Mineral dust aerosol sources in northern Canada:

An investigation using the TROPOMI instrument

By
Kagan Akiyama

A Thesis Submitted to
Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for
the Degree of Bachelor of Science Honours in Environmental Science

April, 2024, Halifax, Nova Scotia

© Kagan Akiyama, 2024

Approved:	Dr. Aldona Wiacek Supervisor
Approved:	Dr. Ian Ashpole Supervisor
Approved:	Dr. Thomas Duck Reader
Date:	April 30, 2024

Mineral dust aerosol sources in northern Canada:

An investigation using the TROPOMI instrument

By Kagan Akiyama

Abstract

Mineral dust aerosols (MDAs) are important for climate regulation, but little is known about their sources in high-latitude regions like northern Canada. This thesis investigates the atmospheric distribution and surface sources of MDAs in northern Canada using the satellite-based Tropospheric Monitoring Instrument (TROPOMI). TROPOMI Absorbing Aerosol Index (AAI) and Aerosol Layer Height (ALH) data were used, together with the TROPOMI cloud cover (CLO) and carbon monoxide (CO) products. (Where CO and CLO are low, spurious impacts on AAI due to non-MDA atmospheric components are minimized.) The three main research questions were: (1) What is the distribution of dust in northern Canada, according to TROPOMI? and (1b) What does ALH data imply about the sources of the observed dust distributions? (2) How do TROPOMI-derived aerosol products compare to those of other well-characterized satellite instruments, like MODIS and OMI? (3) What are the effects of filtering the TROPOMI aerosol products by the carbon monoxide and cloud products? It was found that frequent observations of AAI values greater than 0.5 were made by TROPOMI in high-latitude regions in 2021, suggesting the presence of absorbing aerosols, likely MDAs. Filtering AAI data by ALH isolated likely local MDA sources, previously uncharacterized in field observations. It was also found that TROPOMI AAI data broadly agrees with MODIS Aerosol Optical Depth (AOD) and OMI AAI data, especially when filtered for positive TROPOMI AAI values and statistically significant correlations. Unexpectedly, filtering AAI data by low CO or CLO did not improve the correlation between TROPOMI and MODIS data, but it also did not eliminate the dust sources isolated by ALH filtering. The findings suggest the presence of significant sources of MDAs in northern Canada, but more observational research is needed to confirm their existence and strength, and thus their impact on climate.

April 30th, 2024

Acknowledgements

Firstly, I would like to thank my supervisors Dr. Aldona Wiacek and Dr. Ian Ashpole. Without their constant support and guidance, this work would have never been accomplished. I am grateful to have had the opportunity to learn from them and to have enriched my education with such a unique research project.

With that, I would like to extend my thanks to the WARG research team, including Martin Hellmich, Kyle Yeates, Cameron Power, Lukas Donovan, and Hailey Wigmore, who have shown me kindness and friendship, and extended their own guidance throughout my time with the WARG team. A special thanks to Kyle Yeates for teaching me the basics of Python programming and sharing his code to aid in the preliminary analysis of this work's data.

I would also like to thank the staff at Saint Mary's University who have provided me with the foundation to take on this project. I would also like to thank Dr. Thomas Duck from Dalhousie University for being my reader for this project and for teaching me the importance of radiative transfer.

Lastly, I would like to thank my friends and family for supporting me in times when I needed it the most. Their constant confidence in my abilities and interest in my work kept me motivated when I could not do it myself.

List of Figures

List of Tables

Glossary of Terms and Abbreviations

- AAI: Absorbing Aerosol Index
- ALH: Aerosol Layer Height
- ATBD: Algorithm Theoretical Basis Document
- CLO: CLOud fraction
- CO: Carbon Monoxide
- DOD: Dust Optical Depth
- FoO: Frequency of Occurrence
- Gridbox: The spatial area of one pixel (area varies depending on the spatial resolution chosen) on a map made up of many pixels
- MODIS: MODerate resolution Imaging Spectrometer
- NASA: National Aeronautics and Space Administration
- NaN: Not a Number
- n_days: Number of days (with daily mean observations); interchangeable with n_obs
- n_obs: Number of observations (that are daily means); interchangeable with n_days
- OMI: Ozone Monitoring Instrument
- PUM: Product User Manual
- TROPOMI: TROPOspheric Monitoring Instrument
- UV: UltraViolet
- UVAI: UltraViolet Aerosol Index (similar to AAI)

Table of Contents

1. Introduction

1.1 Mineral Dust Aerosols (MDAs)

Mineral dust aerosols (MDAs) play an important role in global climate patterns (Bullard, et al., 2016). MDAs are suspended particles with a diameter of less than 10 μm that are comprised of windblown dust originating from the Earth's surface (Meinander, et al., 2022). They are strongly absorbing of light in the UV range (100 – 400 nm) and they absorb some light in the visible range (400 – 700 nm), but this absorption becomes nearly negligible at wavelengths greater than 600 nm (Bergstrom, et al., 2007). Another important characteristic of MDAs is that they are comprised of particles that are emitted primarily from surfaces in arid and semi-arid regions (Bullard, et al., 2016; Knippertz & Stuut, 2014; Meinander, et al., 2022). Scattering aerosols, like sulfate particles, disperse solar light and usually have a cooling effect on the climate. Absorbing aerosols, such as smoke from biomass burning, desert dust, volcanic ash, and anthropogenically produced soot, absorb radiation and have a warming effect on the climate (Boucher, 2015). MDAs, in general, play an important role in weather and air quality, marine life, climate, and health (Meinander, et al., 2022). MDAs can impact Earth's incoming and outgoing energy, clouds, precipitation, biogeochemistry, and air quality. For example, when dust is deposited on snow and ice it reduces the surface albedo thus increasing the melting rate (Williamson & Menounos, 2021).

Mid-latitude deserts, such as the Gobi Desert or Taklamakan Desert, have long been recognized as MDA source hotspots (e.g., Wiacek et al., 2010), but emissions from high-latitude sources are expected to increase with climate change (Bullard, et al., 2016). This is partially why there is limited research in high latitudes: since subtropical and mid-latitude deserts are the known major global dust sources, they have received more research focus than the potential newer sources

arising in Canada's north. Climate change is also likely to decrease snow cover and retreat glaciers, thus allowing more surface crustal elements to become suspended by wind and therefore increase the probability of dust storms (Meinander, et al., 2022). MDAs from high-latitude sources are especially important in the Arctic, with around one-third of the dust atmospheric load and 90% of deposition to surface thought to have been from local sources (Groot Zwaaftink, et al., 2016).

MDA importance in the Arctic is linked to the ice-albedo feedback effect. This effect describes the phenomenon where a decrease in snow and ice leads to a decrease of albedo, and then a decrease in albedo leads to a decrease of snow and ice, thus continuing the loop (Yong-Sang, et al., 2020). The reason a decreased albedo increases the melting of ice is that those snow-free areas are now less able to reflect sunlight and therefore increase the temperature of the area since more of the radiation from the Sun is being absorbed, thus increasing the melting. When MDAs are deposited onto surfaces with high albedo, such as snow, the ability of said surface to reflect sunlight is reduced (Williamson & Menous, 2021). In areas like northern Canada where wind speeds are high, this can become a problem as MDAs may be picked up from local sources and deposited onto nearby snowy surfaces (Bullard, et al., 2016). The MDAs reducing the snow albedo have negative impacts on the climate, i.e., accelerated warming as compared to mid-latitudes.

In high latitudes, MDAs add another layer of complexity to the relationship between clouds and climate. MDAs act as cloud condensation nuclei (Garimella, et al., 2014; Shikwambana, 2022). These microscopic particles can provide surfaces for water vapor to condense around, potentially leading to increased cloud formation (Garmiella, et al., 2014). The type of clouds that form, however, depends on the specific type of aerosols present (Garimella, et al., 2014). While increased cloud cover can reflect incoming solar radiation, exerting a cooling effect (Kooreman,

et al., 2020), the presence of clouds can also pose challenges for studying MDAs themselves. Clouds can potentially block aerosol plumes from satellite instruments, making it difficult to track and measure MDAs (Boucher, 2015; Bullard, et al., 2016). Additionally, clouds can be mistaken for other highly reflective surfaces, like snow and ice, further reducing the reliability of space-based measurements of MDAs. Therefore, understanding how MDA detection is influenced by clouds is equally as important as understanding the influence that MDAs have on cloud formation.

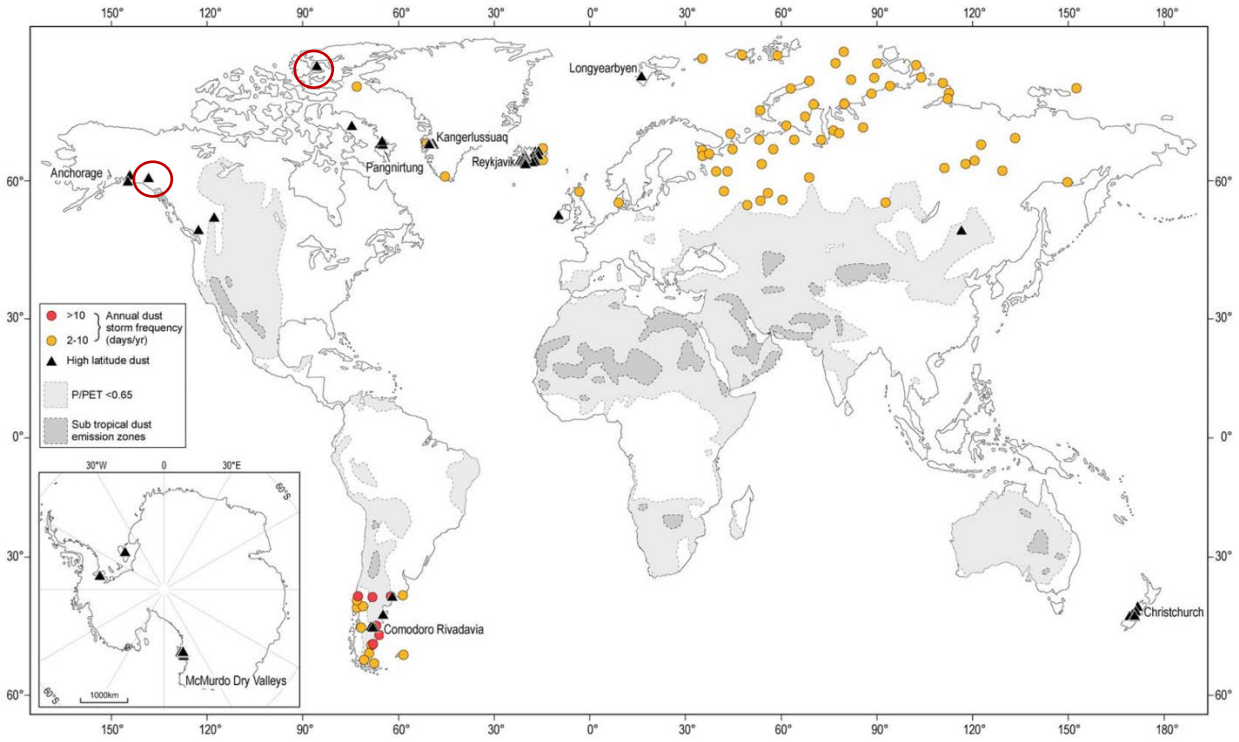

Figure 1: Global observations of high-latitude dust. Black triangles indicate known local sources of MDA emissions. P/PET (potential evapotranspiration ratio) and subtropical dust emission zones are also included for reference. Red circles highlight the Kluane Lake area (Yukon) and the Lake Hazen area (Ellesmere Island) (Adapted from Bullard, et al., 2016).

To date, there has been little research focusing on high-latitude MDA sources, with two well-studied source areas having been characterized in Canada's North. Figure 1 shows the global observations of high-latitude dust where the black triangles indicate known, published observations of dust sources. In this figure, it can be seen that there are two recognized dust sources around (60N, 150W) and (80N, 90W), circled in red. These two sources are found in the Kluane

Lake region in the Yukon Territory, Canada and the Lake Hazen region in Nunavut, Canada. These are not the only sources in northern Canada, but they are the two that are the most well studied, and there are likely even more sources than indicated on Figure 1 that have yet to be widely studied (Ranjbar, et al., 2021; Huck et al., 2023).

1.2 Remote Sensing

Remote sensing can be defined as the gathering and quantification of data concerning specific attributes of phenomena, objects, or materials using a recording device that does not make physical contact with the observed features (Khorram, et al., 2012). Remote sensing has been used across many fields in the past to understand different aspects of a study field. In agriculture, remote sensing is used to monitor crop health and assess agricultural productivity (Thenkabail, 2014). In forestry, remote sensing has been used to detect illegal logging and track deforestation patterns (Hansen, et al., 2013). In general, the benefit of remote sensing comes from its capacity for large-scale coverage and frequent data acquisition (Khorram, et al., 2012). Remote sensing also allows for the monitoring of remote areas that may be difficult to access. It therefore allows for imaging of areas where aerosols may be likely to be found, such as in northern Canada (Mushtaq, et al., 2023). This is not just the case in the north of Canada but is true for anywhere where it would be difficult to measure atmospheric properties from the ground. Dust sources are often linked to remote locations due to their typically dry nature. The prevailing high wind velocities in high latitudes can erode and loosen soils (Bullard, et al., 2016), therefore contributing to the presence of MDAs in the area (Meinander, et al., 2022). The benefit of remote sensing is that it can measure and image remote locations that might be dust sources that would otherwise be difficult to measure.

Up to this point, there has been limited research regarding MDA distribution in high latitudes, largely because monitoring and imaging the atmosphere in those regions is difficult

(Meinander, et al., 2022; Bullard, et al., 2016). One of these reasons is the fact that dust sources near the north are hard to measure, primarily since ground instruments are hard to place in such remote areas. Remote sensing therefore becomes a useful tool for understanding aerosol properties in these areas, but not without its own problems (Bullard, et al., 2016). One reason atmospheric monitoring in those areas is difficult is because of the increased cloud cover. Compared to other regions, high-latitudes tend to have a higher cloud-coverage which makes it difficult for satellite-based instruments to measure the atmospheric properties that lie below the clouds (Huck, et al., 2023). Another problem is that the reflectance due to suspended MDA in the atmospheric column is very small compared to that due to bright clouds or surface ice, making it harder to detect MDA in polar regions (Ma et al., 2023). There is also limited sunlight for a greater percentage of the year at high latitudes compared to other areas, making it difficult for instruments that require sunlight to make their measurements, like the TROPOspheric Monitoring Instrument (TROPOMI) (Zweers, et al., 2022; Lubin, et al., 2006). Problems also arise when attempting to detect MDAs in snow covered areas that are also covered by clouds (Ma et al., 2023). This problem comes from difficulties in distinguishing snow from very low clouds, due to their similar reflectance properties and color (Ma, et al., 2023). These challenges therefore lead to questions in study results derived from remote sensing, like how one would know if a spatial pattern of TROPOMI aerosol measurements are truly the result of MDAs rather than an artefact due to noise effects from surface ice, low cloud cover and/or other aerosol types.

1.2.1 TROPOMI

The TROPOspheric Monitoring Instrument (TROPOMI) was launched on October 13, 2017. TROPOMI is a spectrometer that conducts daily global observations of the crucial atmospheric constituents ozone, nitrogen dioxide, sulfur dioxide, carbon monoxide, methane, and

formaldehyde, as well as properties related to clouds and aerosols (Apituley, et al., 2022b). TROPOMI is a nadir viewing hyperspectral imager with four separate spectrometers that measure radiation in the UV-visible wavelength range (270 – 500 nm), the near infrared (710 – 770 nm) and the shortwave infrared (2314 – 2382 nm) (Babić, et al., 2022; Zweers, 2022). Each of the four spectrometers corresponds to the four wavelength ranges that TROPOMI measures: mediumwave ultraviolet (UV, 270 – 320 nm), longwave ultraviolet combined with visible (UVIS, 320 – 490 nm), near infrared (NIR, 710 - 775), and shortwave infrared (SWIR, 2305 – 2385 nm) (Babić, et al., 2022). Each of these spectrometers measure the solar radiation within its wavelength range that is primarily reflected by the Earth towards the top of the atmosphere (Zweers, 2022). The captured light then undergoes a series of calibrations and instrument processes to extract information about atmospheric properties that underlie the observed signal (Babić, et al., 2022).

The Copernicus Sentinel-5P satellite, aboard which the TROPOMI instrument is found, is in a near-polar orbit with an inclination of approximately 98.7 degrees (Zweers, 2022). The satellite orbits the Earth at an altitude of around 824 km and the orbital cycle is 16 days (14 orbits per day, 227 orbits per cycle). The orbit cycle is the time taken for the satellite to pass over the same geographical point on the ground.

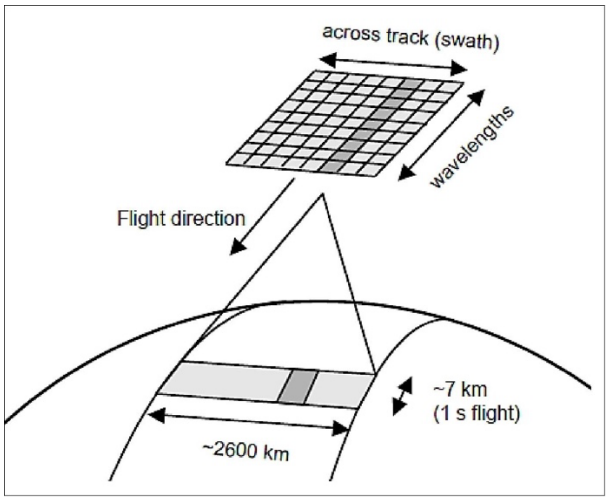

Figure 2: The measurement geometry of the TROPOMI instrument (Zweers, 2022).

TROPOMI observes the Earth's surface in a push-broom fashion, which means it scans across a wide strip of the Earth's atmosphere perpendicular to the satellite's orbital path (Zweers, 2022; Tang, et al., 2023; Trees, et al., 2021). The swath width of TROPOMI is relatively large, approximately 2600 km, which means it can observe a wide area along its orbital track. This measurement geometry can be seen in Figure 2. The instrument images a strip of the Earth for around 1 second. After the 1-second observation, a new measurement is made continuously in time to eventually achieve the swath pattern seen in Figure 2 (Apituley, et al., 2022b; Zweers, 2022). The typical pixel size of TROPOMI is therefore around 7 x 3.5 km^2, an improvement over its predecessor, OMI, with a spatial resolution of 13 x 24 km^2 (Zhao, et al., 2020). The increased spatial resolution of TROPOMI, compared to past satellite instruments, makes it a valuable asset for viewing and understanding the MDAs in northern Canada. With a higher spatial resolution, it is less likely for smaller aerosol plumes to get lost in spatial averaging. Also, the impacts of highly reflective surfaces, like clouds or snow and ice, or other confounding variables on satellite measurements can be minimized with improved spatial filtering in smaller pixels.

The TROPOMI instrument reflectance measurement is the basis for several atmospheric property datasets. For the purposes of this work, the primary dataset that will be used and analysed is the absorbing aerosol index (AAI), discussed below. Briefly, the AAI product is useful in indicating the absorption or scattering of aerosol loadings in certain areas, providing clues as to aerosol types. However, it is limited in that it is not *directly* sensitive to aerosol types, nor aerosol heights. So, in conjunction with the AAI dataset, the aerosol layer height (ALH), carbon monoxide (CO), and cloud (CLO) datasets (discussed below) will be used to better understand the observed aerosol distributions in the AAI dataset. With the use of all of these datasets in combination, it may be possible to attain a better understanding of the locations and abundances of aerosols in the northern parts of Canada.

1.2.1.1 TROPOMI Products

The ultraviolet aerosol index (UVAI), or absorbing aerosol index (AAI), indicates the presence of absorption in the atmosphere attributed to aerosols. The AAI was first introduced with the introduction of the Total Ozone Mapping Spectrometer (TOMS) in 1978 (Tang, et al., 2022), when it was first realized that absorbing aerosols were interfering with the planned and desired ozone measurements of TOMS. The AAI has since become a well-established product used by several satellite instruments (Zweers, 2022; Althaf, et al., 2022). It is calculated by separating the spectral contrast at two ultraviolet (UV) wavelengths caused by aerosol absorption from that of molecular Rayleigh scattering, surface reflection, and absorption by trace gases (de Graaf, et al., 2005; Torres et al., 1998). It is a unitless, numerical index, where higher values typically correspond to denser loadings of dust present in the atmosphere (Trees, et al., 2021). Theoretically, the AAI is zero if there are no absorbing or scattering aerosols present in the scene (Kooreman, et al., 2020). AAI can distinguish between absorbing and non-absorbing aerosols (Parya, et al., 2017).

For this research, the AAI is particularly useful as it was optimized to be a detection tool for dust aerosols (Tang, et al., 2022). AAI is also useful because it can be derived over any kind of surface (i.e. ground or water, and even over a cloud deck) and can therefore give a good idea of what absorbing aerosol distribution might look like in areas with variations in surface types, like areas with many islands, such as in northern Canada (Tang, et al., 2022).

