-July 2008. While the AATS-14 measured AOD_{499} values that frequently exceeded 1 and at times reached 4, the study

20 was entirely focused on biomass burning aerosols from local forest fires and did not include measurements of the oil sands industrial plumes. Indeed, most airborne studies have predominantly focussed on either naturally produced aerosols (such as forest fires or dust events) and/or significantly aged plumes that have been transported for hundreds or thousands of kilometers (Heald et al., 2006, Chin et al., 2007, Stone et al., 2010). While multiple studies have described cases of extreme pollution in China (e.g., Gu et al., 2018, Qin et al., 2018, Sun et al., 2018) significantly less attention has been given to AOD values from

25 isolated industrialized sources outside of Asia. Given that the anthropogenic plumes in the AOSR, far from large urban regions, are nearly entirely attributed to a singular industry, the obtained airborne AOD measurements can be used to evaluate the representativeness of AERONET measurements in the context of industrial emissions. This is especially relevant to the AOSR, where secondary aerosol formation is likely to occur away from source areas and associated AERONET sites.

30 In the spring-summer of 2018, the NASA 4STAR sunphotometer (Spectrometers for Sky-Scanning, Sun-Tracking Atmospheric Research, Dunagan et al, 2013, Shinozuka et al, 2013, Pistone et al, 2019, LeBlanc et al, 2020) was integrated on the National Research Council (NRC) Convair-580 aircraft, as part of state-of-the-art suite of in-situ and remote sensing instruments. The NASA 4STAR acquired airborne hyperspectral AOD data during the Oil Sands Measurement Campaign of

2018 (OSMC), and represented the first airborne sunphotometer deployment to study industrial pollutant plumes from the Alberta oil sands mining and upgrading operations.

In the current work, based on four OSMC flights, we evaluate whether the existing ground based AERONET measurements
5   capture the full extent of AOD variations in the AOSR by comparing the 4STAR AODs of anthropogenic plumes from mining and processing operations with AERONET climatological and diurnal AODs. In addition, we investigate how the AOD evolves as a function of distance downwind of the plume source, in order to better understand the potential effect of secondary aerosol production, i.e. whether the AOD increases as new particles are being formed – a phenomenon that would be missed by a fixed ground-based instrument located near the source. Finally, we consider how the 4STAR FM and CM AODs relate
10 to other in-situ measurements such as particle size, optical absorption and scattering coefficients, and particle chemical composition. The latter analysis provides additional insight into the observed AOD variations and ensures that the 4STAR measurements are physically consistent with other instruments.

2 Instrumentation and data processing

2.1 Oil Sands Measurement Campaign (OSMC)

15 The Oil Sands Measurement Campaign 2018 (OSMC2018, hereafter referred to as OSMC) was a follow-up to the OSMC2013 campaign, where the Convair-580 aircraft was deployed to the Oil Sands region for the first time (see e.g. Liggio et al., 2016, Li et al, 2017). Figure 1 shows the location of the AOSR facilities overlaid on recent satellite imagery. Two types of flights performed during the OSMC focused on (i) pollutant emissions and (ii) their transformation (i.e., chemical evolution/aging) using "box" (latitude/longitude grid) and "screen" (vertical grid) flight patterns, respectively. Box flights encompassed an
20 entire facility to characterize pollution emissions from individual oil sands processing facilities, while screen flights were conducted to study the evolution of the pollutants transported away from the source. Each box/screen started at 500 ft above ground level (AGL) and progressively increased in altitude at intervals of approximately 250 feet or more, until reaching the top of the boundary layer typically resulting in 3-6 altitude levels. The transformation flights were conducted as a Lagrangian experiment designed to intercept the same plume at increasing distances from the source perpendicular to the wind direction.
25 The measurements presented here were obtained during the second phase of the campaign between May 30 and July 5, 2018.

2.2 Sunphotometry measurements

2.2.1 AERONET and CIMEL CE-318

The Cimel CE-318 used at Fort McMurray and Fort McKay are standard ground-based sunphotometers which are part of the global AERONET sunphotometer network. The CE-318 takes measurements of the direct solar irradiance, at nominal sampling
30 intervals of three minutes in eight spectral bands (340, 380, 440, 500, 675, 870, 1020, and 1670 nm) from which spectral AODs

can be calculated. For more information about the AERONET and CIMEL CE318 the reader is referred to Giles et al., 2019 In this work we used AERONET Level 2 (cloud-screened and quality controlled) Version 3 data.

2.2.2 4STAR (Spectrometers for Sky-Scanning Sun-Tracking Atmospheric Research)

The NASA 4STAR instrument combines airborne sun tracking and sky scanning with diffraction spectroscopy. It is comprised
5 of a movable optical head that protrudes through the top of the aircraft fuselage, an accompanying instrument rack inside the aircraft and optical cables that connect the two (Figure S1). The hyperspectral measurements are obtained with two Zeiss spectrometers resulting in a total of 1556 spectral channels between 210 and 1703 nm with sampling resolution of 0.2-1 nm below 1000 nm and 3-6 nm at longer wavelengths. The nominal calibration accuracy of AOD measurements from 4STAR are dependent on wavelength, time of day, tracking stability, and various corrections (such as removal of light absorption by trace
10 gases). This accuracy is typically near 1% in transmittance, for wavelengths outside gas absorption bands between 355 nm and 1650 nm, resulting in an AOD uncertainty of 0.01 when the sun is at zenith. A more detailed technical description of the 4STAR can be found in Dunagan et al, 2013 and Shinozuka et al, 2013.

2.2.3 Spectral Deconvolution Algorithm

We applied the Spectral Deconvolution Algorithm (SDA; O'Neill et al, 2003) in order to separate the total measured AOD
15 into its fine (submicron) and coarse (supermicron) mode components (τ_f and τ_c or FM and CM). While AERONET uses 380, 440, 500, 675, and 870 nm channels as standard inputs to SDA, using the 380 nm channel in the 4STAR analysis was not possible because of its unrealistically high values (thought to be due to stray light at that wavelength) aggravated by high relative humidity within the spectrometer enclosure. Reducing the number of channels to four is thought to produce results commensurate with the standard AERONET wavelength set with RMS differences in retrieved τ_f of less than 0.01 (O'Neill
20 et al., 2008). We expect that, in most cases, the anthropogenically produced aerosols observed in the AOSR are submicron, since coarse-mode aerosols are usually associated with natural sources (e.g., winds lifting up dust particles). Consequently, with the exception of dust aerosols resulting from surface mining, the coarse-mode dominant AODs in the AOSR will usually be associated with clouds. As will be shown below, the CM AODs associated with dust in the AOSR are small (< 0.05) and can easily be distinguished from cloud-dominated events that are associated with significantly higher AODs and larger
25 variations over short periods of time.

2.2.4 4STAR data filtering and cloud screening

In this work we have eliminated all 4STAR measurements with low (raw count) signal to noise ratio and ignored measurements when the aircraft was not in a straight-level track (i.e. when turning or engaged in a spiral). Moreover, in addition to eliminating improbable data (for AOSR) with $AOD_{500} > 3$, we have used the SDA as a spectral cloud-screening where the FM AOD is
30 thought to be free of coarse-mode contamination by either clouds or dust (e.g., O'Neill et al, 2016).

2.4 In-situ measurements

Ambient air was drawn through a forward facing, shrouded isokinetic particle inlet (Droplet Measurement Technologies, Boulder, CO, USA). The fine-mode particle size (~~diameter~~ radius ranging from ~~60~~30 to ~~1000~~500 nm) was measured using a DMT Ultra-High-Sensitivity Aerosol Spectrometer (UHSAS), which employs light scattering techniques to derive particle

5   concentrations and size distributions (Cai et al, 2008, Kupc et al, 2018). The UHSAS sizing was calibrated using NIST traceable polystyrene latex (PSL) nanospheres. Sizing of the UHSAS is dependent on the refractive index and shape of the particles. Differences in refractive index have been estimated to result in a 10% uncertainty in the sizing of the UHSAS (Kupc et al, 2018). For some flights we noticed abnormally high particle counts in 5 bins between the radius of 0.382 and 0.428 μm. This peak was not supported by other particle spectrometers on the aircraft and is likely an instrument artefact. We removed

10   the problematic data from further analysis and suspect that the issue is related to the uncertainties in the UHSAS calibration curve consisting of several individually chosen gains. The UHSAS effective radius (r_{eff}), defined as an area weighted mean radius of the aerosol particles, was estimated using

$$r_{eff} = \frac{\int_0^\infty r^3 n(r) dr}{\int_0^\infty r^2 n(r) dr} \quad (1)$$

where r is the particle radius and $n(r)$ is the particle size distribution (i.e., number of particles per cm^2 having a radius in the

15   range r and $r+dr$ microns). The particle in-situ scattering and absorption coefficients were measured using the TSI 3563 nephelometer and Continuous Light Absorption Photometer (CLAP), respectively. The TSI 3563 measures the light scattered by aerosols ultimately yielding the aerosol total scattering coefficient at 450, 550, 700 nm (Anderson et al., 1996; Bodhaine et al., 1991) while CLAP, developed by National Oceanic and Atmospheric Administration (NOAA), measures light absorption of particles deposited on 47-mm diameter, glass-fiber filters at 467, 529, 653 nm (Ogren et al., 2017 and reference therein).

20   Aerosol single scattering albedo (SSA) values were estimated using the ratio of the light scattering coefficient to the sum of the light absorption and scattering coefficients. Black carbon (BC or soot) mass was measured using the DMT Single Particle Soot Photometer (SP2) based on the laser-induced incandescence between approximately 100 and 600 nm (e.g., Schwarz et al., 2006). Finally, non-refractory particle composition (ammonium, nitrate, sulfate and organics) was measured with a High-Resolution Time-of-Flight Aerosol Mass Spectrometer (HR-ToF-AMS; Aerodyne Research Inc., DeCarlo et al., 2006). The

25   AMS was operated in V-mode with 10 s time resolution. Ionization efficiency calibrations were performed using monodisperse ammonium nitrate particles prior to, during and after the study with a resulting <15% variation. The AMS collection efficiency, ranging from 0.5 to 1.0, was derived by comparing the total AMS mass with the mass derived from the UHSAS size distributions assuming a density consistent with the AMS chemical composition. Background measurements were taken from filtered ambient air approximately 4-5 times per flight. Detection limits, taken from three times the standard

30   deviation of the average of filtered time periods, were determined to be 0.122, 0.024, 0.021, 0.260 ug/m3 for NH_4, SO_4, NO_3, and Org, respectively.

3. Results and discussion

3.1 AERONET AOD climatology

Understanding seasonal AOD variations from the ground-based Fort McMurray instrument is useful for the interpretation of the airborne measurements. While this analysis is required to evaluate whether the AERONET measurements capture the full extent of AOD variations in the AOSR, understanding climatological AODs also provides context for the 4STAR OSMC measurements acquired mostly during a single month of June 2018. Sioris et al., 2017a discussed some of the Fort McMurray AOD statistics, however, their analysis ends in 2015. We have extended the analysis to 2018 in order to include the period leading up to and coincident with the OSMC.

Figure 2 shows that the total AODs at Fort McMurray are relatively low with values typically less than 0.1 except for summer months (May-Aug.) when monthly-averaged AODs can exceed 0.2. There are clear seasonal trends in both the FM and CM AODs. The summer-time FM AOD increase is thought to be predominantly associated with forest fires that frequently occur in the region (Sioris et al., 2017a). For example, the second highest monthly AOD on record of 0.45 occurred during May 2016 when intense forest fires had a major impact on air quality and resulted in the evacuation of more than 80,000 people (Landis et al, 2018). The highest hourly- and daily-averaged AODs during that month were 2.10 and 1.80, respectively. Figure S2 shows monthly-averaged AOD time series for 2005-2018 at Fort McMurray with daily-averaged values shown in pink shading. Also, included for reference is a fine-mode AOD time series for nearby Fort McKay, however, the present discussion focuses on Fort McMurray because of its significantly longer measurement record and data availability during OSMC. In many cases, on days with persistently high AOD values (e.g., FM AOD > 0.5) forest fire smoke plumes were clearly discernable in the vicinity of Fort McMurray from satellite imagery (not shown). In comparison, the highest FM monthly AOD average during the non-summer months (Sep.-Apr.) – considered to be primarily impacted by industrial pollution plumes in the absence of other major aerosol sources – was 0.13 with 96% of all hourly measurements being lower than that value.

With respect to the coarse-mode, the peak monthly-averaged AOD of 0.04 is associated with the month of April and is thought to be due to dust particles. Landis et al. (2017) demonstrated through systematic chemical analysis of summertime CM aerosol from ground-based measurements that a majority of dust is due to local surface mining activities where heavy machinery (trucks and excavators) lift significant amounts of dust into the air, although long-range transport of Asian dust may be more prominent in the spring (e.g., McKendry et al., 2007, AboEl-Fetouh et al., 2020). AboEl-Fetouh et al., 2020 argued, for example, that a springtime coarse mode event was continental in scale (including the North American and European Arctic) and that it was associated with (relatively small) coarse mode particles around 1.3 μm in radius.

Figure 2 also indicates that the mean total AOD during June 2018 was 0.13 ± 0.05 suggesting that the aerosol loading during the OSMC was generally representative of the climatological AOD values for June (0.15 AOD).

3.2 4STAR case studies during OSMC

In this section, based on the OSMC flights of June 9 and 18, we discuss how the 4STAR AOD variations associated with the AOSR emitted plumes relate to the corresponding aircraft in-situ measurements. While direct comparisons between the 4STAR column-based AODs and flight-level in-situ aerosol properties are not strictly possible, evaluating the AOD
5 measurements in the context of time-synchronous in-situ measurements provides insight into the nature of the plume and the associated AOD variations (e.g., Stone et al., 2010, Shinozuka et al., 2013, Pistone et al., 2019). In particular, understanding how changes in aerosol intensive properties such as particle size and chemical composition might affect the AOD response is necessary to evaluate the radiative importance of the emitted plumes and the associated secondary organic aerosol production.

3.2.1 Case study of June 9, 2018

10 On June 9, the Convair-580 conducted a transformation flight to the west of the Oil Sands facilities. The aircraft flew 4 screens perpendicular to the wind direction intercepting the emitted plumes from the majority of the largest surface mining facilities (Figure 21). Figure 3 shows the AOD time series and corresponding in-situ measurements for the relatively cloud-free screen 1, with five nominal altitude levels and "in plume" and "out of plume" periods separated based on elevated values relative to the upwind (background) measurements. While the average fine-mode AODs were ~0.10 outside of the plume, the values
15 were greater than 0.4 within the plume. Moreover, the two consistent peaks at each altitude level in the UHSAS total particle number concentration (pane 2) suggest that the Convair-580 was intercepting two spatially separated plumes where only one is associated with a significant AOD response. Particle composition data (pane 4 of Figure 3 and Figure 4) shows that the first plume (plume "A", detected by the 4STAR and shaded in pink) contained significantly more SO_4 and NH_4 than the second plume (plume "B", shaded in grey) indicating that the two plumes were likely originated from different facilities. The organic
20 aerosol mass, however, was comparable between the two plumes and similar to the background level. Figure 3 pane 3 and Figure 5 show a clear distinction in particle effective radius and volume size distribution between plumes A and B in that the latter lacks optically significant contribution from larger FM particles. The difference is especially pronounced, both in comparison to plume B and out of plume (background) values, for particle radii (r_p) larger than ~0.12 μm with the largest differences observed at radii 0.3 and 0.42 μm. The relative differences between plumes A and B continue to persist for screen
25 two, approximately 20 km downwind (not shown) but the concentration of large FM particles larger than ~ 0.12 μm for plume A was significantly lower (by a factor of 2 or more relative to screen one) as the plume disperses and some of the larger particles settle out of the atmosphere. Based on wind patterns, the plumes on June 9 come mostly from the Syncrude Mildred Lake (SML) and/or Suncor Millenium facilities which perform a number of different processes related to bitumen extraction. The majority of emissions come either from upgrading the bitumen (and in the process emitting SO_2 and SO_4), mining
30 processes including from big trucks digging and transporting materials, and other unknown processes within the plant itself. The two plumes (A and B) mostly represent emissions and subsequent transformation of pollutants from the upgrading (high

SO$_4$ concentrations), but can also contain contributions from mining and tailing ponds. We expect that these two plumes may mix with each other at times.

The June 9 aerosol plume from the Oil Sands processing could be observed in the MODIS satellite imagery, in between the often-prevalent cloud conditions (see Figure 6, left). The aerosol plume was being advected towards the northwest, with the background aerosol loading in between clouds, mostly with AODs below 0.1. The 4STAR-sampled aerosol plume was observable in the satellite imagery, with AOD$_{500}$ exceeding 0.5. For this case, there are only very limited collocated AOD measurements simultaneously acquired with MODIS and 4STAR. In Figure 6 (right) all pixels that were measured by MODIS (within 3km of the flight track) and 4STAR (within 3 hours of Terra overpass) are compared. The differences between the 4STAR measurements (N=2,510) and the average nearest MODIS pixel show that much of the higher AOD observed by 4STAR (above 0.5) in the middle of the plume is not captured by MODIS (with a maximum AOD of 0.5), and that there is an overestimate at the lower end of the 4STAR observed AODs. When considering only the nearest 30 minutes of the MODIS overpass, the cloud free observations by 4STAR (N=327) are not centered on the optically thickest part of the plume, but show MODIS overestimating the AOD by an average of 0.1 (up to a maximum overestimate by 0.25) for the 4STAR AODs at 0.2.

Based on the 4STAR and in-situ data, the aerosol plume was ~19 km wide (spanning plumes A and B) for the first screen (at a distance of emissions of ~20 km, assuming Syncrude Mildred Lake as a plume source). This evolved to a width of ~31 km for the last screen, at a further distance of ~70km from emissions, although the plume boundaries are harder to identify amid a smaller difference with background values. This is related to the dilution of aerosol over a broader volume, potentially reducing AOD for any one column, but also to a (marginally) larger increase of AOD due SOA formation (the last screen was heavily contaminated by clouds resulting in practically no useable AOD data). The AOD variations in this plume can be compared to previous observations in the boreal region as reported by Shinozuka and Redemann (2011), which showed that for the Canadian local emissions from forest fires, variations of near 20% were observed at length scales of 20 km to 35.2 km. The aerosol from long-range transport showed variation of less than 5% (Shinozuka and Redemann, 2011). For the OSMC plumes observed here, a relative change in AOD of about 20% was observed at distances on a distance scale of 50km, but only for a subset of the flights (July 5th, June 24th, and June 18th, example of which showing differences between AERONET and 4STAR AOD as a function of distance, Figure S5). The differences in length scales of the aerosol plumes are indicative of both the industrial processes and the meteorological conditions, and seem to be different than the plumes from biomass burning in the boreal forest.