The calculation of the AAI relies on measured and simulated reflectances at chosen wavelengths (de Graaf et al., 2005; Zweers et al., 2022). It is a measure of the spectral contrast at two wavelengths, λ_1 and λ_2, of the measured surface reflectance in the real atmosphere that contains UV absorbing aerosols (R^{meas}) and the modelled (i.e., simulated) surface reflectance that would result from an aerosol-free atmosphere (R^{sim}) (de Graaf, et al., 2005). The AAI calculation is formally defined as follows (Torres, et al., 1998):

$$AAI = -100 * [log_{10}(\frac{R_{\lambda 1}}{R_{\lambda 2}})^{meas} - log_{10}(\frac{R_{\lambda 1}}{R_{\lambda 2}})^{sim}] \quad (1)$$

The simulated surface albedo (A_s) at λ_1 can be deliberately chosen so that the simulated reflectance is equal to the measured reflectance at this wavelength, which is what is done in the case of the TROPOMI AAI calculation (de Graaf, et al., 2005; Zweers, 2022). This means adjusting A_s until R^{sim} is the same as R^{meas}:

$$R_{R_{\lambda 1}}^{meas} = R_{\lambda 1}^{sim}(A_s) \quad (2)$$

The surface albedo is assumed to be nearly constant within the wavelength pair where the reflectances are calculated and measured, therefore allowing for R^{sim} at λ_2 to be calculated using A_s found at λ_1. Thus, Eq. 1 reduces to:

$$AAI = -100 * \log_{10}\left(\frac{R_{\lambda 2}^{meas}}{R_{\lambda 2}^{sim}}\right) \quad (3)$$

Equation 3 quantifies the statement that the AAI is calculated based on the ratio of the measured (aerosol-influenced) and simulated reflectance. The TROPOMI simulated reflectance calculations are made assuming a purely Rayleigh scattering atmosphere bounded by a Lambertian surface (de Graaf, et al., 2005; Kooreman, et al., 2022; Zweers, 2022). The primary wavelength pair used for the AAI calculation is 340 nm and 380 nm, which is the typical value used for other instruments (de Graaf, et al., 2005; Zweers, 2022). The Algorithm Theoretical Basis Document (ATBD) for the TROPOMI AAI goes into more specific details about the AAI calculation, including specifics about the chosen assumptions and wavelength pairs used in their data product.

One TROPOMI team (of several) primarily focused on the retrieval algorithm of the AAI product and conducted several studies to determine the impact of clouds on the data product (Zweers, 2022). A Lambertian cloud layer was simulated at different heights while cloud fractions were varied to determine the sensitivity of AAI to the presence of clouds (Zweers, 2022). These tests showed that observed aerosol plumes with an AAI above 1.0 are easily detectable when they are above clouds, whereas aerosols with low indices are more difficult to detect both above and below clouds (Zweers, 2022). While there is no precise AAI value that is deemed more significant than another due to it being a qualitative index, an AAI above 1 typically indicates a moderate to significant presence of absorbing aerosols in the atmosphere (Trees, et al., 2021; Tang, et al., 2023). It is also known that the higher the AAI, the greater the indication of highly dense aerosol concentrations, and the closer to 0, the less the presence of absorbing aerosols and the greater the indication of clouds (Tang, et al., 2023) or a clear, purely Rayleigh scattering atmosphere. AAI values close to 0 or negative in the TROPOMI algorithm are assumed to be clear air or clouds (~0)

or scattering aerosols (< 0) and are therefore not indicative of absorbing aerosols, such as MDAs (Zweers, 2022).

TROPOMI also offers an aerosol layer height (ALH) product. ALH is valuable for air quality monitoring, climate studies, and weather forecasting (Griffin, et al., 2020). ALH gives the vertical height, either in altitude, measured in km, or pressure, measured in hPa, at which aerosols may be found in the atmosphere (de Graaf, et al., 2022; Wu, et al., 2016). ALH therefore indicates the proximity of aerosols to the surface of the Earth (Chen, et al., 2020). One of the drawbacks of the AAI product from TROPOMI is that it does not measure what kind of aerosol is being attributed to the measurements, nor where those aerosols have come from. Oftentimes, aerosols that are found higher in the atmosphere are more likely to have been transported from farther regions, whereas aerosols measured closer to the surface are more likely to be near their sources (Buseck et al., 1999; de Graaf, et al., 2022; Gupta, et al., 2021). ALH data therefore provides important information about the vertical distribution of aerosols in the atmosphere, thus when paired with AAI, ALH can be a useful tool in understanding the distribution and sources of aerosols in different locations.

For ALH retrievals, the layer has a fixed pressure thickness, meaning that the difference between the top pressure and the bottom pressure of the layer is held constant, and the reported height parameter is the mid pressure of the aerosol layer (top pressure plus bottom pressure divided by two) (Apituley, et al., 2022b). TROPOMI's ALH product provides a newer aerosol product focused on retrieving tropospheric aerosol height, addressing the lack of daily global observations of this variable (Apituley, et al., 2022b). While active sensors like ground-based lidar or space-borne Cloud-Aerosol Lidar with Orthogonal Polarisation (CALIOP) offer high vertical resolution of the ALH variable, they have limited spatial coverage (Apituley, et al., 2022b). The algorithm

for TROPOMI ALH relies on the absorption within the oxygen A band, occurring in the spectrum between 759 and 770 nm (Griffin, et al., 2020).

The TROPOMI carbon monoxide (CO) product includes information on the vertical column density of carbon monoxide in the Earth's atmosphere. CO-related aerosols, like soot, tend to be aerosols that come from non-dust-related sources, such as forest fires or industrial and automotive combustion emissions. CO is typically not related to MDAs and it can therefore be useful to know whether observed AAI values also correspond to high CO values that are co-located, thus providing additional information about the aerosol type being measured by the AAI. TROPOMI measures the absorption of surface-reflected sunlight in the near-infrared spectral range to infer the concentration of carbon monoxide (Landgraf, et al., 2023). TROPOMI retrieves CO based on the SICOR (SWIR CO Retrieval) algorithm, which is based on SCIAMACHY heritage (Landgraf, et al., 2022). It improves upon the CO retrieval algorithm used for SCIAMACHY, addressing challenges in cloudy and aerosol-loaded atmospheres. The algorithm used in TROPOMI CO retrievals also accounts for multiple scattering effects, therefore aiding in the elimination of overestimations or underestimations in measured CO concentrations (Landgraf, et al., 2022).

The TROPOMI cloud (CLO) product provides information about cloud properties in the atmosphere. Since the AAI product does not have a cloud-clearing feature, the CLO product can be used to clear the AAI data of gridboxes that are known to have clouds, thus acting as a cloud-clearing parameter. There are two main algorithm methods that TROPOMI uses to derive its CLO product. First, is the Optical Cloud Recognition Algorithm (OCRA) which uses the "color" or "whiteness" of a pixel in different spectral bands to estimate cloud fraction (Lutz & Loyola, 2021). The second is the Retrieval of Cloud Information using Neural Networks (ROCINN), which is a

more advanced approach that employs machine learning to analyze spectral information and derive cloud properties (Lutz & Loyola, 2021). The available parameters in this data product include the cloud fraction, using the OCRA algorithm, which is the main measurement analysed in this study. The cloud fraction measurement indicates how much of a ground pixel is covered by clouds: 0 meaning no clouds (clear sky) and 1 meaning all clouds (complete cloud cover).

1.2.2 MODIS & OMI

The Moderate Resolution Imaging Spectroradiometer (MODIS), is an instrument that is aboard two NASA satellites, Terra and Aqua (Zheng Cao, 2020). Launched in 1999 and 2002 respectively, these satellites have been capturing valuable data of Earth Observation (EO) for over two decades. MODIS is a widely used instrument in the world of remote sensing and MDA research (Tang, et al., 2023; Williamson & Menounos, 2021). MODIS has also been verified against other long-standing instruments in the field, therefore making it a key instrument in validation and understanding of newer aerosol instruments and products. For this study, MODIS data will be used to see if TROPOMI data agrees with the MODIS-observed aerosol distribution. This provides a measure of 'validation' of TROPOMI results, although it is strictly an 'intercomparison' since a true validation requires a comparison to a known 'true' measurement.

All MODIS data used in this study was processed, gridded, and saved by Dr. Ian Ashpole. The MODIS data to be used in the research questions is Dust Optical Depth (DOD) data which was derived from Aerosol Optical Depth (AOD) data by Dr. Ashpole in our research group. DOD is often used to depict the spatial distribution of dust aerosols. DOD indicates the proportion of AOD attributed solely to mineral dust particles, distinct from the overall contribution of all aerosols. Originally introduced for global dust source mapping by Ginoux et al. (2012), DOD calculations rely on the retrieval of Ångström exponent (α) and single scattering albedo (ω) to isolate the dust

component in AOD. DOD used in this study is calculated according to the method of Pu and Ginoux (2016), where DOD is calculated from every AOD retrieval based on an empirically derived continuous function relating α to AOD (Anderson et al., 2005):

$$DOD = AOD \times (0.98 - 0.5089\alpha + 0.0512\alpha^2), if\ \alpha < 1$$

The DOD dataset used in this study comes from MODIS aerosol products retrieved using the "Deep Blue" algorithm. MODIS collects data over 36 spectral bands which are each sensitive to specific wavelengths (Zheng Cao, 2020). Deep Blue uses radiance measurements in the blue channels, meaning the shortwave visible blue light wavelengths, to detect aerosols globally over land, even over bright surfaces such as deserts (Hsu, et al., 2013).

The Ozone Monitoring Instrument (OMI) is a space-based spectrometer that covers the globe daily; it is in many ways the heritage instrument to TROPOMI and also measures reflectance, albeit with a coarser spatial resolution of 13 × 24 km^2 at nadir (Althaf, et al., 2022). It was developed and launched as part of the NASA Aura satellite mission in 2004. OMI has provided data for monitoring various atmospheric parameters, namely its UVAI product. One of the aims of the TROPOMI mission was to eventually overtake the OMI heritage by providing improved versions of the same data products in the future (de Smedt, et al., 2021). To continue the OMI data record and for comparison purposes, the TROPOMI AAI is also calculated for the 354/388 nm wavelength pair, the same pair used for OMI's UVAI product, as the additional computational effort was found to be minimal (Zweers, 2022) and it makes for an easier transition between the instruments when the time comes for OMI to retire (Zweers, 2022). The connection between

TROPOMI and OMI therefore supports the use of OMI data in this study to help interpret the TROPOMI results.

1.3 Scope of Work

To guide the research and writing process of the project, the following research questions have been articulated.

1.3.1 RQ1

As previously mentioned, there are only two well documented sources of MDAs in northern Canada, however, current MDA research in northern Canada is limited. Dr. Ashpole has done work on mapping aerosol patterns (like frequency of occurrence) using data from the MODIS instrument. His results indicate that prominent MDA activity in northern Canada is relatively common.

As previously discussed, atmospheric monitoring in high-latitudes is challenging. The results of Dr. Ashpole's remote sensing-based work indicate previously unmapped dust sources that are consistent with other researchers' predictions about where MDA sources may be in northern Canada, based on dry lakebed and peri-glacial locations (Bullard, et al., 2016; Meinander, et al., 2022). This therefore begs the question: *What is the distribution of dust in northern Canada, according to TROPOMI?* (herein referred to as RQ1). A secondary question can further be asked as to the nature of these distributions and the implications of the ALH dataset for the AAI dataset: *What does ALH data imply about the sources of the observed dust distributions?* (herein referred to as RQ1b).

1.3.2 RQ2

TROPOMI relies on solar radiation to make its measurements. Parts of the northern hemisphere do not receive much solar radiation in the winter months due to the axial tilt of the Earth away from the Sun at those times. This means that high quality TROPOMI retrievals are not often possible at times when limited solar radiation is detected by the instrument. It is important to note that not the entire northern hemisphere is affected by this lack of light and that many mid-latitudes and tropics have little issues in the winter months. Despite this, the general high-quality data availability is expected to be lower in the winter months of the year. However, TROPOMI always records data, even with limited sunlight and increased cloud coverage. The quality of the measured data is also tracked and flagged to ensure that only the recommended data is used and analysed. To increase confidence in TROPOMI measurements, a question can be asked: *How do TROPOMI-derived aerosol products compare to those of other well-characterized satellite instruments, like MODIS and OMI?* (herein referred to as RQ2).

The MODIS dust optical depth (DOD) product will be intercompared with TROPOMI's aerosol products and will likely demonstrate similar patterns. This is because both DOD and AAI provide information about the concentration of aerosols in the atmosphere. Higher values of either value suggest a greater abundance of aerosols. The OMI UVAI product measures the aerosol index within the same visible channel as TROPOMI, therefore allowing the two datasets to be easily intercompared (Zweers, 2022). The similarities in both instruments' retrieval algorithms will likely lead to similarities in MDA hotspots when analysed. These comparisons will build confidence in the results found in answering RQ1.

1.3.3 RQ3

Both the AAI and ALH data products have limited sensitivity to the type of aerosols present (de Graaf, de Haan, & Sanders, 2022). It is known that AAI values around 0 are typically indications of a clear atmosphere or clouds, while negative values are typically indicative of UV scattering aerosols, none of which are MDAs (Zweers, 2022). The lack of direct aerosol type detection is a limitation of this study as it does not allow for direct analyses of MDAs. However, it is well known that MDAs can be picked out from the overall aerosol datasets by filtering with the TROPOMI CO and CLO data products, so we can ask the question: *What are the effects of filtering the TROPOMI aerosol products by the Carbon Monoxide and Cloud products?* (herein referred to as RQ3).

The Carbon Monoxide (CO) product provides detailed information about the concentration and distribution of carbon monoxide in Earth's atmosphere (Lambert, et al., 2023). The CO product can be used to filter out the presence of carbon monoxide in the aerosol data, effectively filtering for forest fire aerosols that appear in the aerosol datasets. The primary reason for filtering by CO is because carbon monoxide is highly indicative of soot which also produces high AAI values. A secondary reason to filter for forest fires is because fires produce aerosols that do not typically originate in the north and are more likely to be in northern atmospheres due to transportation from winds (Williamson & Menounos, 2021). For the purposes of this study, interest is only in those aerosols that are MDAs, so filtering by carbon monoxide levels will help filter for aerosol types that are desired for this study.

The cloud (CLO) product provides information about the presence, cover, and properties of clouds in the atmosphere (Lutz & Loyola, 2016). This means that this product can be used to filter the aerosol products by cloud contamination levels. While clouds are typically indicated as

negative or near zero TROPOMI AAI values, this can sometimes be a mischaracterization of the true cloud distributions since those numbers are based on assumptions of cloud layer thickness and height in simulations (Zweers, 2022). The cloud clearing offered by the CLO product can aid in the interpretation of the AAI data as it will remove any measured AAI values contaminated by high cloud coverage, thus giving a more accurate depiction of the aerosols, such as MDAs, in northern Canada.

2. Methods

2.1 NASA/Earthdata Granules

As previously mentioned, TROPOMI is an ESA instrument. ESA is an international partner of NASA, meaning that TROPOMI data is also distributed by NASA. Earthdata is a NASA website that was created for public access to NASA's Earth science data, and this is where TROPOMI data was downloaded from for this study. From this website, one can specify the location and date of the data to be downloaded, thus eliminating the need to download all available data for the given data product.

Satellite data is distributed in three levels. Level 1 (L1) is the raw radiance data from the instrument, corrected for known instrument biases and known atmospheric effects. Level 2 (L2) data contains the retrieved products that rely on the radiances as inputs, e.g., AAI, ALH, CO concentration and cloud fraction (CLO) values in this study. Level 3 (L3) data is further processed data from L2 that offers averaged results in a consistent grid or map, typically at 1 x 1° resolution as monthly means. For TROPOMI, the L2 data is the raw swath file that is stored for every orbital pass of the instrument, which also includes low-quality data typically due to sun glint. There are multiple L2 files for a given day that pass over the same geographical area given the 2600 km swath width. For the purposes of this study, analysing L2 data directly becomes cumbersome due

to the large geographical area and the large number of L2 files available for the study period (all of 2021). A regridding procedure to process the L2 files into L3 files was therefore used to better visualize the data present in the L2 files on a regular Cartesian grid. This meant converting the L2 swath data from the pixel resolution of 7 x 3.5 km^2 to a daily-averaged gridded product with a resolution of 5 x 5 km^2 (cf. the typical 1 x 1° resolution and monthly means in standard L3 products). This included filtering out low-quality data (described in detail below) before averaging all of the swaths passing over the same 5 x 5 km^2 area daily and storing the number of L2 elements in the average. This procedure allowed for the comparison of TROPOMI data with other datasets and a more meaningful data analysis, e.g., calculating frequency of occurrence or means.

All TROPOMI data granules come in three main types: near real-time (NRTI), offline (OFFL) and reprocessed (RPRO). NRTI data is the most approximate but fastest data that can be obtained directly from the instrument. OFFL data undergoes additional processing and operations that have been determined and optimized by the retrieval algorithms given known errors; it is typically available within 12 hours of initial sensing (Zweers, 2022). RPRO data is the newest version of the data that has been re-processed given the most up-to-date information about the satellite, instrument and data processing algorithms, which all affect the final data product, like AAI. When the work in this paper was in its preliminary stages, the first year that had all RPRO data for every day of the year was 2021. So, 2021 was used as the case study year for this project. RPRO data is, in theory, the best version of the data because it has been re-processed according to the newest information revealed about the full chain of processing (satellite, instrument and retrieval corrections). Appendix B has more information regarding differences in OFFL and RPRO data where it can be seen that they are not negligible, thereby justifying the use of the RPRO data.

2.2 RQ1

To visualize the observations from TROPOMI, the data was gridded to fit on a cartesian map (i.e., high-resolution daily mean L3 files) at a resolution of 5 x 5 km^2, for a total of 900 x 3600 gridboxes, starting at (40N, 160W) and ending at (90N, 50W). The method used for regridding was adapted directly from the Atmospheric Toolbox Website where ESA has provided open source software in support of the Sentinel Earth Observation missions, with examples of how to use and process TROPOMI data. HARP is a toolbox designed to read, process, and compare data from various sources like satellites, models, on-site measurements, and ground-based remote sensing. The main purpose of HARP is to make it easier to compare different datasets. In this study, the HARP module was used to load each L2 granule and create a daily file for each day of the study period. The exact regridding parameters and explanations can be found under the *use case* titled "Creating gridded Level 3 data with HARP from multiple TROPOMI Level 2 UVAI files," which can be found on the Atmospheric Toolbox Website. As per the recommendations in the *Product User Manual* (PUM) released by the TROPOMI team, individual data granules were filtered by the associated quality flag for each granule (Apituley, et al., 2022c). This means that data that did not pass quality assurance, meaning that they did not have a quality flag greater than 80%, were excluded from this research. The individual granule files also contained a lot of information and metadata that is useful, but not necessary to preserve. Preserving all parameters stored in the raw data files would have necessitated a great deal of storage and computing power, both of which were often limited. For that reason, when the data was regridded, only the necessary parameters were kept (longitude, longitude_bounds, latitude, latitude_bounds, time, and absorbing_aerosol_index). The data was finally stored and saved as daily files for the full study period.

Frequency of occurrence (FoO) analyses are a valuable method for gaining insights into data patterns, trends, and distributions. FoO analyses show the spatial distribution of the frequency of data points (gridboxes) that match the desired criteria. To create the FoO maps found in this study, an analysis was done daily over specific months in 2021. For each gridbox, a count of how often an AAI value met the threshold was found over the entire month. This is referred to interchangeably as "number of observations (n_obs)" or "number of days (n_days)" henceforth. In the end, I created a dataset with gridboxes, each with an integer number from 1 to 30 or 31, depending on the length of the month, which indicated how many days that gridbox had a daily mean AAI value above the desired threshold. The process of obtaining the mean AAI over a month was similar to that of FoO, but instead of a counter of the AAI above the given threshold, the actual daily mean AAI value was retained and then averaged for each gridbox for the entire month.

To filter the AAI data by ALH data, the same regridding process was conducted. First, all ALH data was gridded to match the AAI data, also on a daily basis. The main difference in the ALH regridding was that the quality flag minimum recommended for the ALH data is 50%. Once the daily ALH files were obtained, the filtering of AAI data began. This was done by looking at each gridbox for each day and specifying whether or not the AAI gridbox should be set to "NaN," based on whether or not the equivalent gridbox on the equivalent day of the ALH data was above the desired ALH threshold. For example, if it was chosen that we only wanted to look at data where daily mean ALH was less than 1000 m, a list of gridboxes where daily mean ALH was greater than 1000 m was found, then used to set all of the equivalent gridboxes in the daily mean AAI dataset to "NaN," since both datasets had the same number of gridboxes. In the end, I created a dataset of daily files that only contained AAI data if the ALH threshold was met. The purpose of filtering AAI data was to have a better understanding of the vertical aerosol distributions. Unlike AAI, ALH

has a cloud-clearing feature that, when the AAI data is filtered by ALH, can help in eliminating aerosol data that is contaminated by clouds (even though we were also looking for positive AAI values that should correspond to absorbing aerosols and not clouds).