3.2.2 Case study of June 18, 2018

During the June 18 emissions flight around the ~~Syncrude Mildred Lake~~ (SML) facility, there was a significant and clearly visible industrial plume emanating from the SML stacks (Figure S3). Figure 6̶7 shows the corresponding data for the first three altitude levels: 500, 750, and 1000 ft with the 500 ft track repeated twice. The 4STAR measurements show that for the duration of the Syncrude box the total AOD at 500 nm was relatively low and stable around 0.05-0.07, with AOD spikes of up to 0.86 associated with the plume on the eastern side of the box. As in the case of June 9, the measurements show that the plume was heterogeneous with two distinct regions identified in Figure 7̶6 in pink (plume "A") and grey ("plume "B"). Plume A is characterized by a significant response in fine-mode AOD and increased (up to a factor of 2) UHSAS effective radius. Plume A is also associated with high values of SSA_{440} (up to 0.96, pane 4), indicating that the particles were mostly of a scattering nature. In contrast, the SSA values dip below 0.75 in plume B suggesting a much stronger presence of absorbing particles. This hypothesis is confirmed by the enhanced concentrations of highly absorptive BC particles (pane 5), with BC mass increasing during certain times by more than a factor of 10 relative to plume A. Similarly to June 9, plume A contains significantly higher concentrations of SO_4 and NH_4 than plume B (not shown). It should be noted that the AOD peaks precede most of the plume A in-situ peaks, which we attribute to the 4STAR viewing geometry in relation to the vertical extent of the plume (i.e. the relative position of the sun, the plume and the aircraft carrying the 4STAR). Plume B contains fewer of the larger FM particles and goes mostly undetected by 4STAR. At the same time, plume B contains higher concentrations of CM particles from surface mining (not shown) but the resulting CM AOD (for example, around 15:22) is < 0.03 which is significantly smaller than the FM AODs.

For both flights, June 9 and June 18, the analysis indicates that differences in particle size within plumes are the driving factor responsible for the AOD response. As a consequence, while certain sections of the plume (e.g., plume B) are clearly associated with elevated in-situ particle and gas concentrations, they have little effect on AOD when particulates are small (in this case $r_p < 120$ nm).

3.3 4STAR-AERONET comparisons

The inherent assumption of using the AERONET data is that it is representative of the regional AOD within a certain distance from the ground-based station. While this distance is often taken as 100 km, the spatial representation will ultimately depend on the local context including proximity to aerosol sources and meteorology (Anderson et al, 2003, Holben et al, 2018). Consequently, in this section we investigate the representativeness of Fort McMurray AERONET measurements of AOSR AODs by studying how the in-flight AOD variations compare to climatological and diurnal ground-based AODs.

Figure 7̶8 and Table 1 show AOD vertical profiles and AOD statistics for 3 transformation flights (June 9, June 24, and July 5) and 1 emission flight (June 18) during the OSMC. For June 9 we only included data from screen 1 and 2, beyond which the measurements were judged to be contaminated by clouds. The in-plume data were selected based on significant increases in the 4STAR, UHSAS, nephelometer, and CLAP signals relative to the background (out of plume) levels. The examples of

the defined "in plume" time periods are shown in Figure 3 and Figure ~~7~~6. While each screen was usually performed at fixed nominal altitudes (e.g., 500 ft, 750 ft above ground level), the terrain variations result in slightly different altitudes above sea-level when compared screen to screen. The bottom of Figure ~~7~~8 also shows fine-mode AERONET daily means and standard deviations for the corresponding dates. It should be noted that on June 24 the AERONET data was significantly affected by the incoming forest fires plume with AODs > 0.4 in the later part of the day. Removing those data after 18:13 (~ 27% of points) changes the daily average for June 24 from 0.13 ± 0.08 to 0.10 ± 0.03.

On June 9 (screen 1) and June 18 ("high-AOD flights") the plumes were sampled in the direct vicinity of the source. Consequently, the June 9 and June 18 cases were associated with particularly high AODs and show relatively large variations in terms of the vertical distribution and AOD variability within the plume. These large AODs and corresponding variations significantly exceed climatological AODs for June and are clearly not captured by AERONET. However, the average 4STAR out-of-plume (background) AODs deviate only marginally from AERONET daily-averaged values (Table 1). During the June 24 and July 5 flights ("low-AOD flights"), the plumes in general travelled longer distances than during the high-AOD flights before being measured by the Convair-580 and hence had more time to mix and disperse. The in-plume and out-of-plume AODs for these dates were indistinguishable (in contrast to clear differences in in-situ measurements, not shown) and the airborne and ground-based averages compared favorably within 0.01-0.02 AOD. In other words, the AERONET measurements capture the background AODs relatively well, but miss some of the spatially constrained high-AOD plumes with sources as close as 30-50 km.

It is not clear what kind of agreement would be considered as reasonable between 4STAR and AERONET given the differences in spatial sampling between the two platforms. While several studies have reported agreements down to RMS differences of 0.01 for wavelengths between 500 nm and 1020 nm (Stone et al., 2010; Shinozuka et al., 2013; LeBlanc et al., 2020), those comparisons often involve aged plumes that are likely to be more (horizontally) homogeneous than newly-emitted AOSR plumes. Furthermore, the comparisons usually include only a few points closest in time to the aircraft overpass of an AERONET station. Factors such as prevalent wind directions, local topography, and distance to emission sources can all bias comparisons between the aircraft and ground-based platforms. In particular, during the June 9 and June 18 flights, the Fort McMurray AERONET station was mostly located upwind of the Syncrude facilities and likely missed most of the plumes measured by the 4STAR. However, even when limiting the AERONET data to periods associated with northern winds (270 to 90° on a wind rose) likely carrying the industrial plumes (Figure 1), the mean and median June 2018 AODs were both 0.09 with only 4 individual measurements exceeding 0.2 (maximum of 0.29). We attribute these disparities to dilution as the plumes are being transported from their sources to where they are captured by the AERONET sunphotometer. Simply moving the AERONET station closer to specific facilities might yield a better agreement in certain cases, but the AOD differences will likely remain because of spatial inhomogeneity of the plumes and the inherent limitations of measurements from a single fixed ground-based station.

Figure 89 (top) shows 4STAR-AERONET scatterplots for selected flights. Only out-of-plume AODs acquired during a straight-level flight (i.e. excluding spirals) at an altitude of less than 1 km were considered for comparison. Each valid AERONET L2.0 direct-sun measurement was matched in time with the nearest 4STAR measurement within a 10 minute window (the 4STAR-AERONET distance varied between 21 and 154 km) and if no such measurements were found, the AERONET point was excluded from analysis. Even though the average out-of-plume AODs compared favourably to AERONET, the actual linear correlation coefficient (R^2) values were relatively low (~0.30) likely indicating spatial variations in the background aerosol loading and/or uncertainties in defining out-of-plume sections (there may well have been moderately higher AODs in the neighbourhood of those plumes that escaped our plume rejection criteria). A notable exception (R^2=0.65) was June 18 when both 4STAR and AERONET showed a slight low-frequency AOD increase (from 0.05 to 0.06) during the flight time (Figure S4, top pane). The bottom pane of that figure also shows that the AERONET $r_{eff,f}$ values agree particularly well with UHSAS but the 4STAR retrievals seem to be overestimated. We suspect that the source of this latter inconsistency is due to the artificially high 4STAR irradiance signals at the shorter wavelengths. Figure 89 (bottom) shows a histogram of τ_{diff} values (defined as $\tau_{diff} = \tau_{f,AST} - \tau_{f,AERONET}$) for all four flights combined. It is evident from this figure that despite the slightly shorter atmospheric column, 4STAR AODs are marginally overestimated relative to AERONET with $|\tau_{diff}|$ mean, median, and standard deviation being 0.02, 0.01 and 0.02. The interpretation of this result is not obvious. In principle, as the aircraft moves further away from sources the AOD should increasingly resemble the background and thus τ_{diff} should decrease with distance. On the other hand, as the distance between the two instruments increases so do the differences in sampled air masses. In fact, Sioris et al., 2017a (their Figure 4) showed that the two AERONET stations become progressively dissimilar as a function of distance between them with the correlation coefficient dropping by ~50 % in the first 500 km. This is consistent with Figure S5 which indicates that despite the significant variance in the data, the average τ_{diff} increases as a function of distance separating the 4STAR and the Fort McMurray AERONET station (up to 160 km) based on the data from the 4 flights. This suggests that either the 4STAR was influenced by dispersed plumes with AODs higher than the AOSR background even at larger distances (i.e. not fully diluted plumes), or that the AERONET AODs were influenced by cleaner air masses.

Another potential explanation accounting for some of the 4STAR overestimation is secondary organic aerosol (SOA) production which may increases with distance/time (e.g. Liggio et al., 2016). This possibility was investigated by targeting the transformation flights of June 24 and July 5, in an attempt to observe any systematic screen-to-screen AOD changes that might be due to SOA production - a phenomenon that could be missed by AERONET. As shown in Table 1 between screen 1 to 3 only a modest increase of 0.01 and 0.02 was observed. While these AOD changes are small and commensurate with the 4STAR measurement error, O'Neill et al., 2008 previously showed that changes on this scale can be linked to real underlying changes in aerosol properties. We compared the in-plume vertical profiles obtained during screens 2 and 3 of each of the two flights, excluding the data from screen 1 because on June 24 it was not properly positioned relative to other screens and on July 5 the

profile data only covered a short vertical range, which we judged to be inadequate for analysis. We found that for both dates the calculated in-situ extinction coefficient at 500 nm was higher at practically each level for screen 3 (Figure 910). Integrating the extinction coefficient over the common vertical range between the screens yielded (in-situ) layer AOD increases between 33 and 67%. Comparing these values to 4STAR was not always possible in part because of the missing sunphotometry data, and in part because the AOD didn't vary sufficiently (i.e. above the measurement uncertainty of 0.01) over these short vertical ranges of 200-500 m to be reliably measured. However, for the two longest profiles during screen 3 on July 5 spanning 800 m, the in-situ layer AOD of 0.013 ± 0.002 compared favorably to 4STAR AOD of 0.01 (taking standard deviation within a 50 m altitude bin as a measurement error for extinction and assuming that the errors in altitude measurements are negligible). Therefore, the observed 4STAR AOD increases, while small, are consistent with layer AOD increases calculated from in-situ extinction measurements. Particle composition data shows that the plumes were dominated by the organic content (> 78%) with the organic mass increasing from screen to screen despite dilution (Figure S6). Using a Top Down Emissions Rate Retrieval Algorithm (TERRA; Gordon et al., 2015) significant SOA formation rates were determined of 1003 ± 193 kg/hr and 443 ± 45 kg/hr for June 9 and July 5, respectively. Consequently, while acknowledging the limited scope of the two flights and the need for more case studies, we believe that the observed AOD increase can be attributed, at least in part, to the process of SOA formation.