2.3 RQ2

Correlation, used in this research question, is a statistical measure that quantifies the degree of association or relationship between two variables. It tells us how changes in one variable correspond to changes in another. For this work, the Pearson Correlation Coefficient was calculated using the following formula:

$$R = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}$$

$$where\ x_i\ and\ y_i\ are\ the\ values\ in\ the\ x\ and\ y\ datasets$$
$$where\ \bar{x}\ and\ \bar{y}\ are\ the\ means\ of\ the\ x\ and\ y\ datasets$$

If R is found to be close to 0, it indicates two datasets that do not correlate, whereas if R is found to be close to 1 it indicates a strong positive relationship, and an R close to -1 indicates a strong negative relationship (Berman, 2016; Faizi & Alvi, 2023). A strong correlation is said to be when R is greater than 0.7 or less than -0.7, which corresponds to R^2 values of around 0.5 or greater. R^2 is a metric of how well a statistical analysis fits the analysed phenomenon, or how well the regression line fits the two analysed datasets (Faizi & Alvi, 2023). Another way to put it is that an R^2 of 0.5 means that the model explains 50% of the variance in the data. Additionally, a *p*-value is representative of the probability that statistical results are due to random chance or if there is a true (i.e. 'significant') relationship between the two datasets (Amrhein & Greenland, 2022; Faizi & Alvi, 2023). If the *p*-value is found to be significant, here meaning a $p < 0.1$, it indicates that there

is a significant relationship indicated by the correlation coefficient that is unlikely to be due to chance alone (Faizi & Alvi, 2023).

In the case of MODIS-TROPOMI correlations, the calculation of correlation describes whether changes in aerosol distributions observed in MODIS data track the changes observed in TROPOMI data. TROPOMI and MODIS do not, however, measure precisely the same quantities. MODIS measures DOD and TROPOMI measures AAI. While the exact quantities are not the same, the qualitative information that the two measures provide has largely the same physical source (dust). An increase in DOD indicates an increase in absorbing aerosols in the atmosphere. An increase in AAI also indicates an increase in absorbing aerosols in the atmosphere. Since both changes qualitatively describe the same phenomenon (apart from biomass burning aerosols, which are also absorbing), calculating the correlation between the two datasets can reveal whether or not the two instruments measure the same changes in aerosol distributions.

The daily gridded MODIS data used in this study was saved at a resolution of 10 x 10 km^2, but the daily gridded TROPOMI data was saved as 5 x 5 km^2. It was therefore necessary to regrid the TROPOMI data to match the coarser resolution of the MODIS data. TROPOMI was used as the dataset to adjust because its resolution was initially finer, meaning that it would require simple averaging instead of oversampling of data onto a finer gridbox. For each gridbox a correlation coefficient and *p*-value were calculated using the daily data for the desired study period. This correlation was only calculated in specific gridboxes. It was specified that there had to be non-NaN data for the same day in the same gridbox for the AAI or DOD value to be included in the correlation calculations, otherwise that day and gridbox was skipped. This ensured that only days with data from both instruments were being correlated.

OMI scans the Earth daily with a spatial resolution of 13 × 24 km^2 (Althaf, et al., 2022). The UVAI is derived from the OMI satellite, provided as gridded daily global L3 data with a resolution of 100 x 100 km^2, which is accessible through Earthdata as well (Althaf, et al., 2022). The process of analyzing the correlation between OMI and TROPOMI data was the same as with MODIS data. The key difference was that the OMI correlation and *p*-value was calculated only for August, rather than for the four-month (June through September) summer period used for the MODIS correlations. The other difference is that OMI data is stored at an even coarser resolution, and TROPOMI data was also rebinned to match it.

2.4 RQ3

To understand the impacts of filtering the AAI dataset by the CO and CLO datasets, a combination of processes of data analysis were followed as in RQ1 and RQ2. The first step was to regrid the CO and CLO data using the same methods as regridding the AAI and ALH data, meaning that individual L2 granules were converted to daily mean maps on a grid of 5 x 5 km^2. For both datasets, a quality flag of at minimum 50% was used, as recommended by both PUMs. Once both CO and CLO data were saved as daily mean gridded files, the next step was to filter the AAI data by CO and CLO data at different thresholds. This was done using the same programming methods as when filtering by ALH data, however thresholds were chosen as follows.

For CO, any gridboxes in the AAI dataset that corresponded to gridboxes in the CO dataset that were below or equal to the desired threshold were kept, otherwise, the gridboxes were set to "NaN." It was decided to do this over a month period in August 2021, due to computing and time requirements, but considering that August was the peak for both dustiness and satellite observations. To determine the desired CO thresholds to be used in the AAI filtering, the first step was to make FoO maps for a wide range of CO thresholds to understand the changes in CO values

as the thresholds changed. A mean CO map was also made to view the different CO values within the chosen geographical area. This was done to aid in visualizing the distributions of CO and what the minimum and maximum CO values were. Once this was done, a set of CO thresholds was determined to encompass the range of measured CO values, without having too few or too many maps. A study by Wan, et al., 2023, reports a total column CO value of between 0.018 mol m^{-2} and 0.032 mol m^{-2} over Australian savanna and temperate forest as "background CO", which was consistent with the observed CO values in the FoO and mean maps in this study, so it was decided that filtering would start with a value for CO of 0.01 mol m^{-2}, then increase by 0.01 mol m^{-2} until a CO value of 0.05 mol m^{-2}. The 0.05 mol m^{-2} value was used as the cutoff since it is nearly three times the lower literature value for background CO (3 * 0.018 = 0.054 mol m^{-2}); the lower value is not unreasonable given that we are working in the Arctic. The filtering of the AAI dataset by the CO dataset was done for every day over the desired period for each gridbox at the desired CO thresholds. In summary, as the threshold for CO filtering decreases, more gridboxes would be set to "NaN" in the AAI dataset. If the gridbox had a higher CO, the corresponding AAI gridbox was set to NaN. "Higher" in this case meant above the chosen threshold, which was varied to properly evaluate the impacts that the CO dataset had on the AAI dataset.

For CLO, any gridboxes in the AAI dataset that corresponded to gridboxes in the CLO dataset that were below or equal to the desired cloud fraction were kept, otherwise, the AAO gridboxes were set to NaN. Similar to the CO process, FoO and mean maps were made to understand and visualize the CLO data. This was done for every day over the month of August for each gridbox at the desired CLO thresholds. This means that AAI gridboxes that had high co-incident CLO were removed to observe the distribution of AAI data that were not cloud contaminated.

Once two new daily mean high resolution L3 TROPOMI datasets were produced daily (AAI filtered by CO and AAI filtered by CLO), the correlation between these new filtered AAI datasets was calculated and again compared to MODIS DOD data. This gave an idea of how the CO and CLO filtering impacted the correlation of AAI data to MODIS, which has a cloud-clearing feature and is calculated so as to represent only dust aerosols. In theory, the AAI dataset filtered by CO should correlate better with MODIS DOD data since non-dust aerosols would be removed from the AAI data. Similarly, the AAI dataset filtered by CLO should correlate better with MODIS DOD since cloud-contaminated AAI gridboxes would be removed, acting as a cloud clearing for the AAI data to match MODIS cloud clearing. In the same fashion as with RQ2, the AAI datasets had to be regridded to match the spatial resolution of MODIS. Then, correlations were calculated for each gridbox for the desired period. This was repeated for the different AAI datasets for each CO and CLO threshold that was tested to filter the AAI data.

2.5 Summary of Data Products and Datasets

Table 1:Summary of data products used for each RQ. Products and datasets are shown in chronological order in terms of when each was used and processed in section 2.

Research Question	Instrument	Product	Time Period	Resolution	File Type	Dataset size
RQ1	TROPOMI	AAI	01/01/2021 – 12/31/2021	L2: 3.5 x 7 km^2	netCDF4	1.03 TB
				L3: 5 x 5 km^2		18.01 GB
		ALH	01/01/2021 – 12/31/2021	L2: 3.5 x 7 km^2	netCDF4	408.44 GB
				L3: 5 x 5 km^2		27.53 GB
		AAI x ALH	06/01/2021 – 09/30/2021	L3: 5 x 5 km^2	netCDF4	133.38 GB
RQ2	TROPOMI	AAI	06/01/2021 09/30/2021	L3: 10 x 10 km^2	netCDF4	4.73 GB
			08/01/2021 – 08/31/2021	L3: 100 x 100 km^2		126.8 MB
	MODIS	DOD	06/01/2021 – 09/30/2021	L3: 10 x 10 km^2	netCDF4	9.42 GB
	OMI	UVAI	08/01/2021 – 08/31/2021	L3: 100 x 100 km^2	hdf5	25.2 MB
RQ3	TROPOMI	CO	06/01/2021 – 09/30/2021	L2: 3.5 x 7 km^2	netCDF4	368.46 GB
				L3: 5 x 5 km^2		4.77 GB
		CLO	06/01/2021 – 09/30/2021	L2: 3.5 x 7 km^2	netCDF4	972.8 GB
				L3: 5 x 5 km^2		6.26 GB
		AAI x CO	06/01/2021 – 09/30/2021	L3: 5 x 5 km^2	netCDF4	230.81 GB
			08/01/2021 – 08/31/2021	L3: 10 x 10 km^2	netCDF4	4.44 GB
		AAI x CLO	06/01/2021 – 09/30/2021	L3: 5 x 5 km^2	netCDF4	155.6 GB
			08/01/2021 – 08/31/2021	L3: 10 x 10 km^2	netCDF4	3.37 GB
	MODIS	DOD	08/01/2021 – 08/31/2021	L3: 10 x 10 km^2	netCDF4	2.41 GB
					Total	3.36 TB

3. Results and Discussion

3.1 RQ1

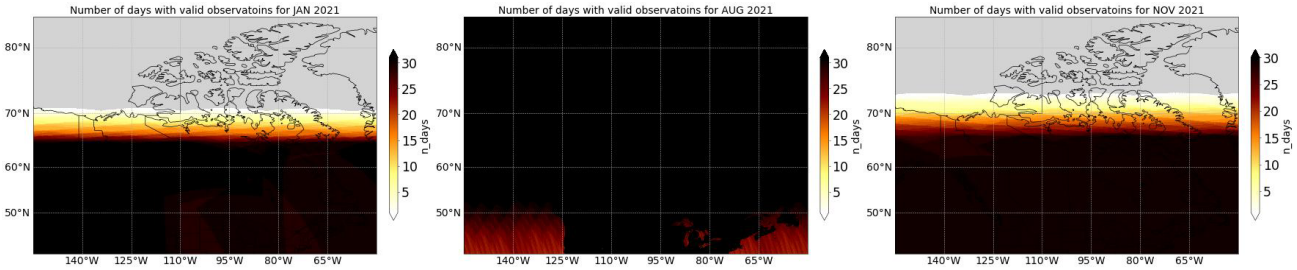

Figure 3: TROPOMI n_days with AAI observations for January (left), August (middle), and November (right) 2021. Grey coloring indicates gridboxes with NaN data for that gridbox for the entire month. NaN data means no good quality observations were made for the entire month. The minimum n_days is 1 and maximum n_days is the total number of days in the month (30 or 31). If n_days was 0, the gridboxes were also set to NaN.

Figure 3 illustrates the n_days with TROPOMI AAI observations for each 5 x 5 km^2 gridbox over three select months. The darker a section is, the greater the number of days those areas have valid data. The grey areas seen in the first and last panels represent gridboxes that are "NaN," i.e., gridboxes that did not have any days with observations over the entire month. This figure shows the phenomenon discussed in section 1.2 regarding the lack of available data in the winter months; in January and November, there was either no good quality data or no data at all due to a lack of sunlight in northern Canada. This gives reason to analyze data from June to September, to ensure that there is sufficient data in northern Canada.

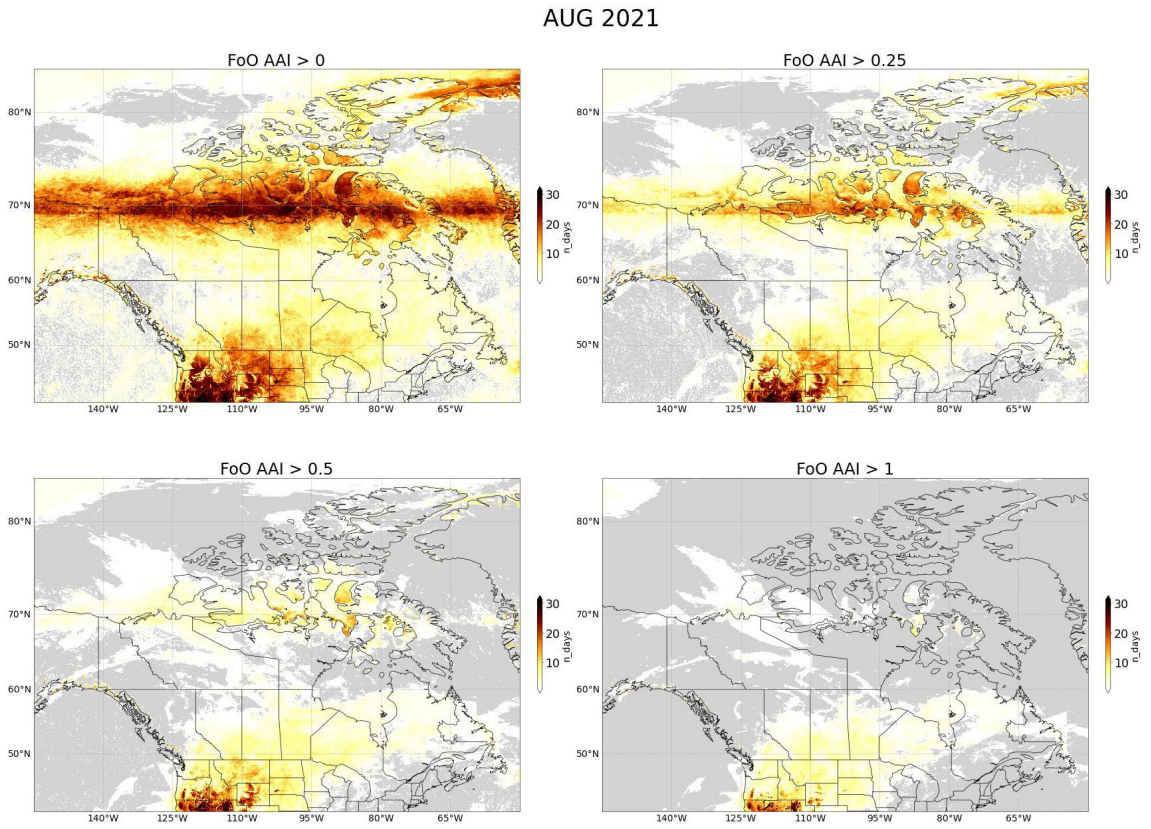

Figure 4: FoO of AAI at increasing thresholds in August 2021. Grey coloring indicates gridboxes either with NaN data or gridboxes with AAI below the given threshold (if n_days = 0, it was set to NaN). The minimum n_days is 1 and the maximum n_days is 31.

Figure 4 shows darker areas (around 70N) where AAI is above the given threshold for multiple days in the month of August. The top left panel shows the FoO of AAI above 0. The 0 threshold was deliberately chosen due to the fact that AAI = 0 is more likely to be clouds or clear skies, meaning that there is little to no absorbing aerosols present or readily detectable by TROPOMI, and AAI < 0 corresponds to small scattering aerosols like sulphate. The top right panel shows the FoO of AAI above 0.25, which has decreased compared to the top left panel. At this AAI threshold, the retained gridboxes are now showing aerosol plumes that are not scattering and are therefore likely to be MDAs; any confounding clouds impacts on AAI are also lower at this threshold.

Looking more closely at northern Canada across all four panels (70N and above), it can be seen that there are areas that show observed AAI values above 0 for multiple days over the 31-day

month. This is apparent especially across the 70N band in the top two panels, where there is a clear concentration of measured AAI values above 0 and 0.25. This observed FoO spatial distribution was unexpected and led to a great deal of thought as to what was causing it. It appeared as an artificial feature and was first thought to have been an issue caused during data analysis. However, a replication of the banding feature was made by Dr. Ashpole, therefore leading to more questions about what was causing it. Before the commencement of this project, the banding feature was intensely investigated over the summer of 2023. The investigation included issues with data versions, sampling geometry, daytime or nighttime observations, or issues with cloud coverage. None of these appeared to yield a strong conclusion as to what was causing the feature. It is important to note that the banding feature disappears with higher AAI values being used as the FoO threshold, notably the bottom panels in Figure 4 (AAI threshold of 0.5 and 1). This is an important observation because MDAs are unlikely to be detected with an AAI threshold close to 0, since they are non-scattering aerosols (and not clear skies). Since no clear explanation was found for the banding feature and since the banding feature disappeared upon analysis of AAI values more closely linked to MDAs, it was decided to focus on the notable FoO values of higher AAI values in northern Canada.

Across all panels in Figure 4, there is evidence of frequent observations of AAI above the given threshold around previously mentioned source areas, such as Kluane lake around (60N, 140W), which will be further discussed in section 3.1.2. The spatial distributions of FoO where AAI is above a given threshold in northern Canada are especially important in the bottom panels since those AAI thresholds are even greater, meaning that those AAI measurements are even less likely to be caused by scattering aerosols and are more likely to be caused by absorbing aerosols, like MDAs. This is an important observation as it indicates that the presence of absorbing aerosols at

substantial levels exists in the areas of interest to this study. From the bottom panels, there is a decrease in the number of days observed to have AAI observations above the given threshold, shown by the increased grey area and the more white (i.e. closer to 1) coloring that appears. This indicates that AAI values that are higher, indicating greater abundance of absorbing aerosol plumes, are less and less frequent. This finding indicates that occurrences of dense and absorbing aerosol plumes are not frequent, but they do exist. It is also important to emphasize that these maps show the frequent observations of AAI values at and above the given thresholds, but they do not uniquely indicate aerosol types (dust vs. soot) or aerosol plume densities. These maps are an indication of geographical locations where particles are absorbing radiation.

Figure 4 also shows frequent observations of highly positive AAI values in the southwestern study areas, roughly below 50N, and between 125W and 100W. The highest AAI values in this area are most likely associated with wildfires, as supported by co-incident (August 2021) elevated mean CO values in northern California and Oregon in this work (Figure 17, Section 3.3.1). The historical distribution of MODIS-derived fire counts in the northwest US has been well documented (e.g., Vaillant et al., 2016), but wildfires exhibit high interannual variability (e.g., Park et al., 2007, van der Werf et al., 2006). It is interesting to note that 2021 was an exceptionally active fire season in North America (Jain et al., 2024), including in BC, Saskatchewan, Manitoba and western Ontario (also evident in Figure 17). These 2021 fires are likely contributing to the diffuse patch of FoO ~ 10 in central Canada below 60 N, transported all the way to northern Quebec by prevailing westerly winds; however, they are spatially separate from FoO enhancements in the Arctic Archipelago, the latter unlikely to be the result of smoke (absorbing) aerosols and most likely the result of dust (absorbing aerosols). We return to these latter dust absorbing aerosols, which are the focus of this study.

The same four-panelled figures seen in Figure 4 were made for June, July, and September (seen in Appendix C). Across all these figures, there is a common theme of observed AAI values for a frequent number of days at high latitudes. Since the observed AAI were filtered for values greater than 0, these are non-scattering – or rather absorbing – aerosols. This further implies that there is a source of some kind that is emitting absorbing aerosols in high latitudes.

3.1.1 RQ1 b)

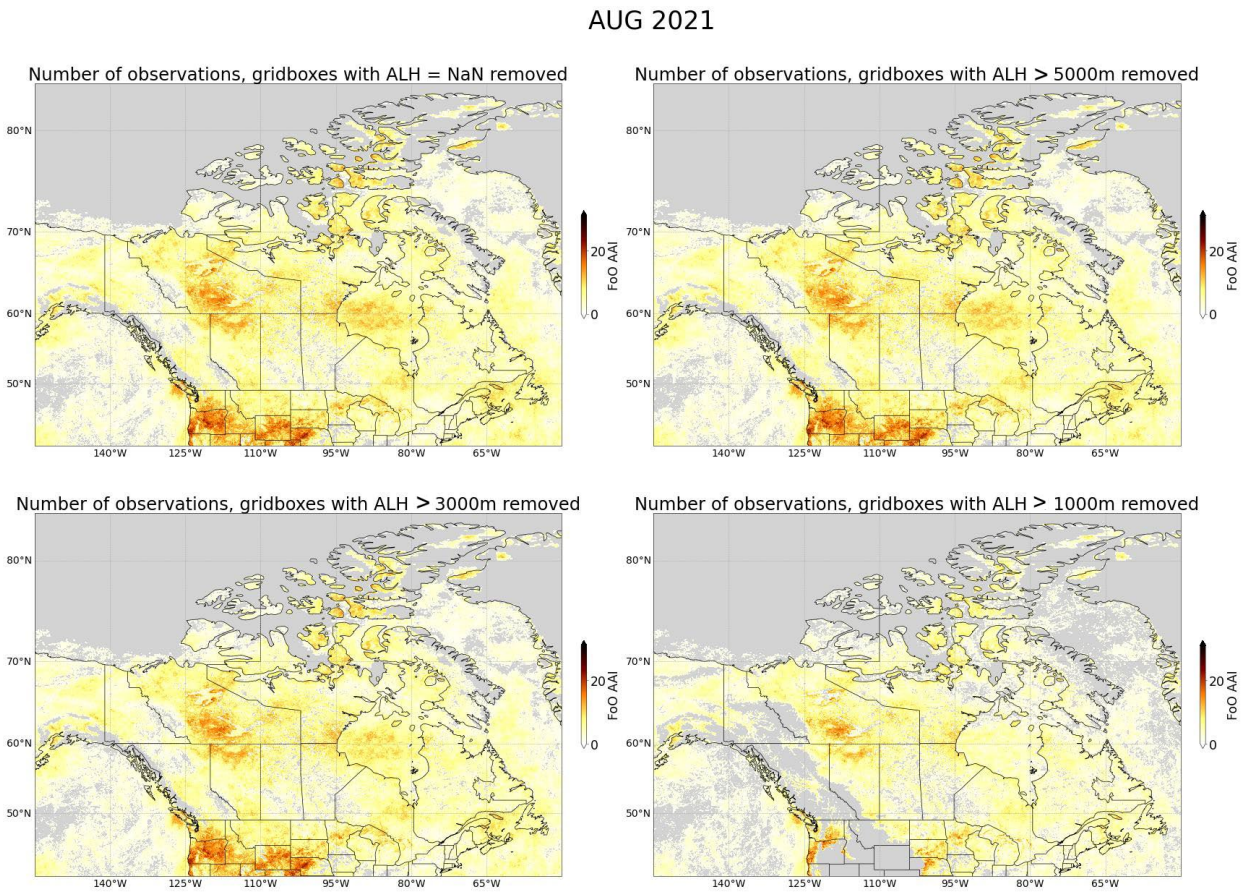

Figure 5: FoO AAI (n_obs) for August 2021 as the AAI data is filtered by ALH data at different height thresholds. Gridboxes that did not have an associated ALH value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum FoO AAI is 1 and the maximum is 31. AAI includes all values (> and < 0).