Summary and conclusions

During the spring-summer 2018 Oil Sands Measurement Campaign, the NASA 4STAR sunphotometer acquired airborne hyperspectral AOD data in the Athabasca Oil Sands Region. In this work, we report on the airborne AOD measurements of anthropogenic plumes from mining and processing operations in AOSR in the context of AERONET climatological and diurnal AODs at Fort McMurray. The monthly-averaged AERONET AOD for June 2018, when most of the 4STAR sunphotometer data was acquired, was 0.13 ± 0.05 (mean $\pm$ std) which was in line with summertime climatological averages, indicating that the observational conditions during OSMC were representative of the typical aerosol loading in the area for that month.

During the two flights of June 9 and June 18 (high-AOD flights), while flying into the visually opaque industrial plumes, 4STAR registered high fine-mode in-plume AODs of up to 0.4 and 0.9, respectively, in the vicinity of the plume source (< 20 km). These high AODs significantly exceed climatological averages and were clearly not captured by the nearby AERONET instrument, which reported mean daily AODs of 0.10 ± 0.01 and 0.07 ± 0.02 (mean $\pm$ std) for those days. At the same time the average 4STAR out-of-plume (background) AODs deviated only marginally from AERONET daily-averaged values. We attribute the disparities principally to the horizontal inhomogeneity of the plumes, the distance between the plume source and the AERONET instrument as well as prevalent winds which, in certain cases, were carrying the plume away from the ground-based instrument.

For both flights, the 4STAR AOD peaks were generally well correlated in time with peaks in in-situ measured particle and gas concentrations for the optically active sections of the plume (plume "A", i.e. those sections that were associated with significant increases in in-situ extinction and/or AOD), while other sections, clearly associated with elevated in-situ concentrations (relative to out-of-plume measurements), produced no significant 4STAR response (plume "B"). We attribute this difference

5 to the presence of larger FM particles within plume A (FM effective radius increasing from 0.10 to 0.18 µm) relative to the out-of-plume values and plume B. Particle composition analysis shows that plume A contained elevated levels (by up to a factor of 6) of sulphates and ammonium while the organic mass remained largely comparable between the two plumes.

During the June 24 and July 5 flights (low-AOD flights), the plumes likely travelled longer distances before being measured
10 by the Convair-580 onboard instruments (the Syncrude plume for example would have travelled ~60 km, or three times the distance than on June 9 before being first measured) and hence had more time to mix and disperse. The in-plume and out-of-plume 4STAR AODs for these dates were practically indistinguishable (average AODs of ~0.12 and 0.05 for June 24 and July 5, respectively) and the airborne and ground-based averages compared favorably to within 0.01-0.02 AOD. Comparing screen 1 to 3 AODs for these flights shows a small increase of 0.01-0.02, which was supported in several cases by the increases in
15 layer AOD calculated from the in-situ extinction measurements. Based on these observations and the screen 1 to 3 increases in organic content and organic fraction, we attribute the observed AOD increase, at least in part, to the process of SOA formation.

Comparing out-of-plume 4STAR AODs to AERONET measurements shows that despite the slightly shorter atmospheric
20 column 4STAR, AODs are marginally overestimated ($\left|\tau_{diff}\right| = 0.02$) with the difference increasing as a function of distance from the AERONET station. This suggests that either the 4STAR was still influenced by plumes with AODs higher than the AOSR background (presumably measured by AERONET) even at larger distances or that the AERONET AODs were generally underestimated during those flights (due for example to being influenced by cleaner air masses).

25 Finally, it is noted that low AOD AERONET measurements should be interpreted with care when they are used to represent the AOSR average. We showed that industrial plumes can be associated with significantly higher AODs in the vicinity of the emission sources and their full extent is not captured by the ground-based instruments even with sources as close as 30-50 km. The impact of aerosol in-plume to the instantaneous change in solar irradiance is larger than the background aerosol for any other part of the screens and flight segments in the surrounding, with changes occurring at the km-scales. By using the change
30 in measured AOD and in-situ aerosol optical properties (measured single scattering albedo and asymmetry parameter derived from Mie calculations of the aerosol size and refractive index), we calculated the change in instantaneous direct aerosol radiative effect (using the Fu-Liou radiative transfer code as part of LibRadtran; Emde et al., 2016). The aerosol within the industrial plume cooled the surface, through a decrease of the integrated net irradiance change (W/m^2), by an average reduction of incident light by 25%, as compared to the immediately surrounding background aerosol. This shows the potential to effect

estimations of the aerosol's regional radiative impact, particularly if only using the AOD from the nearby AERONET site as sole indicator.

Acknowledgements

Funding for the NRC Convair-580 participation in the Oil Sands 2018 field project was provided by ECCC and NRC. We would like to acknowledge the NRC and ECCC project crew that supported the field project. We also acknowledge Warren Gore (retired) from NASA Ames Research Center for his efforts in bringing the 4STAR to Canada for the duration of the project. We thank Ihab Abboud and Vitali Fioletov for establishing and maintaining the Fort McMurray and Fort McKay AERONET sites used in this investigation. This work was partially funded under the Oil Sands Monitoring (OSM) program. It is independent of any position of the OSM program. The U.S. authors were supported in part by NASA Radiation Science Program, under the direction of Hal Maring.

Author contribution

KB, MW, SL and JR conceived the study. MW, JL, JR and SL acquisitioned the funding. MW and JR oversaw the use of 4STAR in Canada. KB and MW collected the 4STAR data on board the NRC Convair-580. KB, SL and KP analysed the 4STAR direct-sun measurements. KR, NTO and KB performed the SDA analysis. KP, SL and CF processed the sky scan data. TC and KB analysed nephelometer and CLAP data. MJW, KB and LN analysed UHSAS and SP2 data. KB analyzed the AERONET data with contributions from KR, NTO and SL. RJ and SL provided engineering support for integrating and running the 4STAR on board the NRC Convair-580. KB wrote the paper with contributions and reviews from all authors.

References

AboEl-Fetouh, Y., O'Neill, N. T., Ranjbar, K., Hesaraki, S., Abboud, I., & Sobolewski, P. S. (2020). Climatological-scale analysis of intensive and semi-intensive aerosol parameters derived from AERONET retrievals over the Arctic. *Journal of Geophysical Research: Atmospheres*, 125,https://doi.org/10.1029/2019JD031569

Ackerman, A., Toon, O., Stevens, D., Heymsfield, A., Ramanathan, V., and Welton, E.: Reduction of tropical cloudiness by soot, Science, 288, 1042–1047, https://doi.org/10.1126/science.288.5468.1042, 2000.

Anderson, T. L., et al. (2003), Climate forcing by aerosols—A hazy picture, Science, 300, 1103– 1104, doi:10.1126/science.1084777.

Anderson, T. L., Charlson, R. J., Winker, D. M., Ogren, J. A., and Holmén, K.: Mesoscale variations of tropospheric aerosols, J. Atmos. Sci., 60, 119–136, https://doi.org/10.1175/1520-0469(2003)060<0119:MVOTA>2.0.CO;2, 2003.

Anderson, T.L., Covert, D.S., Marshall, S.F., Laucks, M.L., Charlson, R.L., Waggoner, A.P., Ogren, J.A., Caldow, R., Holm, R.L., Quant, F.R., Sem, G.J., Wiedensohler, A., Ahlquist, N.A., and Bates, T.S. (1996), Performance characteristics of a high-sensitivity, three-wavelength total scatter/backscatter nephelometer, Journal of Atmospheric and Oceanic Technology, 13, 967-986.

Andrews, E., Ogren, J. A., Kinne, S., and Samset, B.: Comparison of AOD, AAOD and column single scattering albedo from AERONET retrievals and in situ profiling measurements, Atmos. Chem. Phys., 17, 6041–6072, https://doi.org/10.5194/acp-17-6041-2017, 2017.

Baray, S., Darlington, A., Gordon, M., Hayden, K. L., Leithead, A., Li, S.-M., Liu, P. S. K., Mittermeier, R. L., Moussa, S. G., O'Brien, J., Staebler, R., Wolde, M., Worthy, D., and McLaren, R.: Quantification of methane sources in the Athabasca Oil Sands Region of Alberta by aircraft mass balance, Atmos. Chem. Phys., 18, 7361–7378, https://doi.org/10.5194/acp-18-7361-2018, 2018.

Bodhaine, B.A., Ahlquist, N.C., and Schnell, R.C. (1991). Three-wavelength nephelometer suitable for aircraft measurement of background aerosol scattering coefficient, Atmospheric Environment, 25A, 2267-2276

Boucher, O., Randall, P., Artaxo, C., Bretherton, G., Feingold, P., Forester, V.-M., et al. (2013). 2013: Clouds and aerosols, Climate change2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 571–658, Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.016.