The first panel in Figure 5 shows the n_obs of AAI (both above and below zero) where only those AAI values were removed that corresponded to "NaN" ALH values. Since there is no specified ALH threshold being used, the AAI data has been effectively filtered by either low

quality ALH data, seen as “NaN” values in the ALH dataset, or by cloud contaminated ALH data, also seen as “NaN.” This first set of filtering practically acts as a cloud clearing filter for the AAI data, since ALH has a cloud clearing feature, discussed in section 1.2.1.1; 34.71% of grid boxes were removed in this step. As the ALH threshold gets lower, meaning only considering aerosol layers closer to the ground, any gridboxes with aerosol layers above that threshold were excluded. Filtering for ALH < 5000 m has little additional effect (35.09%) compared to filtering for all ALH (34.71%) because most aerosol plumes identified by TROPOMI are below 5000 m. As the dataset was filtered for aerosol layers closer to the surface of the Earth, the percentage of valid AAI gridboxes in many regions decreased. The exact numbers and changes in valid AAI gridboxes can be seen in Table 2. The total number of AAI gridboxes for ALH < 5000 m is 2,115,443 and for ALH < 3000 m is 2,103,099 (1.68% decrease). Restricting AAI to plumes below 1000 m leaves 1,703,033 gridboxes (a further 17.64% decrease). A summary of this information is found in Table 2 where it can be seen that there is a decrease in the number of gridboxes across all datasets, but that decrease is not overly large after the initial drop (34.75%) corresponding to cloud filtering. It can also be seen, however, that the FoO (plot colour) does not drastically decrease as ALH gets closer to the surface.

I

Table 2: (Row 1) - Changes in the number of non-NaN gridboxes for the n_obs for August 2021 for the filtered AAI datasets by ALH data. (Rows 2:5) - Percent differences in number of non-NaN gridboxes from the other AAI datasets. All gridboxes are within the same geographical area represented in Figure 3.

Row		No ALH filtering	ALH = NaN	ALH < 5000 m	ALH < 3000 m	ALH < 1000 m
1	Total non-NaN gridboxes	3240000	2115443	2103099	2067817	1703033
2	Percent difference from no ALH filtering		34.71 %	35.09 %	36.18 %	47.44 %
3	Percent difference from ALH = NaN			0.58 %	2.25 %	19.50 %
4	Percent difference from ALH < 5000 m				1.68 %	19.02 %
5	Percent difference from ALH < 3000 m					17.64 %

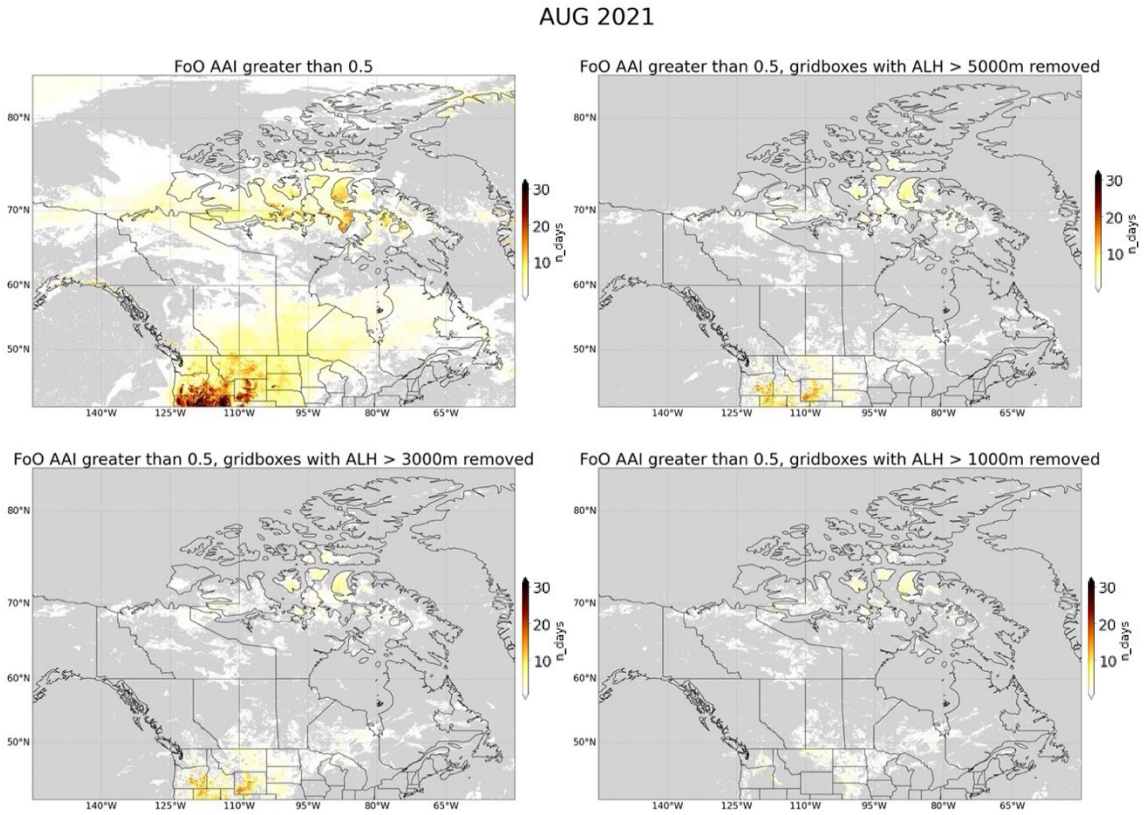

Figure 6: FoO of AAI greater than 0.5 for August 2021. Each panel is the same AAI dataset that has been filtered by ALH, except the top left which is unfiltered. NaN data is represented by grey coloring. Here, NaN data is either bad quality ALH data or gridboxes where n_days was 0.

Figure 6 shows the FoO of AAI that has been filtered by decreasing values of ALH. For this figure, only FoO of AAI > 0.5 is shown. This is primarily because the banding feature previously discussed does not appear in the FoO maps for AAI values greater than 0.5, as these AAI values are more strongly associated with absorbing aerosols and not clouds. A paper published by Penning de Vries et al. (2009) also shows, via simulations, that AAI is below 0.5 when examining the impacts of various cloud configurations on UVAI (their Figure 2). The top left panel in Figure 6 shows the unfiltered FoO of AAI greater than 0.5. There are significant frequencies of aerosol observations around 125W and below 50N, as discussed already. It can also be seen that there are frequencies of occurrences in the higher latitudes, around 70N. These distributions, while not as frequent as the distributions around 125W and 50N, tell us that there are frequent absorbing aerosols in those regions. The top right panel shows the FoO of AAI above 0.5, but all gridboxes with an ALH greater than 5000 m have been removed. The general distribution of AAI data is comparable to that of the first panel but with fewer spatially and less frequent (at a given gridbox) observations. These exact numbers can be seen in Table 3. The bottom left panel of Figure 6 shows the FoO of AAI above 0.5, but all gridboxes with an ALH greater than 3000 m have been removed. Again, the general distribution of the FoO of AAI greater than 0.5 is comparable to the first two panels discussed, but with less frequent observations, (percent difference seen in Table 3). This is to be expected.

Table 3: (Row 1) - Changes in the number of non-NaN gridboxes for FoO AAI > 0.5 for August 2021 for the filtered AAI datasets by ALH data. (Rows 2:4) - Percent differences in number of non-NaN gridboxes from the other AAI datasets. All gridboxes are within the same geographical area represented in Figure 3.

Row		No ALH filtering	ALH < 5000 m	ALH < 3000 m	ALH < 1000m
1	Total non-NaN gridboxes	1308713	318416	292838	151459
2	Percent difference from no ALH filtering		75.67 %	77.62 %	88.43 %
3	Percent difference from ALH < 5000 m			8.03 %	52.41 %
4	Percent difference from ALH < 3000 m				48.28 %

It is important to notice that there are still distributions of FoO of AAI greater than 0.5 in the northern latitudes, indicated by the non-grey gridboxes in Figure 6. The bottom right panel shows the FoO of AAI above 0.5, but all gridboxes with ALH greater than 1000 m have been removed. This panel shows the distribution of AAI which is also the closest to the surface of the Earth. It can be seen that there are areas in the high latitudes, like around 75N and 90W, where there are still frequent observations of AAI above 0.5. This indicates that the observed AAI values are also close to the surface. As previously mentioned, aerosols that are closer to the surface are less likely to have been transported from other regions and are therefore indicative of aerosol distributions that are in close proximity to their source. These high latitude regions are displaying characteristics, based on the AAI and ALH data, that indicate that there may be sources of dust in northern Canada, located in regions not previously studied/mentioned.

Another important observation to be made upon inspection of Figure 6 is that as high ALH data is removed, the FoO of aerosols in the southwest region, like around (45N, 110W), decreases. This means that the aerosol distributions observed in those locations are less frequently observed close to the surface of the Earth and are more commonly found in aerosol layers higher in the

atmosphere. As previously mentioned, aerosols that are found to be higher in the Earth's atmosphere are more likely to have been transported from other regions or other source areas. The northwest US is known for its active fire season in the summer months (e.g., Griffin, et al., 2020) but also especially in 2021, as already noted in section 3.1. Wildfire aerosols are typically found several kilometers above the Earth's surface, due to the small size of these aerosols and the increased buoyancy of fire-heated air parcels (Griffin, et al., 2020). So, the observed decrease in FoO in these south-western map regions is to be expected because of the nature of fire aerosols, therefore implying that the ALH filtering on the AAI dataset is effective in scrutinizing aerosol types and aerosol source locations. It should, however, be noted that terrain in the northwest US is elevated such that filtering for below 1000 m intercepts the local topography in much of the region, which can be seen in the bottom right panel of Figure 6.

3.1.2 Implications in northern Canada

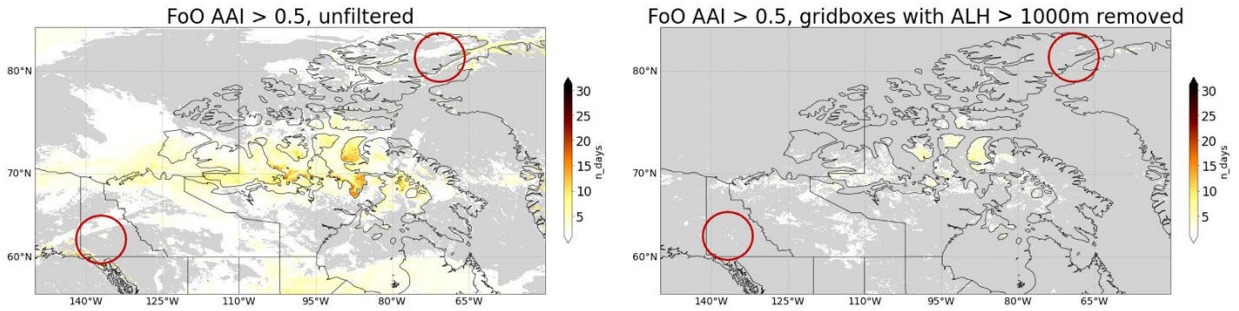

Figure 7:Zoomed-in FoO of AAI greater than 0.5 for August 2021. Left panel is the AAI dataset with no filtering. Right panel is the AAI dataset that has been filtered by ALH < 1000 m. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0. Red circles added to show Kluane Lake and Lake Hazen.

Figure 7 is a zoomed-in comparison of some of the key results found in section 3.1, focusing on FoO of AAI > 0.5 for latitudes above 60N. It can be seen that there are multiple days where AAI values above 0.5 are observed in northern Canada. The red circles, specifically, show known source areas at Kluane Lake (left circle) and Lake Hazen (right circle). The left panel shows the AAI dataset where there is no ALH filtering. It can be seen that there are AAI observations

around Kluane Lake and Lake Hazen, more so for the latter. AAI observations near Lake Hazen are sometimes found at heights greater than 1000 m but the elevated topography near Kluane Lake leaves no observations in its vicinity. Another important observation to be made is that the n_days around both known source locations is not very high. This indicates that known sources of MDAs may not always have the most frequent observations, i.e., both are in mountainous terrain that presents challenges and Lake Hazen is also in the extreme north where light angles are a challenge. This analysis is further important because it implies that other source areas may also show low FoO values, which are observed in many locations above 60N, such as around (70N, 85W) or (70N, 110W). Figure 7 shows many areas where the n_days is either close to or greater than the n_days around Kluane Lake and Lake Hazen, indicating that there are aerosol observations that are comparable to those around known aerosol source areas, therefore indicating that sources of MDAs might exist in these unstudied areas.

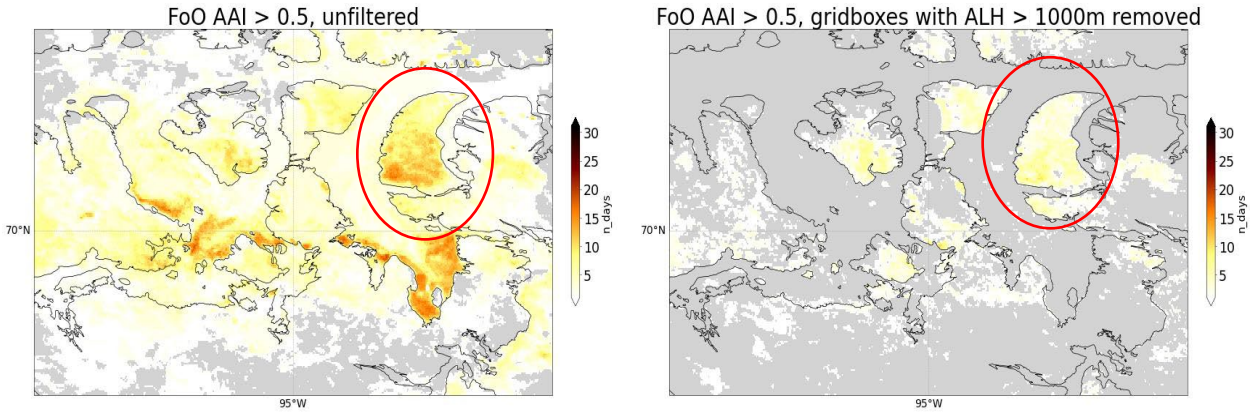

Figure 8: Zoomed-in map of the Arctic Archipelago showing the FoO of AAI greater than 0.5 for August 2021. Left panel is the AAI dataset with no filtering. Right panel is the AAI dataset that has been filtered by ALH < 1000 m. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.

Figure 9: Google Maps image of the Arctic Archipelago showing the composite image of vegetation cover from multiple satellite sources (Adapted from Google Maps, 2024).

Figures 8 and 9 are showing parts the Arctic Archipelago area of northern Canada. This area specifically demonstrated a high number of days with AAI values above 0.5, both in the unfiltered dataset (left panel, Figure 8) and when ALH is below 1000 m (right panel, Figure 8). This hotspot across both FoO maps indicates that there are absorbing aerosols that are low in the atmosphere, suggesting that there might be a source area in this location previously unknown. Figure 9 shows the same area in visible imagery from multiple satellite sources, thus providing a clue to the nature of the AAI observations made. Notably, Figure 9 reveals that the Brodeur Peninsula of Baffin Island (the area encompassed by the red circle) appears to have less dense vegetation cover compared to the surrounding areas. This is apparent from the muted green, grayish coloring in the circle compared to the surrounding greener areas. Since sparse vegetation cover translates to looser soils that are more susceptible to wind erosion (Li, et al., 2021), this finding is consistent with the TROPOMI observations of frequent occurrences of high AAI values. After consulting a Government of Canada Land Use Atlas, the Brodeur Peninsula is confirmed as barren land. This analysis further suggests that there is a source area for MDAs in this region not previously researched. These results demonstrate a need for further investigation of the cause of the observed AAI values, preferably with ground-based observations of dustiness.

3.2 RQ2

3.2.1 MODIS

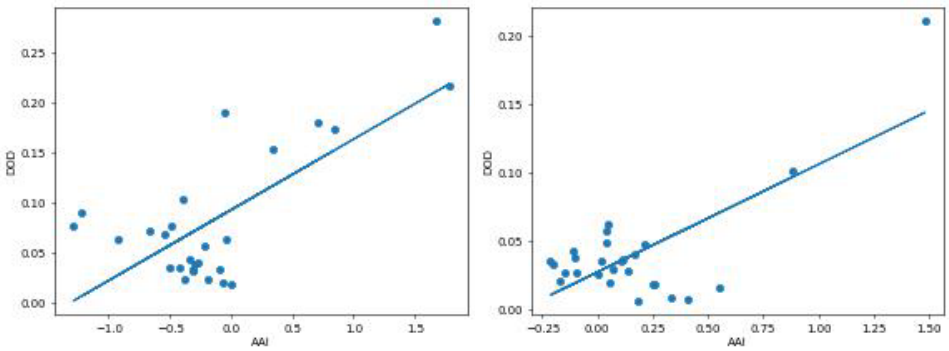

Figure 10: Scatterplots showing the correlation between TROPOMI AAI data and MODIS DOD data for two individual gridboxes. Data includes coincident non-NaN values from June to September 2021. Left and right panels show statistically insignificant and significant ($p > 0.1$) correlations, respectively. See text for details.

To better understand the validity of TROPOMI's aerosol measurements, a comparison of TROPOMI AAI and MODIS DOD was carried out. For each gridbox, DOD and AAI were tracked over time (June to September, 2021) and a correlation between the values was found. Figure 10 illustrates the scatter of data around the line of best fit for two selected gridboxes for the coincident AAI-DOD dataset. These gridboxes were chosen as they were found to have an R-value greater than 0.7. This corresponds to an R^2 value of 0.5. Typically, an R^2 of 0.5 is said to indicate a moderately strong correlation between the two variables, where the line of best fit explains 50% of the variance in the dataset (Faizi & Alvi, 2023). In the left panel, there is a significant amount of scattering around the line of best fit, as well as some outliers, such as the AAI value around 1.5. The right panel shows the scatter of data around the line of best fit from a different gridbox than in the left panel, again with some outliers around AAI of 1 and 1.5. The difference here, however, is that the *p*-value in the right panel was found to be significant (< 0.1). Compared to the spread of the left panel, the right panel appears to have more data points that lie close to the line of best fit. This is because a statistically significant gridbox will display less spread from the linear regression model. It can also be seen that there is a lower number of obvious outliers in the right panel

compared to the left, again showing the statistical significance difference between the two gridboxes.

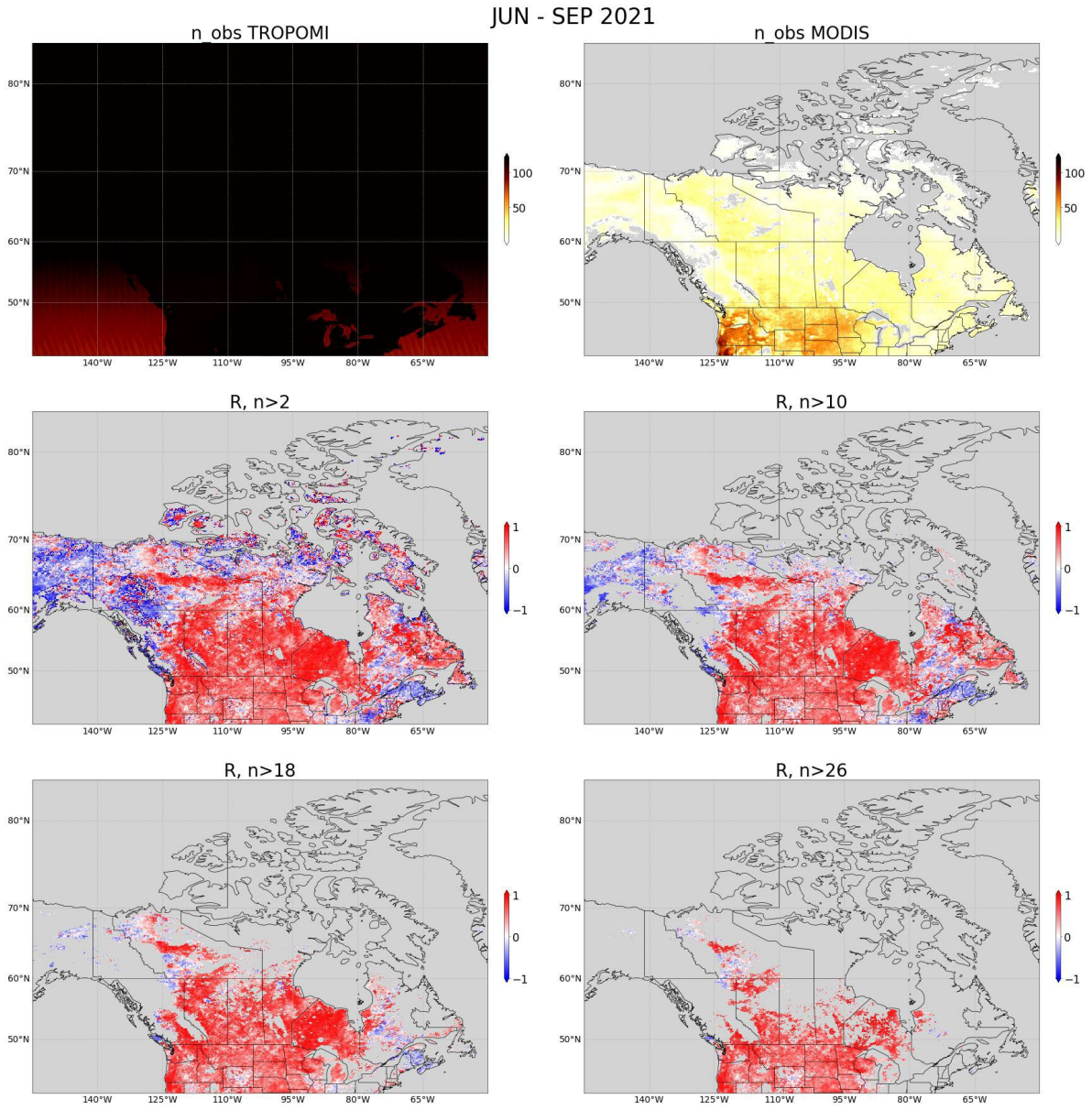

Figure 11: Basis of the coincident data found for the TROPOMI AAI (all data) and MODIS DOD datasets. Top row shows the n_obs for TROPOMI (left) and MODIS (right). Middle and bottom rows show the correlation coefficient (R) of MODIS and TROPOMI data from June to September 2021. Panels in the middle and bottom rows show the change in R as the number of minimum days (n) used in the correlation calculation increases. Red indicates high, positive correlation (close to or at 1) and blue indicates low negative correlation (close to or at -1).

Figure 11 shows the results from the correlation calculations made between MODIS and TROPOMI. The n_obs for TROPOMI and MODIS are shown in the top two panels. These panels show the distributions of observations over the four summer months, which impacts interpretation. Many of the gridboxes in the top right panel in northern Canada have 0 observations for the entire period, which is represented by the grey color. This indicates that, for this section of the analysis, there is a significant lack of data in northern Canada in the DOD dataset, which is the focus of this section. While the purpose of this section is to understand the agreement, or lack thereof, of the instruments, it is important to note that this analysis is limited by MODIS data availability.

When conducting a correlation analysis, a key component lies in the amount of data in each dataset that is being correlated. With more data points, represented by "n", the correlation coefficient becomes a more reliable estimate of the true relationship observed between the two datasets (Berman, 2016). A larger sample size can also reduce the impact of sampling errors, meaning the results are less likely to be skewed by random measurement errors not caught by the quality filtering (Berman, 2016). The TROPOMI AAI and MODIS DOD datasets have a high, positive correlation in many locations (Figure 11), especially as n increases. However, there are several places where there is a strong negative correlation. This could be due to AAI having negative values, and because AAI does not have a cloud-clearing feature, so many of the AAI values are not indicative of pure dust in the atmospheric column (i.e., dust can be over a cloud), like MODIS is. So, removing negative AAI values might eliminate some of the negative correlations observed. However, negative correlations can also arise from other factors, such as different spatial resolution and averaging, sensitivity differences, or assumptions in retrieval algorithm differences. Understanding what causes the negative correlations is ultimately beyond the scope of this study.

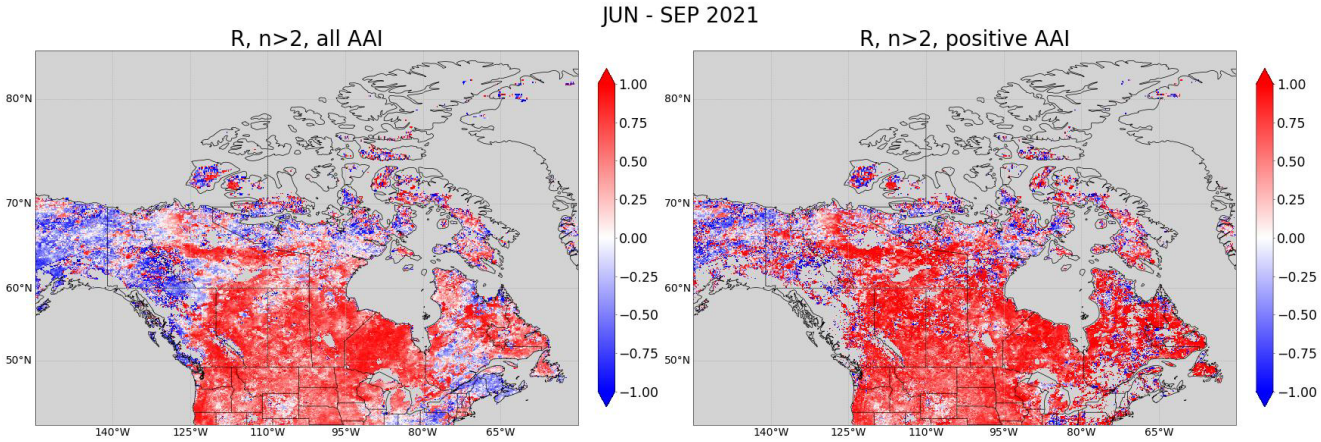

Figure 12: The correlation (R) of TROPOMI AAI (all data) and MODIS DOD for all values of AAI (left) compared to only positive values of AAI (right). Strong red and strong blue coloring indicate strong positive and strong negative correlations respectively. All data comes from June through September 2021 and the minimum number of day (n) used in the calculations was 2.

Figure 12 shows the difference in correlation when the TROPOMI AAI data only includes positive AAI values, which correspond most closely to MODIS DOD (cf. Figure 10, where the negative AAI values are now eliminated, along with their coincident DOD values). There is a slight increase in positive correlation when negative AAI values are removed, but visible negative correlations remain. The observed increase in red gridboxes, indicating positive correlations, is to be expected since MODIS DOD does not include scattering aerosols (i.e. aerosols that lead to AAI < 0) like TROPOMI AAI does. The change in gridboxes after AAI filtering is quantified next.

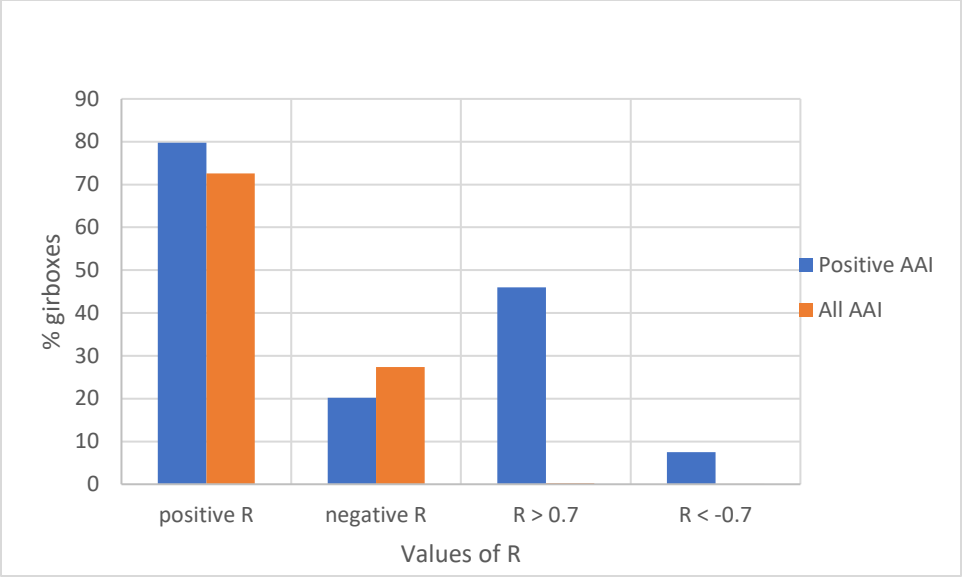

Figure 13: Changes of percent gridboxes for different values of R (n > 2) when either positive AAI values (blue) or all AAI values (orange) are included. Values for the "All AAI" data set (orange bars) for R > 0.7 and R < -0.7 are 0.22% and 0.03% respectively. Gridboxes are in the geographical area shown in Figure 11.

Figure 13 summarizes the percent gridboxes for different values of R for the correlation between MODIS and all TROPOMI AAI, compared to MODIS and only positive TROPOMI AAI. The percent gridboxes are fairly similar (within 10%) for all negative and positive correlations, no matter if AAI is filtered to include only positives, or not. Nevertheless, the percentage of positive R values increases after filtering (72% → 80%) and the percentage of negative R values decreases after filtering (28% → 20%). However, the percentage of gridboxes with R > 0.7 or R < -0.7 is much higher for only positive AAIs (~0% → 46% and ~0% → 8%, respectively). These R values correspond to an $R^2 > 0.5$, meaning that the regression line explains 50% of the variance between the datasets. (Appendix C has a map that shows the R^2 values for all AAI compared to positive AAI.) After filtering, the regression model between MODIS DOD and positive TROPOMI AAI fits better (46% of coincident data have R > 0.7 as opposed to ~0% before). This also indicates that, in general, the instruments display generally positive correlations, but they have even stronger positive correlations when negative AAI values, meaning scattering aerosols or clouds, are removed. This is to be expected as MODIS DOD specifically looks at dust aerosols (i.e. highly

absorbing and non-negative) and therefore should track better when TROPOMI also looks at AAI values that are more likely to be dust. It therefore appears that the agreement between MODIS and TROPOMI is improved when looking only at positive AAI values from TROPOMI.

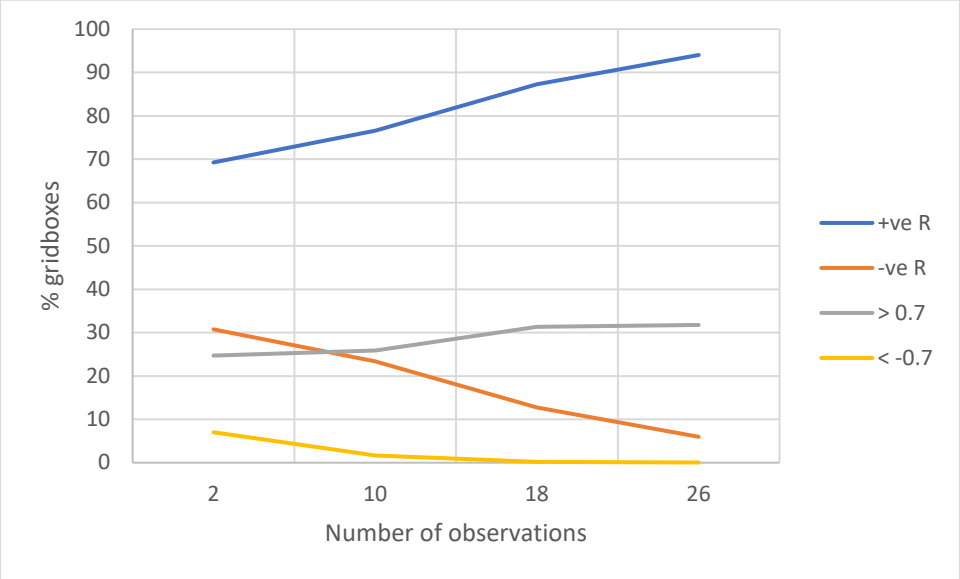

Figure 14: Changes in the percentage of gridboxes that correspond to different values of R as n increases. Gridboxes are in the same geographical area shown in Figure 11. All data comes from June through September 2021.

Figure 14 shows the percentage of gridboxes for different R values as n changes (cf. Figure 10, where n was 26 over June to September in two example gridboxes; cf. also Figure 11 that showed the number of gridboxes decreasing as more coincidence is required). As n increases, the percentage of gridboxes that have a positive correlation also increases, indicating that the instruments agree better as more coincident observations are required, although at the expense of losing gridboxes that do not satisfy the more stringent coincidence criteria (this will have an impact on p-values). Of note is that the percentage of negative correlation gridboxes decreases significantly as n increases, with strongly negatively correlated gridboxes dropping to 0% for 18 coincident observations or more. This supports the idea that MODIS and TROPOMI aerosol data agree for the geographical area shown in Figure 11.

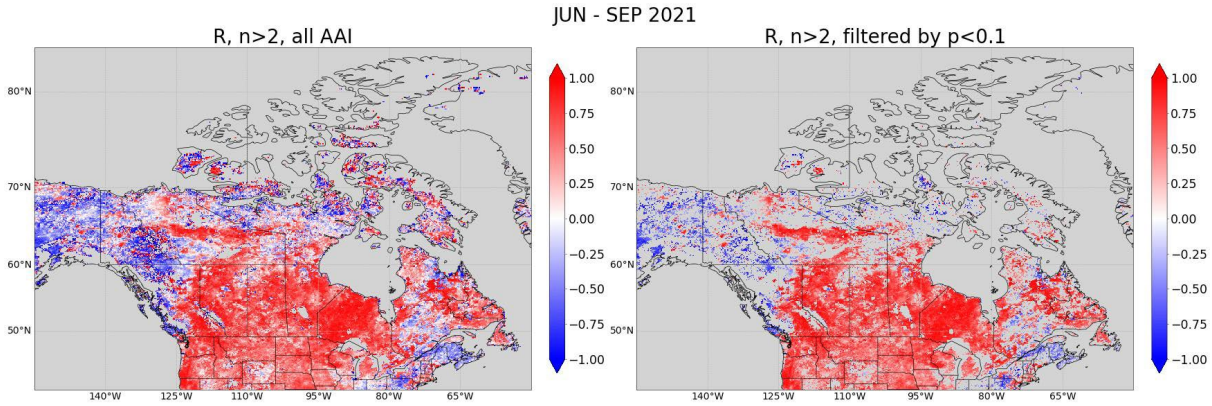

Figure 15: Correlation (R) between TROPOMI AAI (all data) and MODIS DOD for all p (left, same as Figure 12) and filtered by p < 0.1 (right). Strong red and strong blue coloring indicate strong positive and strong negative correlations respectively. All data comes from June through September 2021 and the minimum number of day (n) used in the calculations was 2.

Figure 15 illustrates how filtering the correlation data for significant *p*-values leaves stronger correlations between MODIS DOD and TROPOMI AAI. This is evident from the decrease in correlations close to zero, apparent in the west of Canada near (70N, 125W) where much of the white and faint coloring disappears in the right panel. Although some regions, particularly in northern Canada, lack sufficient data, overall, the *p*-filtered dataset highlights areas of significant correlation between the two datasets. (cf. Figure 10 for a single grid box, which shows the high scatter in a non-significant *p*-value gridbox.) Notably, strong negative correlations remain, but this figure requires the least coincident data points (n > 2), and Figure 14 showed that negative correlations will decrease as n increases (at the expense of geographic coverage as seen in Figure 11). The results shown in Figure 15 further characterize the agreement between the two instruments, now with *p*-value filtering. The overall impression of this analysis is that the TROPOMI AAI > 0 data product agrees reasonably well with the MODIS DOD product. As we require higher levels of data coincidence (increasing n, Figure 11), higher correlation coefficients (R > 0.7, Figures 13 and 14), and higher statistical significance (p < 0.1, Figure 15), the number of gridboxes in the region of interest (northern Canada, > 60N) decreases. However, this does not mean that the abundant TROPOMI

observations in northern Canada (Section 3.1) are suspect. In the next section we compare TROPOMI to its heritage instrument, OMI, which produces the very similar UVAI product, and which is impacted by clouds in the same way as TROPOMI (as opposed to cloud-cleared MODIS DOD data).

3.2.2 OMI

The top two panels in Figure 16 show the n_obs for each instrument for each gridbox over the entire month of August 2021. The important feature to notice here is that both instruments show n=31 for nearly every gridbox for the entire month. This means that the correlation plots will not change as we increase n_obs, especially in northern Canada. The middle-left panel shows the correlation of TROPOMI AAI and OMI UVAI for each gridbox over the entire month. Table 4 shows the percent change in gridboxes when the minimum value of n increases. This data shows that, as expected, no changes to the calculated correlations are observed with increasing n. Looking back at Figure 16, it can be seen that there is a significant amount of positive correlations between both instruments (72.6 % positive gridboxes), as shown by the red on the map in the middle left panel. There is also a significant amount of deeper red gridboxes in higher latitudes, which means stronger positive correlations; however, very few gridboxes in Northern Canada have an R greater than 0.7 (see Table 4). Since both of these datasets are measuring the same quantity, absorbing aerosol index, there is no obvious need to filter out negative AAI values. This is because, in theory, negative AAIs should be measured in both datasets at the same time, therefore giving a high correlation. However, there are some locations where there is a negative correlation, indicated by the blue. There are not, however, large amounts of strongly negative correlations. This is further supported by Table 4 (below), which shows that negative R values account only for 27.4 % of the total gridboxes, with only 0.03% being less than -0.7.

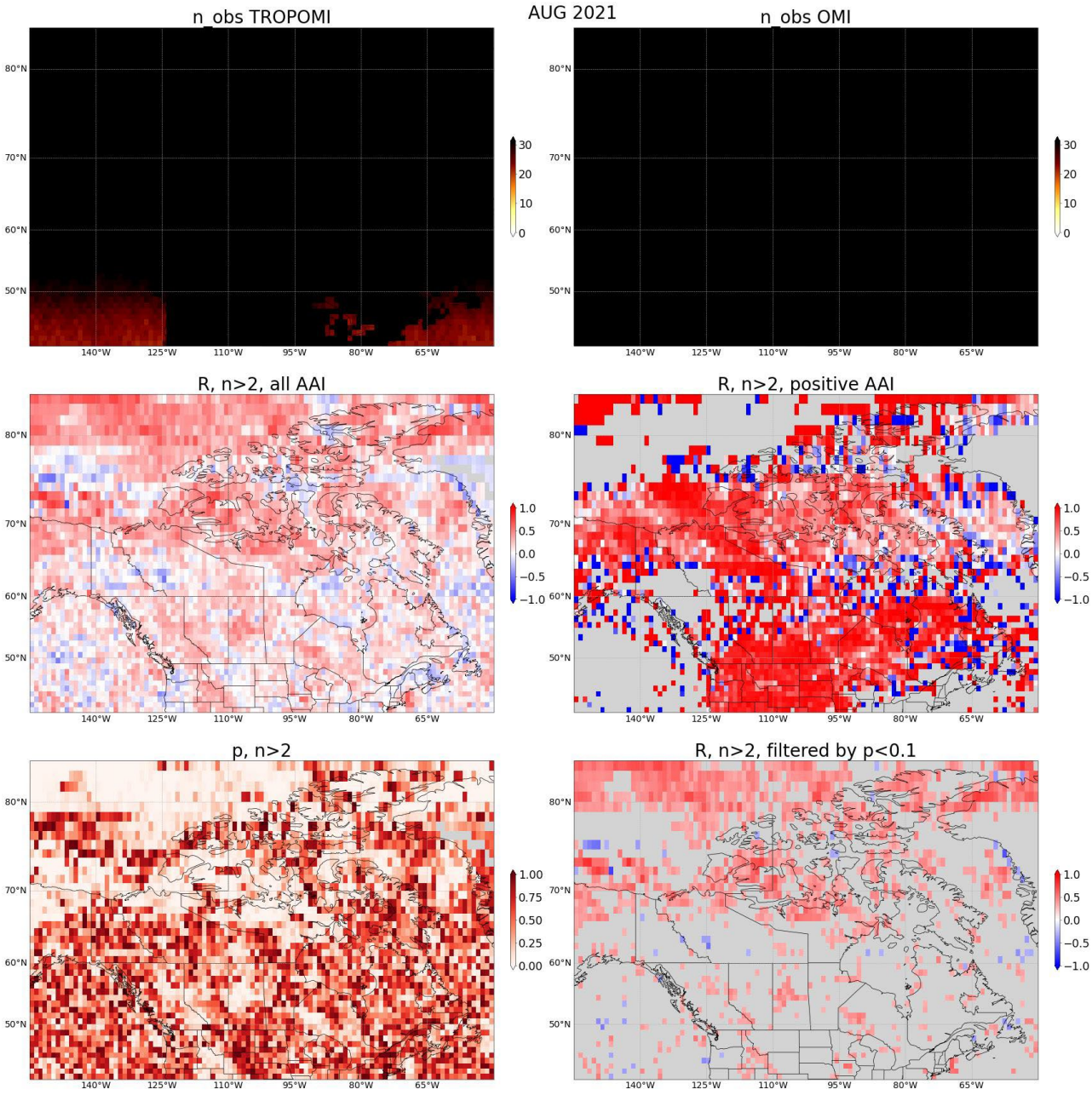

*Figure 16: Correlations between TROPOMI AAI and OMI UVAI datasets for August 2021. Top row shows the n_obs for TROPOMI (left) and OMI (right). (**Middle left**) – Correlation (R) between all TROPOMI AAI and OMI UVAI data. (**Middle right**) – Correlation for all OMI UVAI data and only positive TROPOMI AAI data. (**Bottom right**) – Correlation (R) of all TROPOMI AAI and OMI UVAI data, but only gridboxes corresponding to a significant p-value ($p < 0.1$) are kept. (**Bottom left**) – p-value for all TROPOMI AAI and OMU UVAI data. Grey coloring indicates NaN data.*

The right panel in the middle of Figure 16 shows that filtering TROPOMI for positive AAI increases the intensity of the calculated correlations. In fact, filtering for only positive AAI values increases the number of gridboxes with $R > 0.7$ from 0.23% to 47.6%. This is a significant

improvement from the first set of correlations with all AAI being included. Supported by the ATBD document and the discussion in section 1.2.1.1, filtering for positive AAI values strengthens the correlation with OMI AAI since TROPOMI can more readily detect high AAI even with clouds (Zweers, 2022). This filtering step helps ensure that a comparison of data less affected by cloud contamination is being made and that the data more likely represents true absorbing aerosols.