Bytnerowicz, A., Fraczek,W., Schilling, S., and Alexander, D.: Spatial and temporal distribution of ambient nitric acid and ammonia in the Athabasca Oil Sands Region, Alberta, J. Limnol., 69, 11–21, 2010.

Cai, Y., D. Montague, W. Mooiweer-Bryan, and T. Deshler, "Performance characteristics of the ultra high sensitivity aerosol spectrometer for particles between 55 and 800 nm: Laboratory and field studies," Journal of Aerosol Science 39 (2008) 759-769.

Charlson, R.J., Valero, F.P.J. and Seinfeld, J.H.: In search of balance, Science, 308, 806–807, https://doi.org/10.1126/science.1108162, 2005.

Charpentier, A. and Bergerson, J. M.: Understanding the Canadian oil sands industry's greenhouse gas emissions, Environ. Res. Lett., 4, 014005, https://doi.org/10.1088/1748-9326/4/1/014005, 2009.

Cheng, Y., Li, S.-M., Gordon, M., and Liu, P.: Size distribution and coating thickness of black carbon from the Canadian oil sands operations, Atmos. Chem. Phys., 18, 2653–2667, https://doi.org/10.5194/acp-18-2653-2018, 2018.

Chin, M., Diehl, T., Ginoux, P., and Malm, W.: Intercontinental transport of pollution and dust aerosols: implications for regional air quality, Atmos. Chem. Phys., 7, 5501–5517, https://doi.org/10.5194/acp-7-5501-2007, 2007.

Dunagan, S.E., Johnson, R.; Zavaleta, J.; Russell, P.B.; Schmid, B.; Flynn, C.; Redemann, J.; Shinozuka, Y.; Livingston, J.; Segal-Rosenhaimer, M. Spectrometer for Sky-Scanning Sun-Tracking Atmospheric Research (4STAR): Instrument Technology. Remote Sens. 2013, 5, 3872-3895.

DeCarlo, P. F. et al. Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. Anal. Chem. 78, 8281–8289 (2006).

Emde, C., Buras-schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., Kylling, J., Richter, B., Pause, C., Dowling, T. and Bugliaro, L.: The libRadtran software package for radiative transfer calculations (version 2 . 0 . 1), Geosci. Model Dev., 9, 1647–1672, doi:10.5194/gmd-9-1647-2016, 2016.

Giles D.M., Sinyuk A., Sorokin M.G., Schafer J.S., Smirnov A., Slutsker I., Eck T.F., Holben B.N., Lewis J.R., Campbell J.R., Welton E.J., Advancements in the Aerosol Robotic Network (AERONET) Version 3 database–automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. Atmospheric Measurement Techniques. 2019 Jan 1;12(1).

Gordon, M., Li, S.-M., Staebler, R., Darlington, A., Hayden, K., O'Brien, J., and Wolde, M.: Determining air pollutant emission rates based on mass balance using airborne measurement data over the Alberta oil sands operations, Atmos. Meas. Tech., 8, 3745–3765, https://doi.org/10.5194/amt-8-3745-2015, 2015.

Gu, X., Bao, F., Cheng, T. *et al.* The impacts of regional transport and meteorological factors on aerosol optical depth over Beijing, 1980–2014. *Sci Rep* **8,** 5113 (2018). https://doi.org/10.1038/s41598-018-22803-x

Heald, C. L., Jacob, D. J., Park, R. J., Alexander, B., Fairlie, T. D., Yantosca, R. M., and Chu, D. A. (2006), Transpacific transport of Asian anthropogenic aerosols and its impact on surface air quality in the United States, *J. Geophys. Res.*, 111, D14310, doi:10.1029/2005JD006847.

Holben, B. N., T. F. Eck, I. Slutsker, D. Tanré, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan, Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov (1998), AERONET—a federated instrument network and data archive for aerosol characterization, Remote Sensing of Environment, 66 (1), 1–16.

Holben, B. N., Kim, J., Sano, I., Mukai, S., Eck, T. F., Giles, D. M., Schafer, J. S., Sinyuk, A., Slutsker, I., Smirnov, A., Sorokin, M., Anderson, B. E., Che, H., Choi, M., Crawford, J. H., Ferrare, R. A., Garay, M. J., Jeong, U., Kim, M., Kim, W., Knox, N., Li, Z., Lim, H. S., Liu, Y., Maring, H., Nakata, M., Pickering, K. E., Piketh, S., Redemann, J., Reid, J. S., Salinas, S., Seo, S., Tan, F., Tripathi, S. N., Toon, O. B., and Xiao, Q.: An overview of mesoscale aerosol processes, comparisons, and validation studies from DRAGON networks, Atmos. Chem. Phys., 18, 655–671, https://doi.org/10.5194/acp-18-655-2018, 2018.

Kahn, R.A. Reducing the Uncertainties in Direct Aerosol Radiative Forcing. *Surv Geophys* **33,** 701–721 (2012). https://doi.org/10.1007/s10712-011-9153-z

Kaufman, Y., Tanré, D. & Boucher, O. A satellite view of aerosols in the climate system. *Nature* **419,** 215–223 (2002). https://doi.org/10.1038/nature01091

Kupc, A., Williamson, C., Wagner, N. L., Richardson, M., and Brock, C. A.: Modification, calibration, and performance of the Ultra-High Sensitivity Aerosol Spectrometer for particle size distribution and volatility measurements during the Atmospheric Tomography Mission (ATom) airborne campaign, Atmos. Meas. Tech., 11, 369-383, https://doi.org/10.5194/amt-11-369-2018, 2018.

Landis, M.S, J.P. Pancras, J.R. Graney, E.M. White, E.S. Edgerton, A. Legge, K.E. Percy, 2017, Source apportionment of ambient fine and coarse particulate matter at the Fort McKay community site, in the Athabasca Oil Sands Region, Alberta, Canada, Sci. Total Environ., 584–585 (2017), pp. 105-117.

Landis, M.S.; Edgerton, E.S.; White, E.M.; Wentworth, G.R.; Sullivan, A.P.; Dillner, A.M. The impact of the 2016 Fort McMurray Horse River Wildfire on ambient air pollution levels in the Athabasca Oil Sands Region, Alberta, Canada, (2018), Sci. Total Environ. 618, 1665–1676.

LeBlanc, S. E., Redemann, J., Flynn, C., Pistone, K., Kacenelenbogen, M., Segal-rosenheimer, M., Shinozuka, Y., Dunagan, S., Dahlgren, R. P., Meyer, K., Podolske, J., Howell, S. G., Freitag, S., Small-griswold, J., Holben, B., Diamond, M., Wood, R., Formenti, P., Piketh, S., Maggs-Kölling, G., Gerber, M. and Namwoonde, A.: Above-cloud aerosol optical depth from airborne observations in the southeast Atlantic, Atmos. Chem. Phys., 20, 1565–1590, doi:10.5194/acp-20-1565-2020, 2020

Li et al, 2017, Differences between measured and reported volatile organic compound emissions from oil sands facilities in Alberta, Canada, Proceedings of the National Academy of Sciences May 2017, 114 (19) E3756-E3765; DOI: 10.1073/pnas.1617862114

Liggio J, et al. (2016) Oil sands operations as a large source of secondary organic aerosols. Nature 534:91–94.

Liggio, J., Li, S., Staebler, R.M. *et al.* Measured Canadian oil sands CO_2 emissions are higher than estimates made using internationally recommended methods. *Nat Commun* **10,** 1863 (2019). https://doi.org/10.1038/s41467-019-09714-9

Levy, R., Hsu, C., et al., 2015. MODIS Atmosphere L2 Aerosol Product. NASA MODIS Adaptive Processing System, Goddard Space Flight Center, USA: http://dx.doi.org/10.5067/MODIS/MOD04_L2.061

Matus, A., L'Ecuyer, T. and Henderson, D. (2019), New estimates of aerosol direct radiative effects and forcing from A-Train satellite observations. Geophysical Research Letters, 46, 8338–8346, https://doi.org/10.1029/2019GL083656

McComiskey, A. and Feingold, G.: The scale problem in quantifying aerosol indirect effects, Atmos. Chem. Phys., 12, 1031–1049, https://doi.org/10.5194/acp-12-1031-2012, 2012.

McKendry, I.G., Strawbridge, K.B., O'Neill, N.T., Macdonald, A.M., Liu, P.S.K., Leaitch, W.R., Anlauf, K.G., Jaegle, L., Fairlie, T.D., Westphal, D.L., 2007. Trans-Pacific transport of Saharan dust to western North America: a case study. J. Geophys. Res. 112, D01103. http://dx.doi.org/10.1029/2006JD007129.

NRCAN (National Resources Canada), 2020, Energy Fact Book 2019-20, accessible at: https://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/energy/energy-factbook_EN-feb14-2020.pdf

O'Neill, N. T., Eck, T. F., Smirnov, A., Holben, B. N., and Thulasiraman,S.: Spectral discrimination of coarse and fine mode optical depth, J. Geophys. Res.-Atmos., 108, AAC8.1–AAC8.15, oi:10.1029/2002JD002975, 2003.

O'Neill, N. T., T. Eck, A. Smirnov, B. Holben, Spectral Deconvolution Algorithm (SDA) Technical memo, AERONET Technical report, NASA-GSFC, 2008

O'Neill, N. T., Baibakov, K., Hesaraki, S., Ivanescu, L., Martin, R. V., Perro, C., Chaubey, J. P., Herber, A., and Duck, T. J.: Temporal and spectral cloud screening of polar winter aerosol optical depth (AOD): impact of homogeneous and inhomogeneous clouds and crystal layers on climatological-scale AODs, Atmos. Chem. Phys., 16, 12753–12765, https://doi.org/10.5194/acp-16-12753-2016, 2016.

Ogren, J.A., Wendell, J., Andrews, E., and Sheridan, P.J.: Continuous light absorption photometer for long-term studies, Atmos. Meas. Tech., 10, 4805-4818, 2017.

Pistone, K., Redemann, J., Doherty, S., Zuidema, P., Burton, S., Cairns, B., Cochrane, S., Ferrare, R., Flynn, C., Freitag, S., Howell, S. G., Kacenelenbogen, M., LeBlanc, S., Liu, X., Schmidt, K. S., Sedlacek III, A. J., Segal-Rozenhaimer, M., Shinozuka, Y., Stamnes, S., van Diedenhoven, B., Van Harten, G., and Xu, F.: Intercomparison of biomass burning aerosol optical properties from in situ and remote-sensing instruments in ORACLES-2016, Atmos. Chem. Phys., 19, 9181–9208, https://doi.org/10.5194/acp-19-9181-2019, 2019.

Qin, K.; Wang, L.; Xu, J.; Letu, H.; Zhang, K.; Li, D.; Zou, J.; Fan, W. Haze Optical Properties from Long-Term Ground-Based Remote Sensing over Beijing and Xuzhou, China. *Remote Sens.* **2018**, *10*, 518.

Schwarz, J. P., et al. (2006), Single-particle measurements of midlatitude black carbon and light-scattering aerosols from the boundary layer to the lower stratosphere, *J. Geophys. Res.*, 111, D16207, doi:10.1029/2006JD007076.

Shinozuka, Y., et al (2011), Airborne observation of aerosol optical depth during ARCTAS: vertical profiles, inter-comparison and fine-mode fraction, Atmos. Chem. Phys., 11, 3673-3688, https://doi.org/10.5194/acp-11-3673-2011.

Shinozuka, Y. and Redemann, J.: Horizontal variability of aerosol optical depth observed during the ARCTAS airborne experiment, Atmos Chem Phys., 11(October 2002), 8489–8495, doi:10.5194/acp-11-8489-2011, 2011.

Shinozuka, Y., et al. (2013), Hyperspectral aerosol optical depths from TCAP flights, J. Geophys. Res. Atmos., 118, 12,180–12,194, doi:10.1002/2013JD020596.

Sioris, C. E., Abboud, I., Fioletov, V. E., McLinden C.A. (2017a): AEROCAN, the Canadian sub-network of AERONET: aerosol monitoring and air quality applications. Atmos. Environ. 167:444–457. https://doi.org/10.1016/J.ATMOSENV.2017.08.044

Sioris, C. E., McLinden, C. A., Shephard, M. W., Fioletov, V. E., and Abboud, I. (2017b): Assessment of the aerosol optical depths measured by satellite-based passive remote sensors in the Alberta oil sands region, Atmos. Chem. Phys., 17, 1931-1943, https://doi.org/10.5194/acp-17-1931-2017.

Stone, R. S., et al. (2010), A three-dimensional characterization of Arctic aerosols from airborne Sun photometer observations: PAM-ARCMIP, April 2009, J. Geophys. Res., 115, D13203, doi:10.1029/2009JD013605.

Sun, T., Che, H., Qi, B., Wang, Y., Dong, Y., Xia, X., Wang, H., Gui, K., Zheng, Y., Zhao, H., Ma, Q., Du, R., and Zhang, X.: Aerosol optical characteristics and their vertical distributions under enhanced haze pollution events: effect of the regional transport of different aerosol types over eastern China, Atmos. Chem. Phys., 18, 2949–2971, https://doi.org/10.5194/acp-18-2949-2018, 2018.

Thermo Scientific (2015a): Model 42i Chemiluminescence NO-NO2-NOx Analyzer Part Number 101350-00, Instruction Manual, 347 p.

Thermo Scientific (2015b): Model 43i Instruction Manual Pulsed Fluorescence SO2 Analyzer Part Number 101589-00, Instruction Manual, 316 p.

Toledano, C., Cachorro, V.E., Berjon, A., de Frutos, A.M., Sorribas, M., de la Morena, B.A. and Goloub, P. (2007), Aerosol optical depth and Ångström exponent climatology at El Arenosillo AERONET site (Huelva, Spain). Q.J.R. Meteorol. Soc., 133: 795-807. doi:10.1002/qj.54