Despite the TROPOMI algorithm being adapted from the OMI algorithm, there is clear disagreement between the instruments in some cases. This is indicated by negative correlation shown on the map in Figure 16. If the two datasets were exactly measuring the same aerosol changes in the atmosphere, a uniform color map would be observed where all gridboxes have an R = 1, but this is not the case. There is a large difference in the spatial resolution of TROPOMI compared to OMI, where the OMI is significantly coarser. This means that during the analysis process, the TROPOMI data had to be binned on a coarser scale, meaning that more AAI data is being averaged between each gridbox than for the OMI data. In fact, 400 5 x 5 km^2 TROPOMI gridboxes were averaged into one 100 x 100 km^2 gridbox to match the spatial resolution of OMI L3 products (the actual footprint of OMI is 13 x 24 km^2). This is perhaps the reason why there is disagreement in some areas between the two instruments, i.e, having to do with AAI retrieval quality and robustness around clouds, which are sampled differently at different instrument measurement (native) resolutions.

Another important feature of Figure 16 is that when filtering for positive TROPOMI AAI values there is an increase in strong correlations, but many of these strong correlations are in areas where the *p*-value was found not to be significant. In a direct comparison to the middle and bottom right panels in Figure 16, it can be seen that most of the stronger correlations in the southern parts of Canada are not statistically significant at the $p < 0.1$ threshold (bottom right panel). However,

it is important to note that the *p*-values used to filter the correlation data in the bottom right panel were found using all AAI data, while the middle right panel shows correlations found using only positive TROPOMI AAI data. While the feature of missing data is comparable, the datasets used between the two are not. A further analysis could be done to filter for positive AAI and statistically significant *p*-values.

The last panel of Figure 16 shows that there are some statistically significant correlations between TROPOMI and OMI in northern Canada. However, there is a lack of strong and significant correlations across much of the rest of Canada. Additionally, the filtering approaches used and the difference in spatial resolution require further analyses to fully understand the agreement and disagreement between the two instruments. Overall, these findings suggest that TROPOMI shows reasonable AAI retrievals in agreement with OMI, notably also at the higher latitudes identified as potentially dust emitting (Figure 9).

Table 4: Percent gridboxes for OMI and TROPOMI correlation (all AAI data) for August 2021. Bottom row shows the percent of gridboxes for the correlation dataset that has been filtered by significant p-value ($p < 0.1$).

	Percent Gridboxes			
n_obs	Positive R	Negative R	R > 0.7	R < -0.7
> 2	72.6 %	27.4 %	0.23 %	0.03 %
> 10	72.6 %	27.4 %	0.23 %	0.03 %
31	72.6 %	27.4 %	0.23 %	0.03 %
R filtered by p < 0.1	92.9 %	7.45 %	0.98 %	0.11 %

Table 4 summarizes the distribution of gridboxes across various correlation coefficient (R) values for different threshold numbers of observations (n_obs). Notably, the percentage of gridboxes associated with each R value remains consistent even when n_obs reaches 31, as expected, since 31 observations are always present. The table further reveals that most gridboxes

show positive correlations, though the majority fall below an R of 0.7, indicating weak to moderate agreement between the two instruments. When the R dataset is filtered only for significant p values, the positive, strongly positive, and strongly negative correlations increase, while the total amount of negative correlations decrease. There is evidence therefore that some deeper analysis of the TROPOMI AAI data is required to ensure that its measurements are comparable to those made by OMI. The filtering tested in this study used for deeper analyses were positive AAI values and significant p values, but further testing or a combination of both of these might better reveal how well both instruments track aerosol changes in the atmosphere. The remaining disagreement likely stems from OMI's lower spatial resolution, causing TROPOMI gridboxes to average a wider range of AAI values and deviate from OMI's data, particularly in the presence of clouds. Alternatively, the discrepancy might be attributed in some part to TROPOMI's advancements in technology and retrieval algorithms compared to OMI. The OMI-TROPOMI correlation calculations were also only made over one month. This means that there are only 31 data points being used to calculate the correlation. It is possible that, if the correlation were to be made over a greater time span, say over the four-month summer period studied in the MODIS-TROPOMI correlation, the amount of strong correlations may increase. However, a deeper understanding of this disagreement falls outside the scope of this study.

3.3 RQ 3

3.3.1 CO

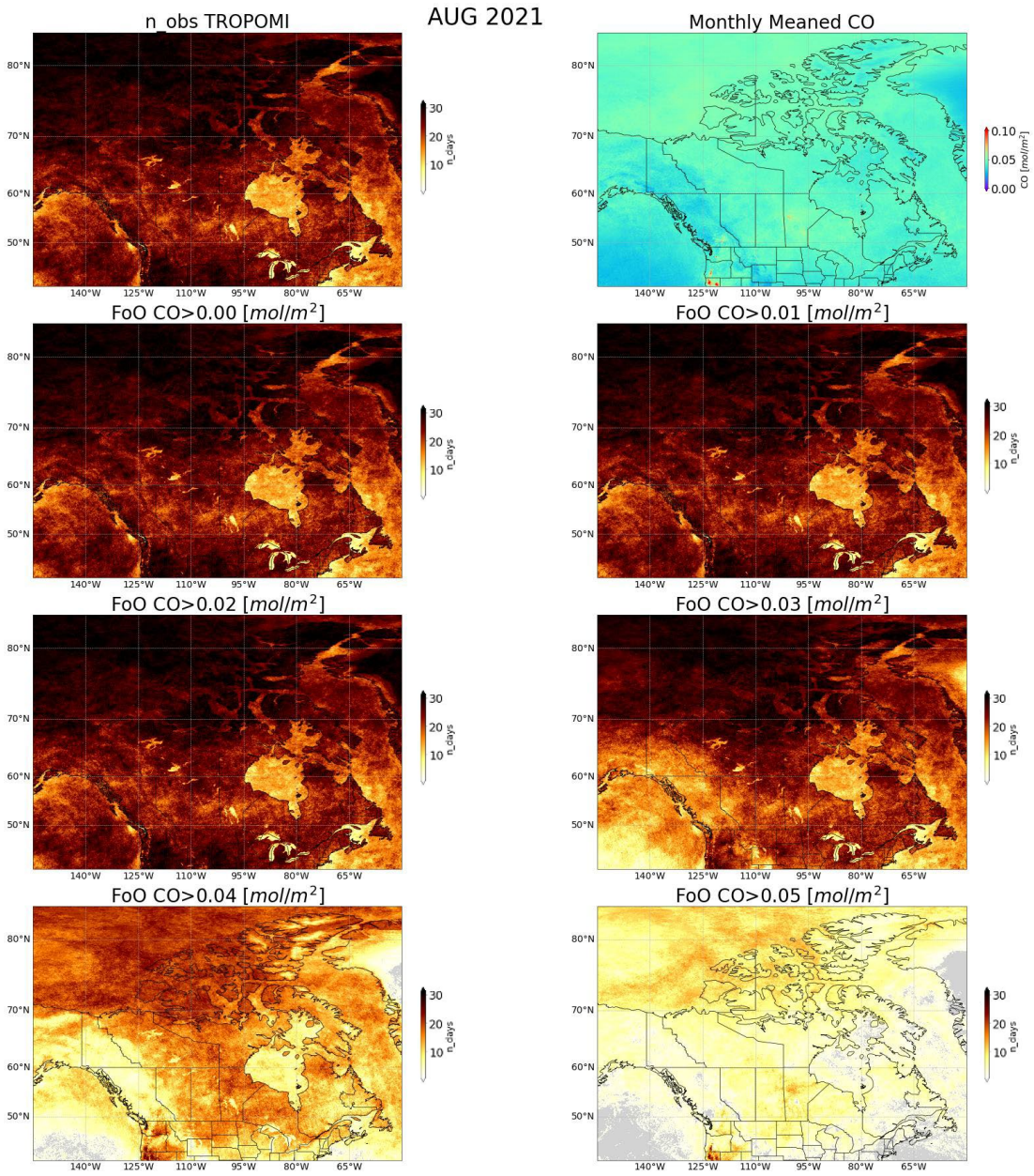

Figure 17: FoO maps showing the distribution of CO column density (measured in x 10^{18} [mol/m^2]) for August 2021. ***(Top left)*** *– n_obs for the TROPOMI CO data.* ***(Rows 2:4)*** *- FoO of CO above different thresholds.* ***(Top right)*** *- Monthly mean CO. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.*

Figure 17 illustrates the distribution of CO column density for August 2021. It specifically shows the FoO of CO for different CO thresholds, as well as the total n_obs (top left panel) and monthly mean (top right panel). The bottom six panels show that there are frequent amounts of

CO above ~ 0.02 x 10^{18} mol/m^2 while the frequency of even higher CO values reduces beyond that threshold. This is in line with a study by Wan, et al., 2023, which reports a CO value of 0.018 as "background CO". These figures also show, specifically in the first panel, that most of the gridboxes over land have an n_obs greater than 1 and close to 31, (depending on the number of days in the month). There is a significant reduction in frequency of CO as the threshold value increases. This is to be expected since high amounts of CO in an atmospheric column are less frequent due to the finite amount of fuel available to emit CO. Appendix C shows the same plots but for the rest of the summer months, June, July and September. The patterns observed in Figure 17 are consistent across all four months.

With the information presented in Figure 17, it was decided to use a threshold of CO 0.01 to begin the filtering of AAI data. The decision of what CO values to use to filter the AAI data depends on the desired outcome of the analysis. For the purposes of this study, several CO thresholds (ranging from 0.01 to 0.05) were used to filter the AAI data to understand how the CO dataset would impact the distributions of aerosols observed in the dataset. It was desired to use a CO value below the background value to compare and ensure that the literature value of 0.018 was consistent with the observations made in TROPOMI. What this means is the AAI dataset filtered by CO of 0.01 should leave little to no data in the AAI dataset after filtering. This is primarily because the n_obs for all CO (first panel in Figure 17) and the n_obs for CO greater than 0.01 (second row right panel) shows the same thing, meaning that all CO data is contained at a value greater than 0.01. It is expected to remove all AAI data because CO is below the accepted background value, meaning that the "normal" amount of CO that is naturally found in the atmosphere is accepted to be 0.018 mol/cm^2 so any CO values below 0.01 mol/cm^2 would likely represent measurement noise or other outliers, and not true atmospheric CO concentrations.

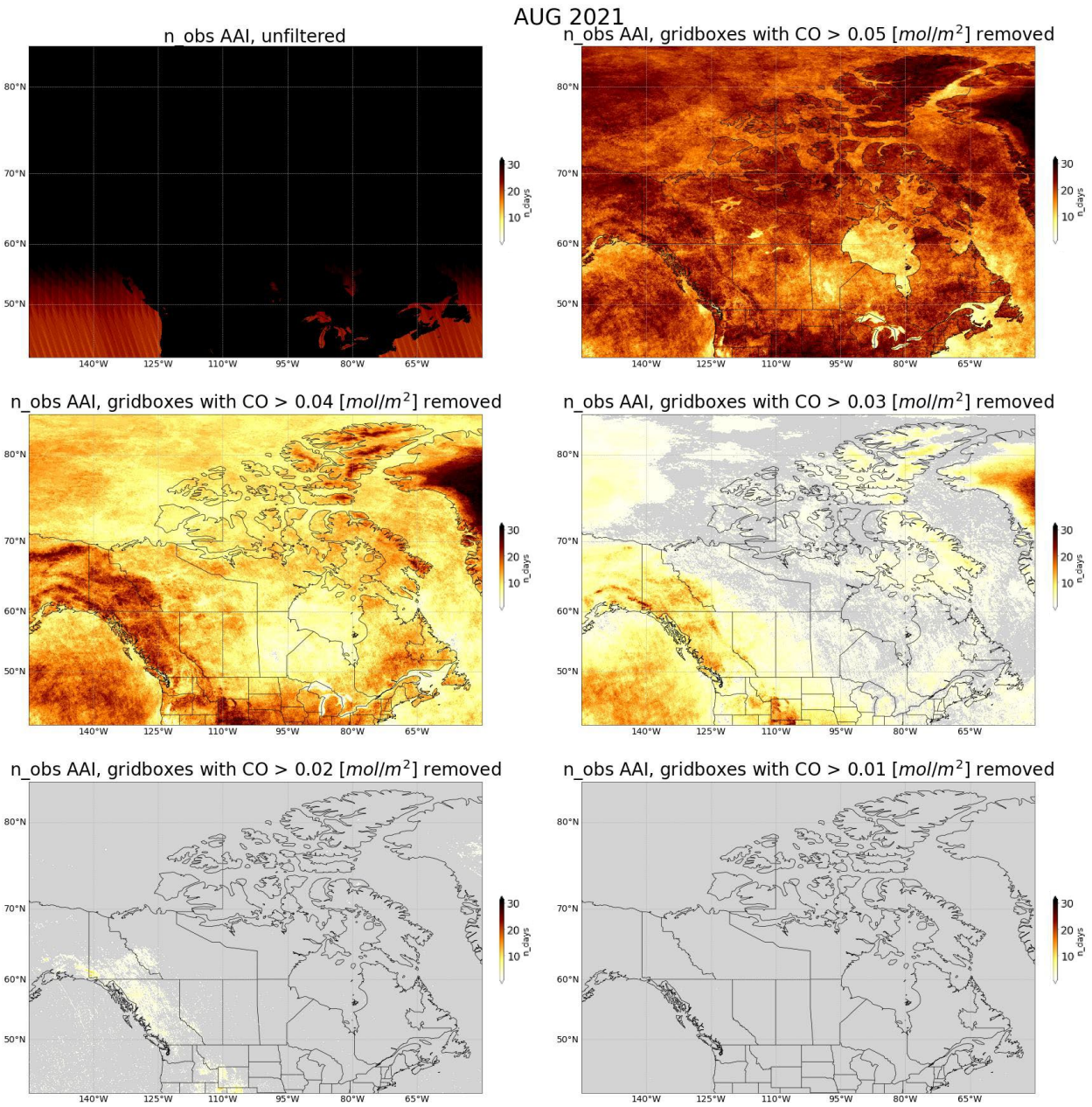

Figure 18: The changes in n_obs for August 2021 as the AAI data is filtered by CO data at different thresholds. (Top left) – n_obs for the unfiltered dataset. (Rest of panels) – n_obs for the AAI datasets that have been filtered by different values of CO. Gridboxes that did not have a CO value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum n_days is 1 and the maximum is 31.

The top left panel in Figure 18 shows the n_obs for unfiltered data. In subsequent panels, as the CO threshold increases, so does n_obs. The bottom right panel shows zero n_obs for AAI data where coincident CO values are also 0.01 or below, as expected based on the literature value for background CO of 0.018. The rest of the panels in Figure 18 show the n_obs increasing as the

threshold for CO increases. This increase in n_obs indicates that there are a lot of AAI data that are not coincident with high CO values, meaning that the observed AAI are unlikely due to forest fire absorbing aerosols (soot) associated with high CO gas values, i.e., confounding of dust.

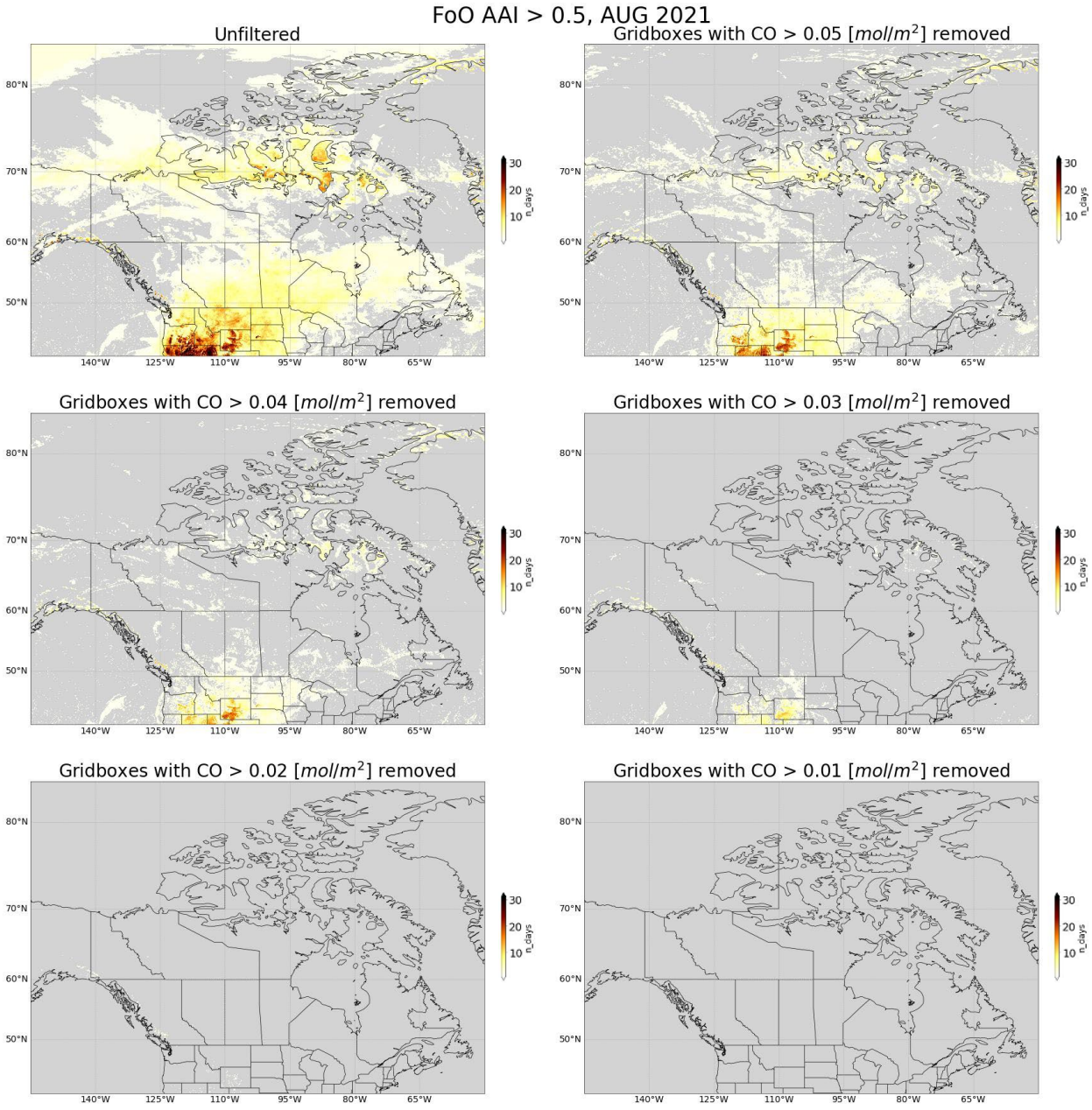

Figure 19: FoO of AAI greater than 0.5 for August 2021. Each panel is the same AAI dataset but filtered by different CO thresholds, except the top left which is the unfiltered AAI dataset. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.

Figure 19 illustrates the distribution of FoO of AAI > 0.5 for different thresholds of CO applied to filter the AAI data. The last panel in Figure 19 shows an entirely grey map, which means that there are no distributions of AAI above 0.5 that also have a coincident CO value of 0.1 or below. This, again, is to be expected from Figure 18. As the CO threshold increases, increasing areas with frequencies of occurrence of AAI appear on the maps, with increasing FoO values. For example, looking at the region near 110W and 50N, as more CO data is kept, more frequently we find AAI above 0.5 in that region. This indicates that those distributions of AAI are areas with low concentrations of CO and are therefore more likely to be MDAs instead of urban or biomass burning absorbing aerosols like soot. Looking at the northern part of Canada, around 70N and 110W to 80W, there are frequencies of occurrence of AAI above 0.5, even after gridboxes with high CO were removed. This indicates that those regions show distributions of absorbing aerosols that are not associated with CO, meaning they are not smoke or soot and are therefore more likely to be MDAs. Removing data from the AAI dataset where there are high coincident CO values effectively eliminates distributions of AAI data that are unlikely to be MDAs, since MDAs are not co-emitted with CO. The general effect of the CO dataset on the AAI dataset is that it aids in filtering for different types of absorbing aerosols, which the AAI product does not distinguish alone. It helps in understanding the distribution of AAI data by removing hotspots with a lot of CO that are likely to be forest fires or perhaps transported dust mixed with CO, effectively removing areas that would not be the dust source areas (i.e., Figure 9) of most direct interest to this study.