Wilcox, E. M.: Stratocumulus cloud thickening beneath layers of absorbing smoke aerosol, Atmos. Chem. Phys., 10, 11769–11777, https://doi.org/10.5194/acp-10-11769-2010, 2010.
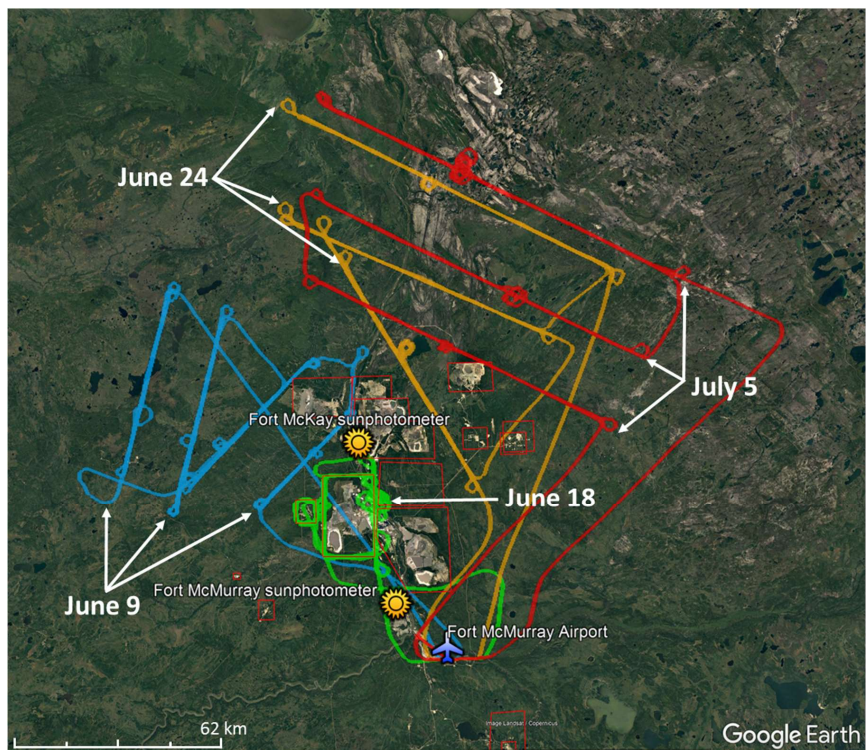
Figure 1: Selected flight tracks during the Oil Sands Measurement Campaign. Also shown are locations of AOSR facilities (red outlined boxes) and Fort McMurray and Fort McKay AERONET stations overlaid on satellite imagery (image source: Landsat/Copernicus as used by © Google Earth).
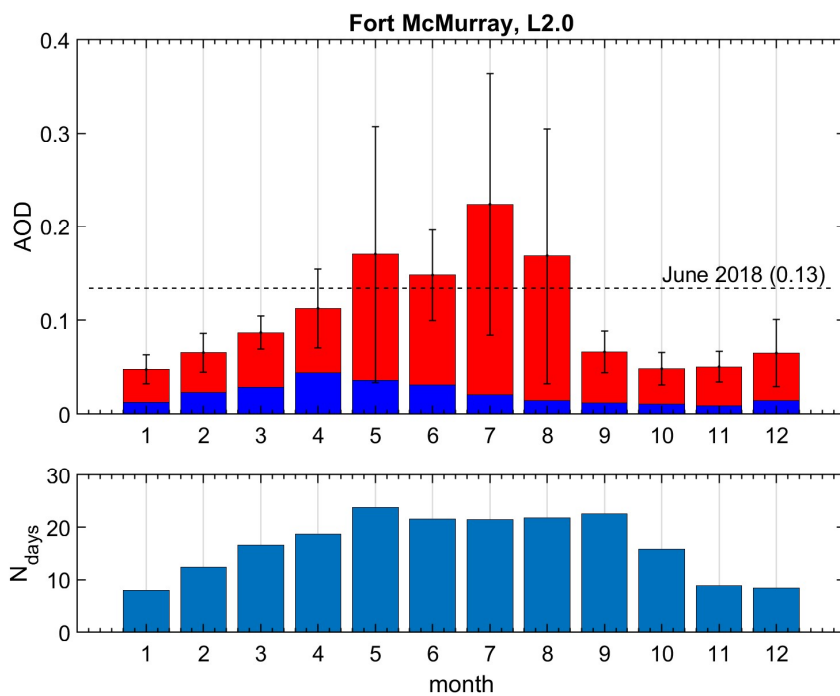
Figure 2: Monthly fine- (red) and coarse-mode (blue) AOD averages for the Fort McMurray AERONET station. The dashed line shows the June 2018 mean total AOD. The error bars represent standard deviations of the (total) monthly mean AODs. The bottom graph shows the mean number of days associated with each monthly-binned average. These statistics are for AERONET V3, level 2.0, extinction AODs averaged over the data range from 2005 to 2018 inclusively.
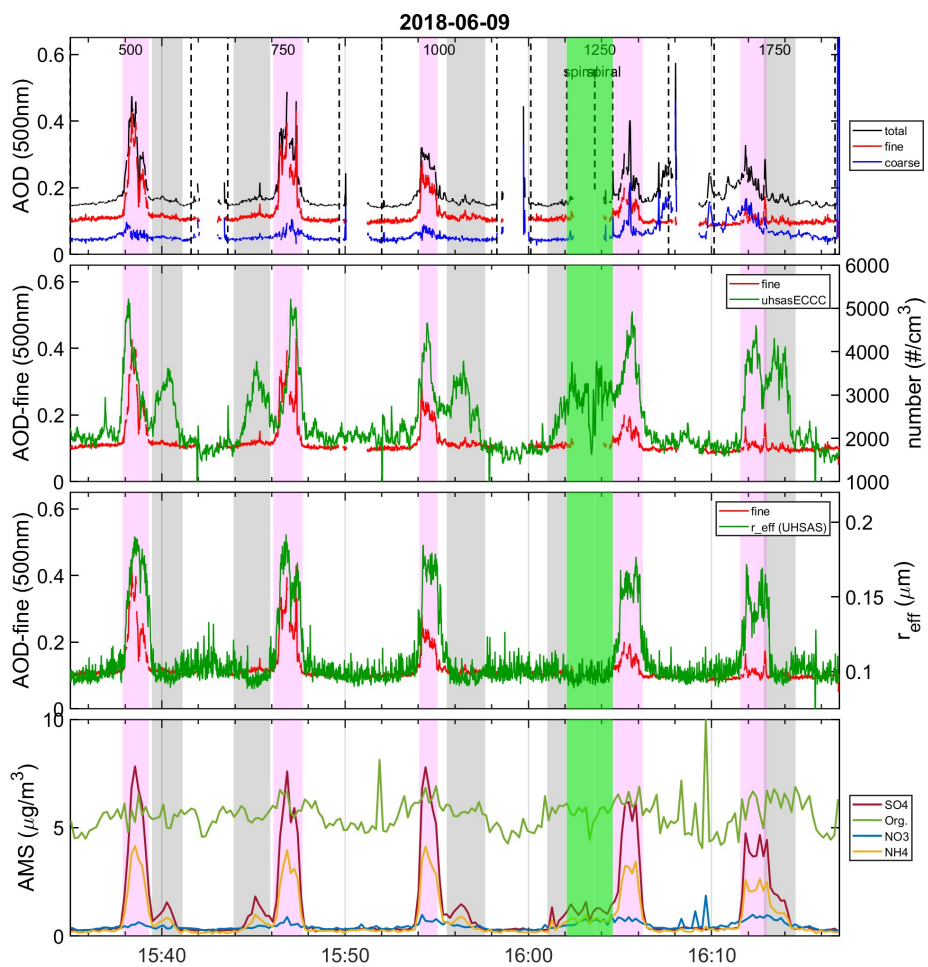
Figure 3: Screen 1 time series on June 9 (UTC). From top to bottom: a) total, fine and coarse mode AODs at 500 nm derived from the SDA of the 4STAR data; b) UHSAS total particle number concentration (all sizes), $^\#/cm^3$, c) UHSAS effective radius, μm, d) AMS particle composition. In the top panel, the dashed vertical lines delineate each altitude level with nominal (planned) altitudes listed in feet above ground level at the top. The fine-mode AOD is reproduced on several plots to facilitate the comparisons. The pink and grey rectangles represent periods associated with plumes "A" and "B", respectively, while the green shading indicates periods when the aircraft was engaged in a spiral manoeuvre.
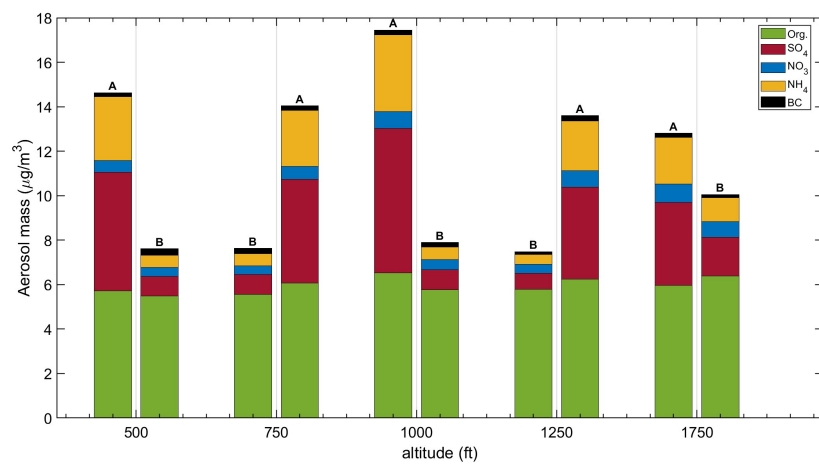
Figure 4: Particle chemical composition for plumes A and B during screen 1 of the June 9 transformation flight.
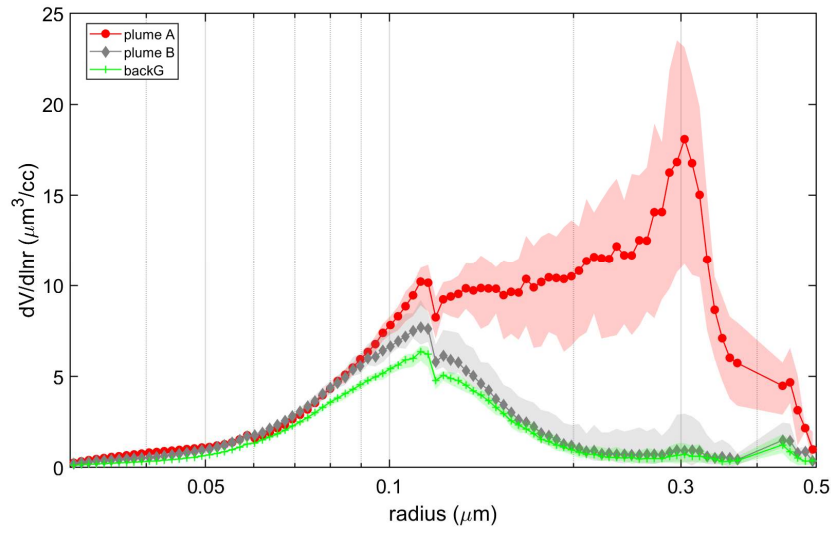
Figure 5: Averaged UHSAS volume PSDs during screen 1 on June 9. Color patches represent variations with altitude while solid lines show the screen-averaged PSDs for plumes A and B as well as the background (out-of-plume) signal.
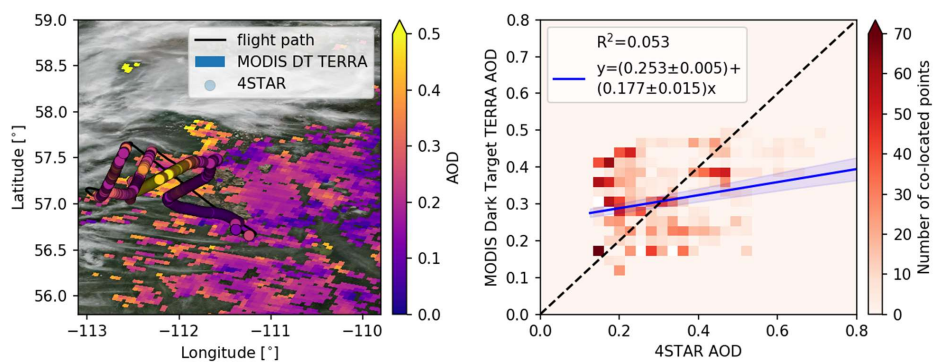
Figure 6. Comparison of AOD measured by 4STAR and retrieved by MODIS Dark Target (DT, MOD04_3K; Levy et al., 2015) for June 9th, 2018. (Left) The aerosol scene is overlaid on the True Color imagery from MODIS TERRA, with the colorbar showing the AOD sampled by MODIS DT (in rectangular pixels) and 4STAR (circles put along the flight track of the NRC Convair-580). (Right) The pixel-by-pixel comparison of the MODIS DT AOD and 4STAR AOD, with a best fit line in blue, and 95% confidence intervals of the fit line in blue shade.
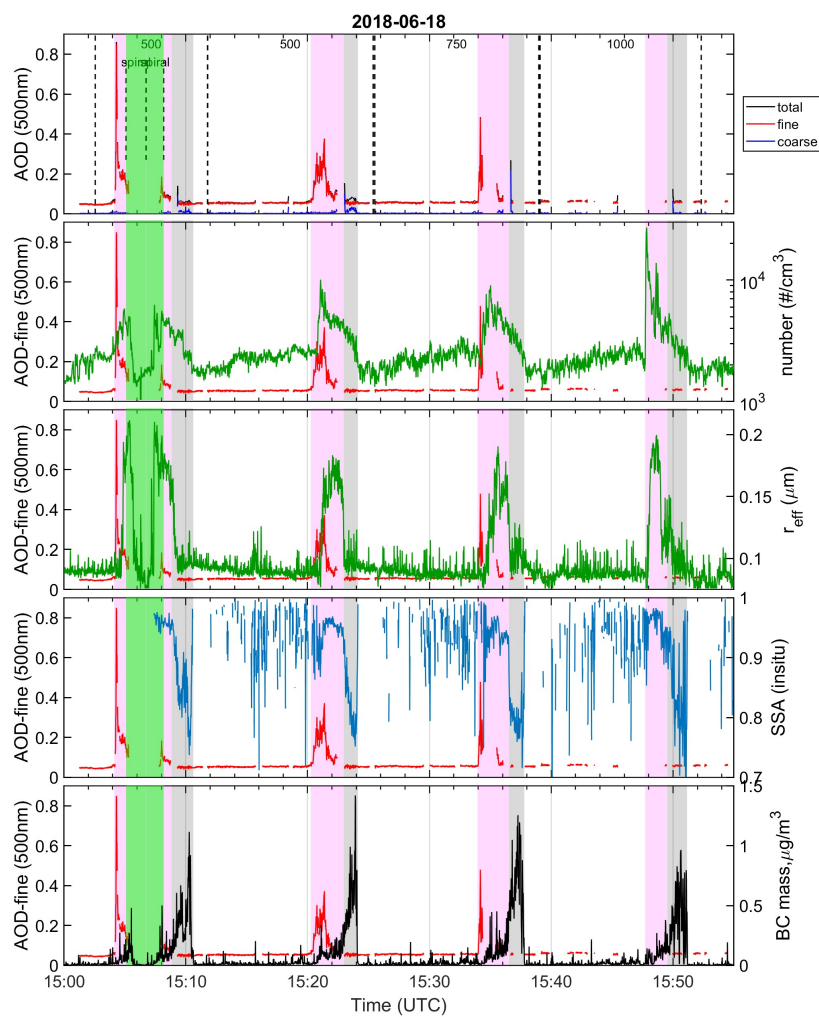
Figure 67: Top three panes as in Figure 3. Panes 4: SSA440 from nephelometer and CLAP, unitless, pane 5: black carbon mass, $\mu g/m^3$. The pink and grey rectangles represent plumes A and B respectively as discussed in the text. UTC time.
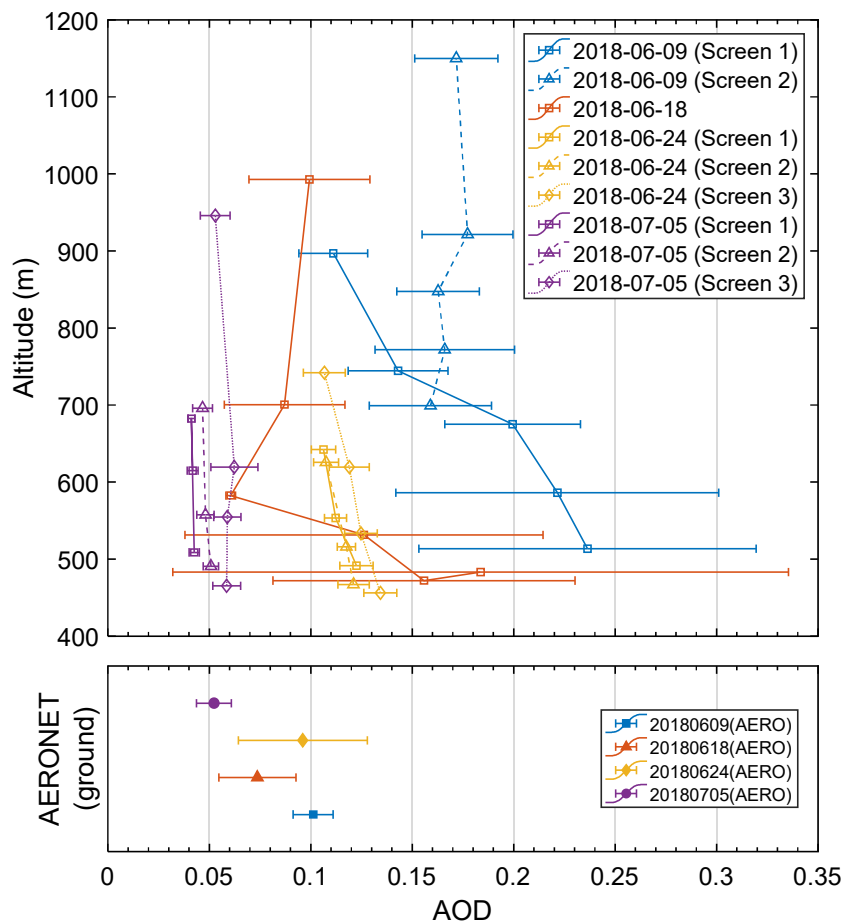
Figure 7~~7~~8: Top: 4STAR in plume mean fine-mode AODs for transformation flights of June 9 June 24, and July 5 and an emission flight of June 18. Bottom: Fort McMurray ground-based AOD daily averages and standard deviations for selected dates (staggered vertically to show the error bars). ~~In both panels T~~the horizontal error bars indicate the AOD standard deviation representing spatial variability of the plumes (per altitude std values for 4STAR and daily std values for AERONET). For June 24 the AERONET data after 18:13 was removed because it was associated with forest fires.
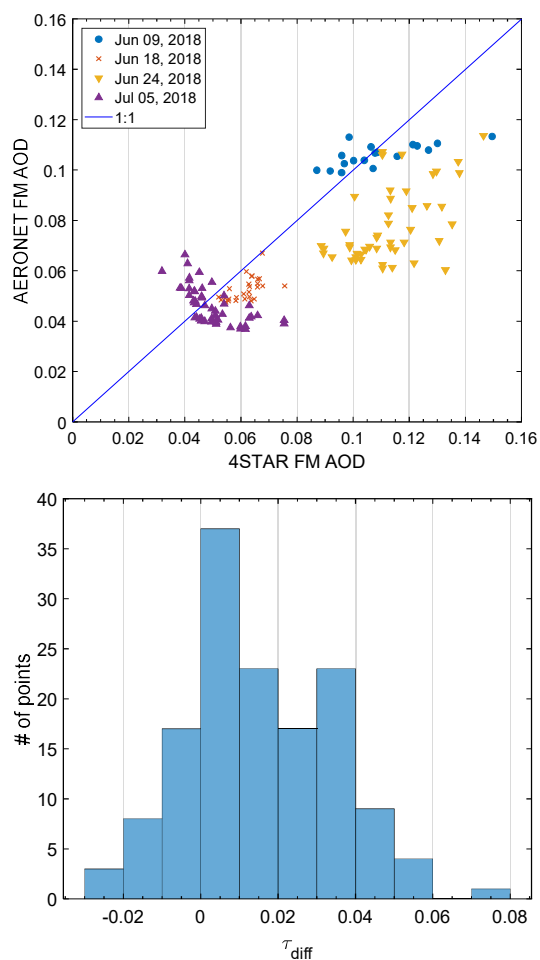
Figure 89: Top: AOD fine-mode comparisons between 4STAR and AERONET for the selected flights of June 9, 18, 24, and July 5. Bottom: corresponding histogram of fine-mode AOD differences between the 4STAR and AERONET. AERONET data was matched with 4STAR within 10 minutes. 4STAR data includes only out-of-plume AODs acquired during a straight-level flight (i.e. excluding spirals) at an altitude of less than 1 km. For June 24 the AERONET data after 18:13 was removed because it was associated with forest fires.
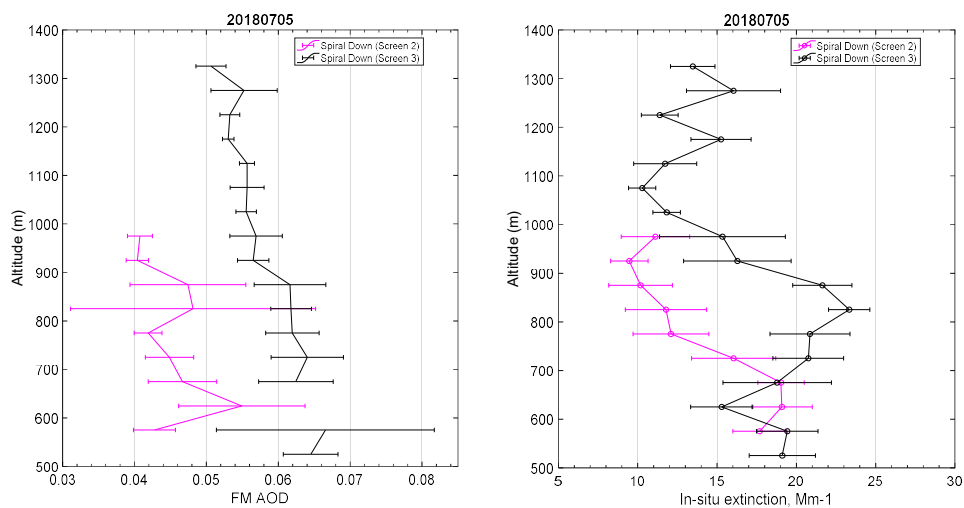
Figure 910. In-plume binned profile data obtained on July 5 during screens 2 and 3. The error bars represent standard deviations within each 50 m bin. Left: 4STAR FM AOD, right: calculated in-situ extinction coefficient at 500 nm. The in-situ layer AOD between 575 and 975 m has increased by 0.002 (~33%) from screen 2 to screen 3.

Table 1. FM AOD statistics for selected flights during the OSMC (mean ± std). 4STAR in-plume and out-of-plume statistics are reported for the lowest flight level.

Date (Flight #)	4STAR in-plume average	4STAR in-plume maximum	4STAR out-of-plume average	AERONET daily average	AERONET daily max
June 9 (F14)				0.10 ± 0.01	0.12
scr. 1	0.24 ± 0.08	0.43	0.11 ± 0.00		
scr. 2	0.16 ± 0.03	0.23	0.11 ± 0.01		
June 18 (F19)	0.18 ± 0.15	0.85	0.05 ± 0.00	0.07 ± 0.02	0.12
June 24 (F24)				0.13 ± 0.08	0.37
scr. 1	0.12 ± 0.01	0.15	0.13 ± 0.01	0.10 ± 0.03*	0.18*
scr. 2	0.12 ± 0.01	0.14	0.11 ± 0.01		
scr. 3	0.13 ± 0.01	0.15	0.12 ± 0.00		
July 5 (F30)				0.05 ± 0.01	0.08
scr. 1	0.04 ± 0.00	0.05	0.04 ± 0.00		
scr. 2	0.05 ± 0.00	0.06	0.05 ± 0.01		
scr. 3	0.06 ± 0.01	0.07	0.05 ± 0.01		

5   * with data from forest fire plumes excluded (~27% of data)