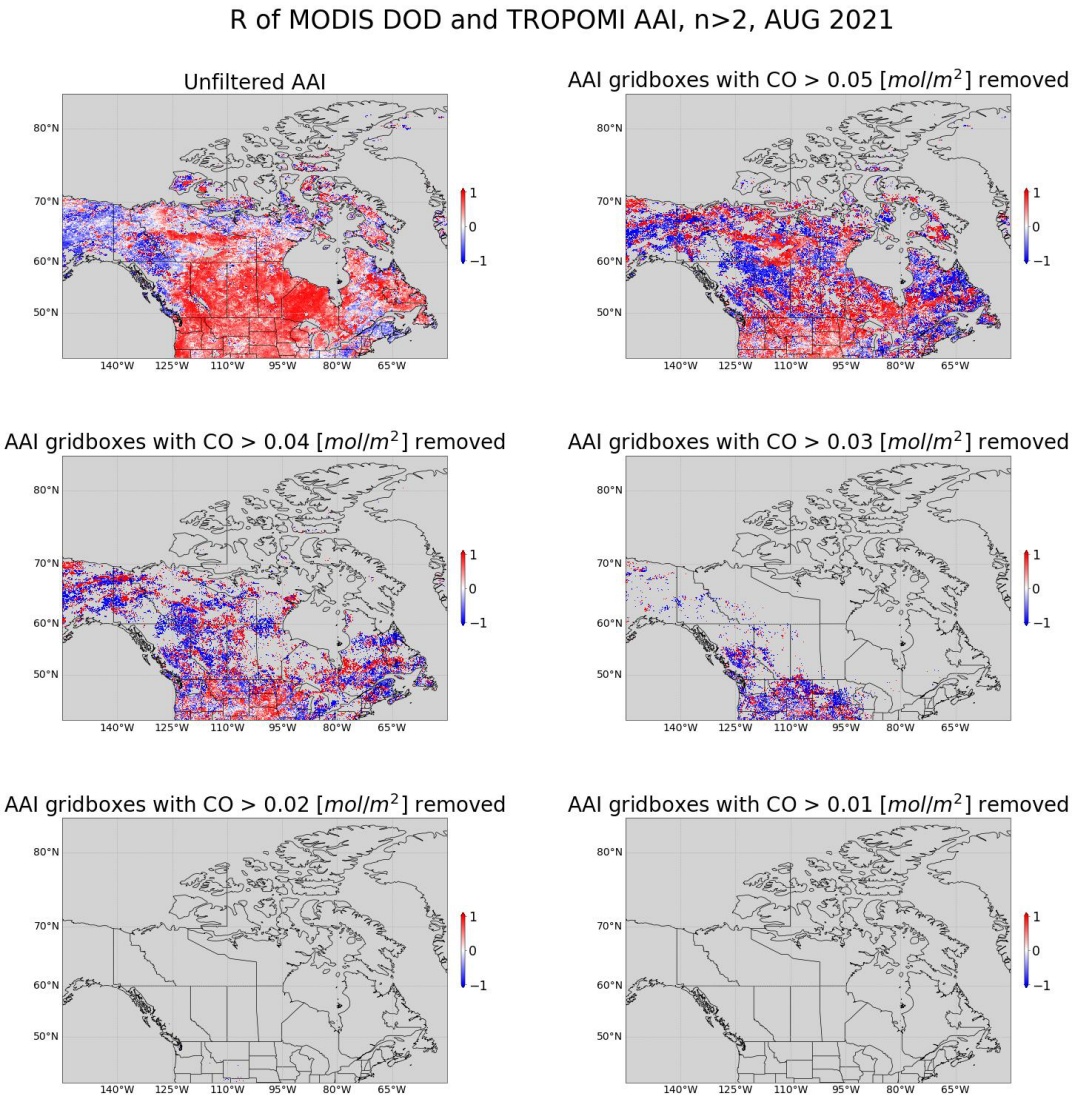

Figure 20: Correlation (R) values for August 2021 between TROPOMI AAI and MODIS DOD. The TROPOMI AAI data has been regridded to match the 10 x 10 km² resolution of MODIS DOD data. The AAI data has also been filtered by different values of CO. Grey areas indicate areas where there is no data, meaning that no correlation could be calculated for those gridboxes.

Figure 20 shows the correlation between TROPOMI and MODIS aerosol data. The bottom two panels show nearly no gridboxes where correlations could be calculated, since they have been filtered to remove any AAI data that is coincident with CO data near or below the background value. The bottom left panel, in reality, has only 64 gridboxes with available data, which is essentially negligible with the total number of gridboxes of 3,240,000. The rest of the panels show significant variations between red and blue coloring, indicating heterogeneously distributed positive and negative correlations. This indicates that, while there is agreement between the

instruments in some locations, there is also disagreement. To better understand how CO filtering impacts the correlation between MODIS and TROPOMI, we look at the specific numbers of strong and significant correlations that remain before and after filtering. However, it is a limitation of this study that the correlation calculations made are severely lacking in the area of interest (i.e. northern Canada).

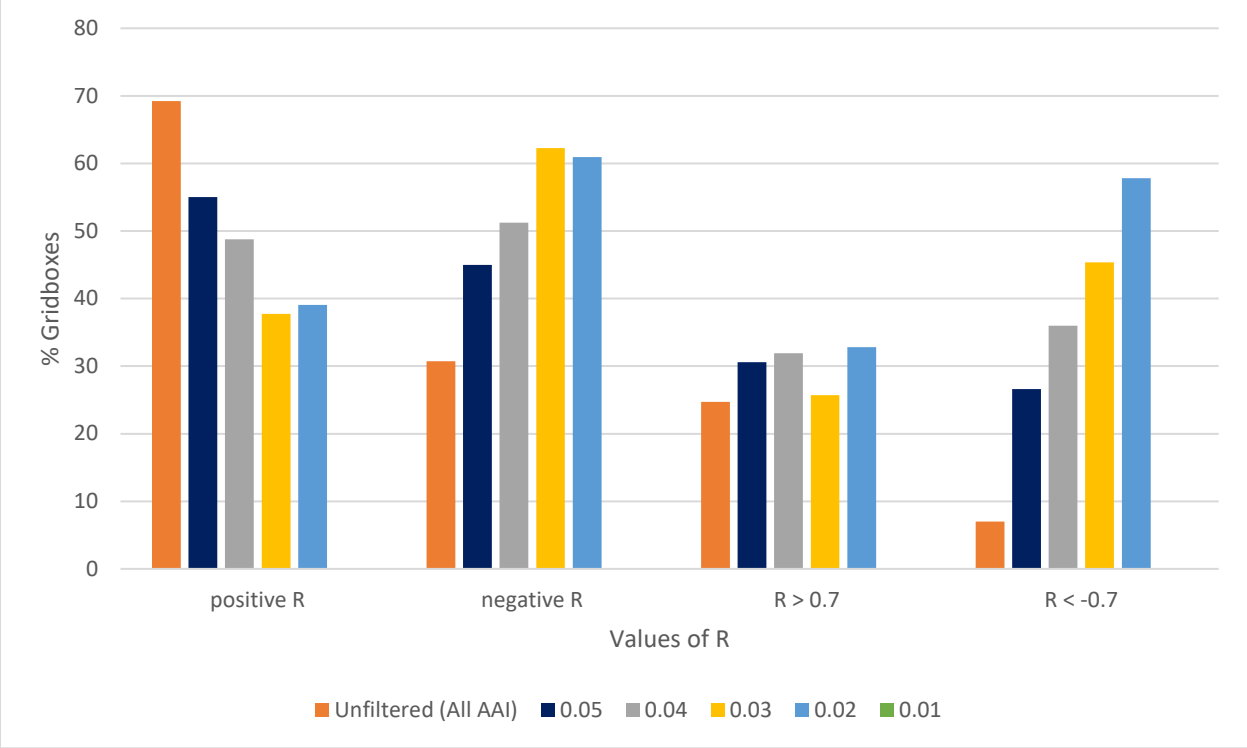

Figure 21: Graph showing the changes in percentages of gridboxes for the different AAI datasets plotted in Figure 20. The AAI datasets are where CO has been used to filter the data and the unfiltered AAI data (orange bar). Each color represents a different CO filter that had been used on the AAI data. An n>2 used in each dataset indicates a minimum of 2 data points (i.e. two days) were used in the correlation calculations for each gridbox.

Figure 21 shows the variation in the percentage of gridboxes for the differently filtered AAI datasets. The dataset where gridboxes with a CO value greater than 0.01 were removed had 0 gridboxes remaining, hence why it is not seen on the graph. For the datasets filtered by CO of 0.03, 0.04, and 0.05, there are more negative correlations than positive ones. This observation can also be made by looking at Figure 20 and observing the large amounts of blue on the maps in panels in the last two rows. While the dataset with CO filtering of 0.02 shows similar patterns as

the other three datasets, it is important to note that there is a significant lack of gridboxes for this dataset compared to the others (i.e. 64 total gridboxes compared to 23945 or 89344 gridboxes). This analysis shows that filtering the AAI dataset does not improve the correlation between MODIS DOD and TROPOMI AAI, as anticipated.

Figure 21 also shows the dataset where CO greater than 0.05 has been removed from the AAI data. It can be seen that this dataset has more positive correlations than negative ones, indicating a better agreement between the two instruments compared to the other CO filters. This is also the dataset where the least AAI data has been removed, meaning that there is more AAI data available to make correlations. This is indicated in Figure 20, as well as by the decrease in grey areas. The fact that there is better agreement in this dataset indicates that filtering CO data and thereby removing certain AAI hotspots does not necessarily increase the correlation between the two instruments. The orange bar in Figure 21 represents the AAI dataset that has not been filtered by any CO. This dataset has the greatest percentage of positive gridboxes and the lowest amount of negative gridboxes. This means that the greatest amount of agreement between MODIS and TROPOMI is found in the unfiltered dataset. This also aligns with the results in section 3.2.1 where the strongest correlations are observed when the greatest number of observations are used. In the case of the filtered AAI datasets, retrievals for individual grid boxes are removed if they do not meet the CO filtering criteria. This effectively reduces the number of observations used for comparison within those grid boxes, thereby decreasing the observed agreement between the two instruments.

The lack of improved agreement between the two instruments from the filtered datasets compared to the unfiltered datasets, as seen in Figure 20, has several implications. First, the filtering may not be effectively removing all fire aerosols from the AAI data. While CO is often a

good indicator of fires, it is not a perfect one in this case. Other fire emissions, like organic matter or black carbon, might still be present in the AAI data and contribute to the lack of increased correlation with MODIS DOD. It is also possible that factors unrelated to fire or dust aerosols are influencing the datasets. This may be due instrument calibration issues, retrieval errors, or the presence of other aerosol types not captured by the filtering or DOD measurements. While the lack of increased correlation after CO filtering was not anticipated, it was based on a logical assumption of correlation between non-MDA absorbing aerosols like soot and CO. It indicates the need to consider the limitations of the filtering approach in detail, and the potential dominance of other aerosol types in the analysis, which are beyond the scope of this study.

3.3.2 CLO

Figure 22 shows the different FoOs for CLO for August 2021. In most areas, frequency of occurrence decreases as the CLO threshold increases. This indicates that days with gridboxes covered by high percentages of clouds are less frequent. The first panel in Figure 22 shows the total n_obs for the month. The nearly uniform dark color indicates that most gridboxes had observations for each day of the month. The n_obs distribution is almost entirely over land in the higher latitudes, besides around 75N and 140W. The lack of observations over the water in these regions indicates that the CLO collected over water in higher latitudes either does not exist or did not pass the quality filtering when gridding the data. This makes sense because remote sensing over water with large amounts of sea ice can be confounding for remote sensing instruments. The TROPOMI CLO product does collect data over oceans but increasing sun glint from sea ice is high, especially in the summer months. Appendix C, showing the same information as Figure 22 but for the other three months, shows the same phenomenon with lack of data over oceans.

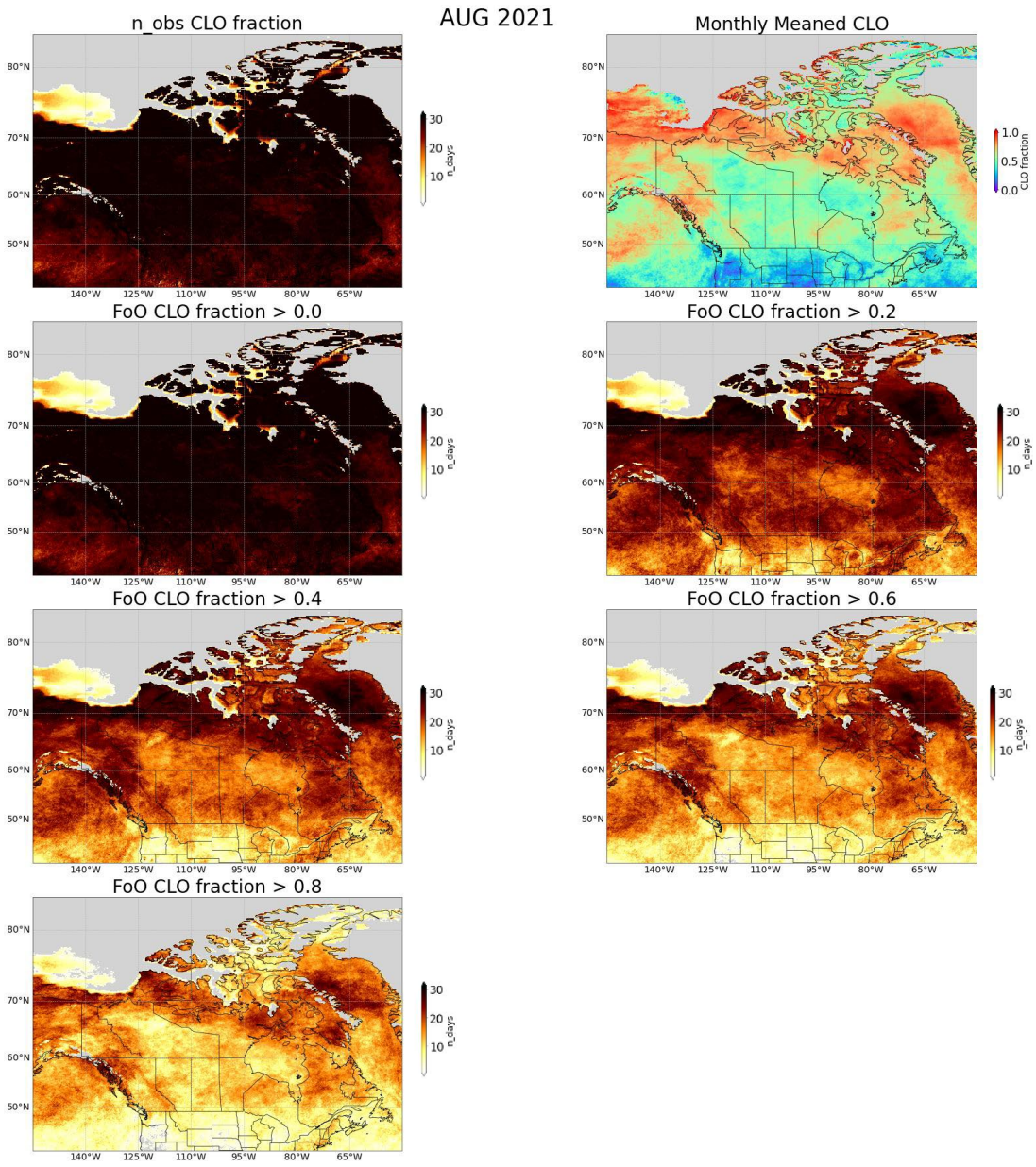

*Figure 22: The changes in n_obs for August 2021 as the AAI data is filtered by CLO data at different thresholds. (**Top left**) – n_obs for the unfiltered dataset. (**Rest of panels**) – n_obs for the AAI datasets that have been filtered by different values of CLO. Gridboxes that did not have a CLO value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum n_days is 1 and the maximum is 31.*

The top right panel in Figure 22 shows the monthly mean CLO for August 2021, which varies significantly in space. The regions where the mean CLO is closer to 1 are cloudier than regions with a lower cloud fraction. Average CLO around the 70N band is fairly high (> 0.7). It can therefore be inferred that a lot of the AAI data around that band may be obscured by clouds,

which may be confounding for the AAI dataset since it does not clear clouds, but rather exhibits values near and below zero for clouds (Zweers, 2022).

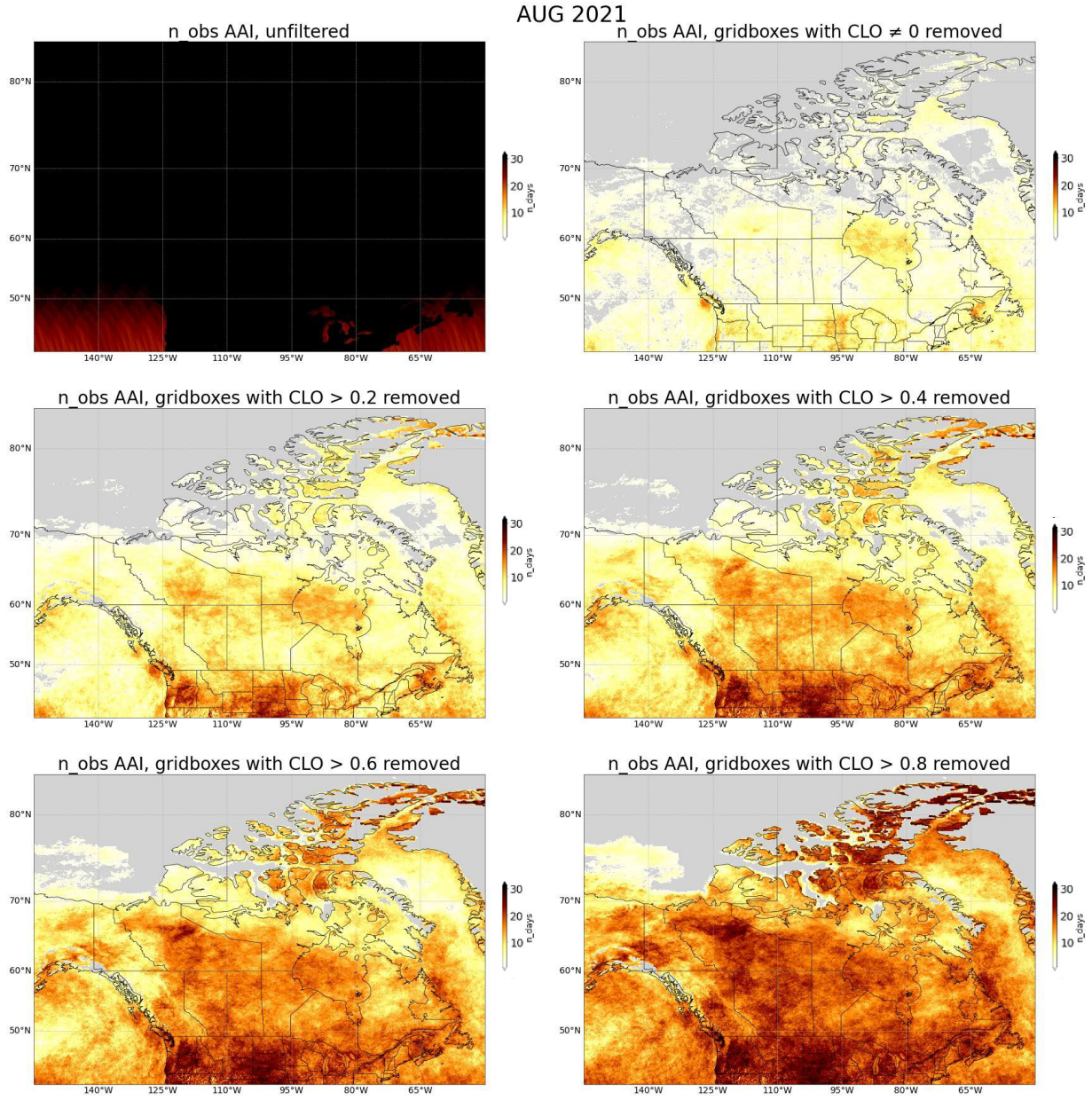

*Figure 23: The changes in n_obs for August 2021 as the AAI data is filtered by CLO data at different thresholds. (**Top left**) – n_obs for the unfiltered dataset. (**Rest of panels**) – n_obs for the AAI datasets that have been filtered by different values of CLO. Gridboxes that did not have a CLO value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum n_days is 1 and the maximum is 31.*

The first panel in Figure 23 shows the unfiltered n_obs for the month of August 2021. As the CLO threshold increases, the n_obs also increases. This is because more gridboxes are being

included with the higher CLO threshold. The right panel in the top row shows the n_obs where CLO not being 0 is removed. This means that any gridboxes with any cloud contamination have been removed from the AAI data. This panel is the one with the least n_obs, which is to be expected because the most data has been removed from that AAI dataset. Throughout the rest of the panels there is a significant frequency of observations in the northern part of Canada. Looking specifically at the second and third panels, many of the regions in northern Canada, say around (65W, 85N), frequently have non-NaN data. This indicates that the distributions of AAI data in that region are not just from cloudy observations, therefore indicating the likelihood of the presence of other kinds of aerosols, particularly when AAI > 0.

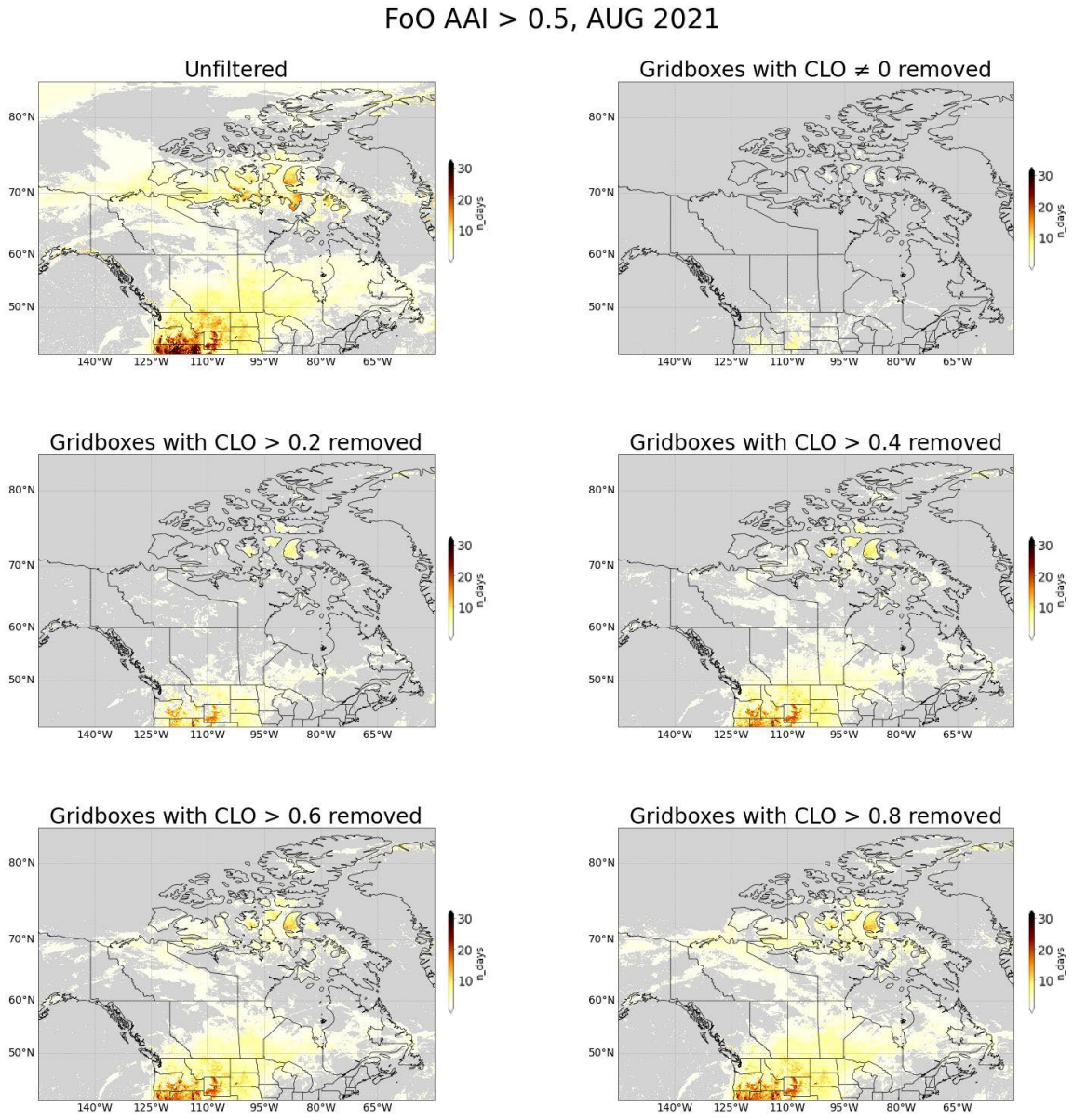

Figure 24: FoO of AAI greater than 0.5 for August 2021. Each panel is based on the same AAI dataset that has been filtered by a different CLO threshold, except the top left which is the unfiltered AAI dataset. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.

The first panel in Figure 24 shows the FoO AAI above 0.5 that is unfiltered by CLO data. The top right panel shows the FoO AAI above 0.5 where only gridboxes with no cloud contamination are kept. Despite a lot of "NaN" data, there are some regions in northern Canada that have frequencies of occurrences, and they correspond to those highlighted in Figure 8. Looking specifically at 80W – 95W and 70N, there are regions with AAI above 0.5 for multiple

days over the month. This indicates that those AAI distributions are not from clouds as the only gridboxes kept in this second panel are gridboxes with a cloud fraction of 0.

As the cloud fraction threshold increases, more frequent observations are noted, indicated by the increased number of gridboxes and increased color. This is particularly noticeable around 125W and 50N where more frequent observations are made as more cloud-covered retrievals are kept. This increase in FoO across CLO thresholds indicates that the observed distributions of AAI data may be from clouds rather than absorbing aerosols. The fact that, with a cloud fraction of 0.4 or 0.2 kept only (i.e. middle row of Figure 24), there are still distributions of AAI data in northern Canada indicates that those aerosols are more reliable indicators of non-scattering aerosols or clouds. This is an important observation as the AAI data does not have a cloud-clearing feature, but using the CLO data to remove gridboxes with high CLO values is essentially manually creating a cloud-clearing. This is an important step in the analysis as it aids in removing distributions of AAI data that are contaminated by clouds and therefore less reliable indicators of MDAs, therefore allowing for closer examination of regions left with low cloud coverage and frequent AAI values. As previously mentioned, clouds can be confounding for satellite instruments, especially for the AAI product. It is known that AAI values are easily detectable in the presence of clouds if the AAI is greater than 1 (Zweers, 2022), but removing cloudy data from the AAI datasets further adds a level of confidence in the measured AAI values that are indicative of MDAs (AAI > 0.5), but are less reliable in the presence of clouds.

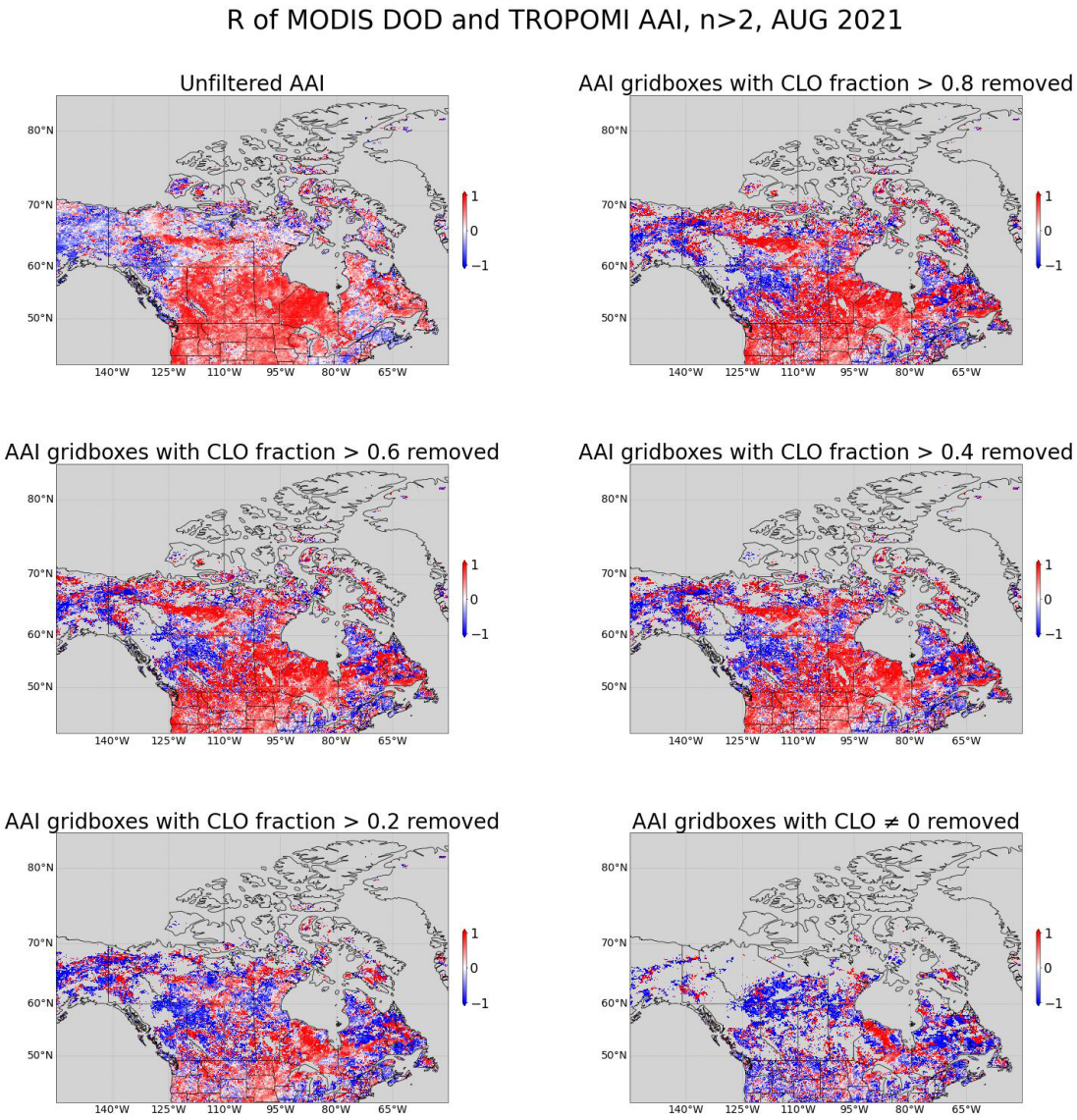

Figure 25: Correlation (R) values for August 2021 between TROPOMI AAI and MODIS DOD. The TROPOMI AAI data has been regridded to match the 10 x 10 km^2 resolution of MODIS DOD data. The AAI data has also been filtered by different values of CLO. Grey areas indicate areas where there is no data, meaning that no correlation could be calculated for those gridboxes.

Figure 25 shows the correlation map between TROPOMI AAI and MODIS DOD where the AAI data has been filtered by the cloud data. It can be seen that throughout all panels, there tends to be a spatial heterogeneity of positive and negative correlations. The coloring across all four panels also appears to be intense, indicating that there are many, strong positive or negative correlations. While there is a lack of data availability in the northern latitudes, there appear to be mainly red, therefore positive, correlations in the higher latitudes. This indicates that there is

potential agreement between the instruments when CLO data has been used to remove gridboxes where there are known clouds in the AAI data. Next we quantify the changes in correlation and the precise numbers of strong and significant correlations when the AAI data has been filtered by CLO.

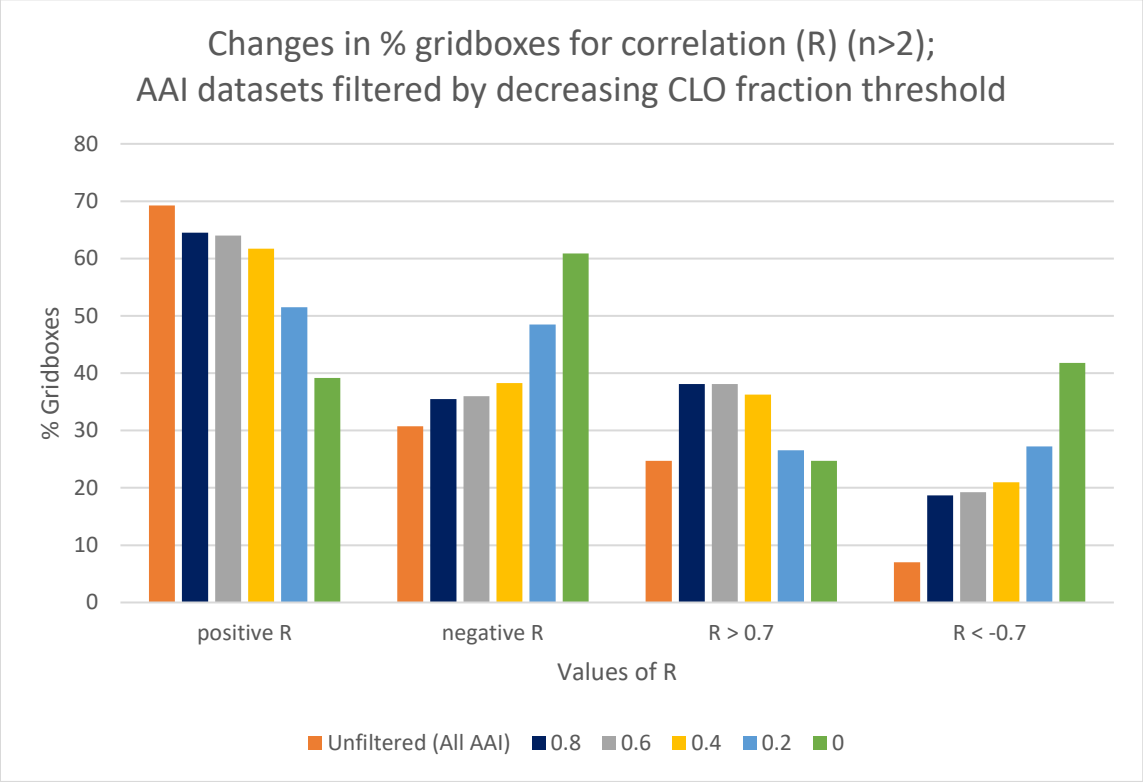

Figure 26: Graph showing the changes in percentages of gridboxes for the different AAI datasets. The AAI datasets are where CLO has been used to filter the data and the unfiltered AAI data (orange bar). Each color represents a different CLO filter that had been used on the AAI data. An n>2 used in each dataset indicates a minimum of 2 data points (i.e. two days) were used in the correlation calculations for each gridbox.

Figure 26 shows a bar graph representing the changes in percentages of gridboxes for different values of R. The different colors represent the different CLO thresholds used to filter the AAI data. In most cases, there are more positive correlations than negative ones, indicating overall agreement between the two instruments. The exception is the dataset where only cloud-free gridboxes were kept, shown by the green bar in Figure 26. The greater amount of negative correlation existing for this dataset implies that filtering for purely cloud-free scenes is removing

instances of aerosol detection from the AAI dataset that are mixed with cloud-contaminated pixels, thus leading to the mismatch in measurements in aerosol properties between the two instruments. It can also be seen that as more gridboxes with a higher percentage of cloud coverage are kept in the AAI data, the greater the percentage of gridboxes that are greater than 0.7. This is seen in Figure 26 where, from right to left (meaning lower CLO threshold to greater CLO threshold), the bar gets larger, indicating a higher percentage of gridboxes. This R value also indicates an R^2 greater than 0.5. This therefore implies that as more gridboxes are kept, the greater the amount of strong R^2 values, therefore implying that those datasets with less cloud clearing are a better representation of the aerosol distributions, in relation to MODIS DOD data. The fact that the orange bar, indicating the unfiltered dataset, is largest for positive correlations suggests that the most agreement between MODIS and TROPOMI is seen when the AAI data is unfiltered.

The increase in positive correlations from CLO 0.2 to CLO 0.8 (12.5%), meaning an increased amount of gridboxes are kept with a cloud fraction up to 80%, suggests that more AAI data, regardless of the cloud cover, improves the correlation and therefore the agreement between MODIS and TROPOMI. This means that including more cloud-covered data may increase the chance of capturing these positive relationships. The increase in positive correlations and decrease in negative correlations as less AAI data is removed indicates that the filtering of AAI data by CLO data does not increase its correlation to MODIS DOD. As with the CO results, this is inline with the results in section 3.2.1 where more observations lead to more positive correlations.

Filtering the AAI data by cloud data has unexpected consequences for the correlation between TROPOMI AAI and MODIS DOD. Clouds can act like a physical curtain that blocks the view of the Earth's surface from satellites, including areas with high concentrations of dust aerosols that MODIS DOD is designed to detect. When filtering out these cloud-covered gridboxes from

the AAI data, it may be disregarding datapoints where AAI is below 1 (i.e. where AAI values are easily detectable in the presence of clouds (Zweers, 2022)) and therefore contaminated by clouds and not easily detectable. This can lead to underestimating the actual positive correlation between dust aerosols and the total aerosols measured by TROPOMI.

3.3.3 Implications in northern Canada

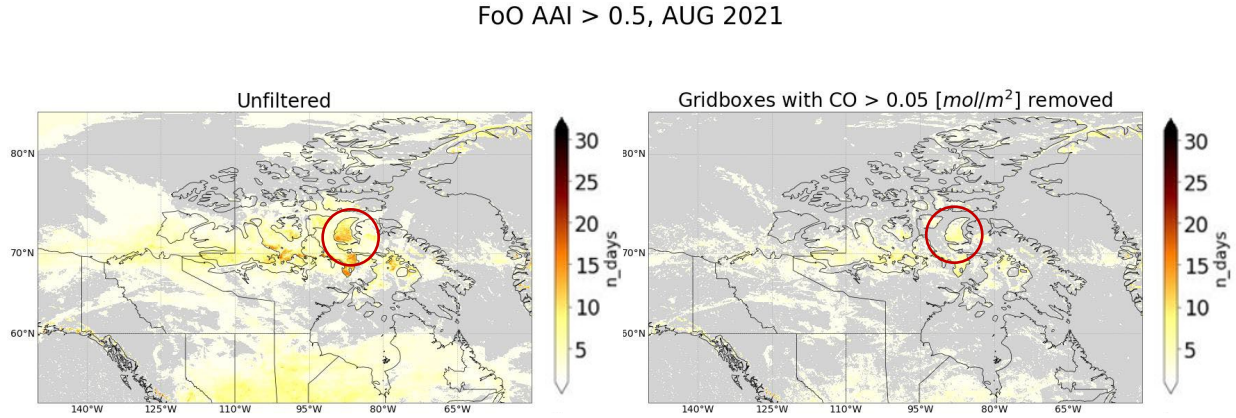

*Figure 27: Zoomed-in FoO of AAI greater than 0.5 for August 202. (**Left**) - unfiltered AAI dataset. (**Right**) – AAI dataset filtered by CO greater than 0.05. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0. Red circle shows the same area discussed in section 3.1.2 with low vegetation cover (see Figures 8 and 9).*

Figure 27 is a zoomed in map of the unfiltered AAI dataset compared to the AAI dataset where CO > 0.05 has been removed. This figure shows the FoO of AAI greater than 0.5. In both cases, there are clear observations of AAI values for multiple days over the month. Looking at the right panel, there are places where aerosol loadings exist for multiple days that have an AAI greater than 0.5, indicating that these distribution of AAI values are due to aerosols that are not coincident with high CO. This is especially important to note as a CO of 0.05 (around three times the background value of CO ($0.018 \times 3 = 0.054 \sim 0.05$)), does not significantly alter the FoO distribution of AAI greater than 0.5, implying that these observed absorbing aerosols are not strongly associated with smoke aerosols. The FoO of AAI circled in red in both panels is the same area within the Arctic Archipelago that had low vegetation coverage and high FoO of AAI close to the surface. This is significant as it further indicates that the results in section 3.1.2 are showing

MDA distributions and may be originating from a close by source and are unlikely to be linked to absorbing aerosols mixed with high CO concentration.

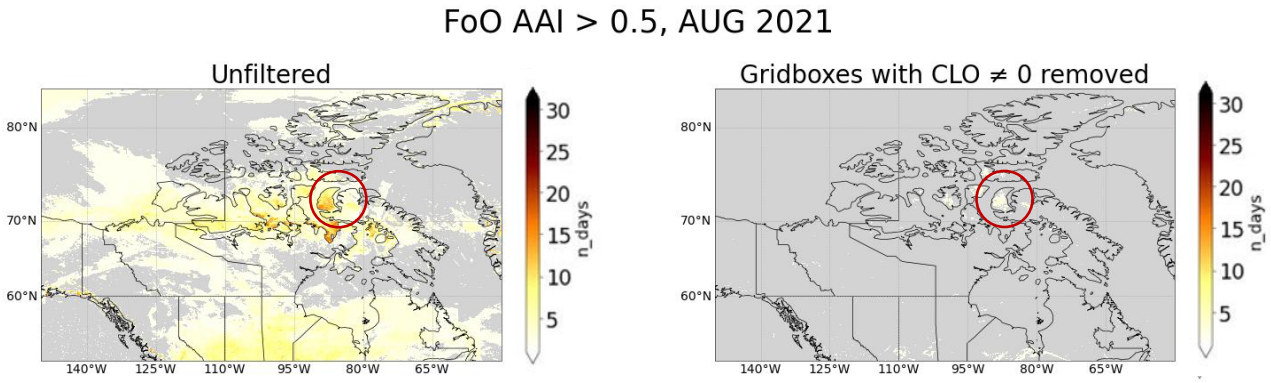

*Figure 28: Zoomed-in FoO of AAI greater than 0.5 for August 202. (**Left**) - unfiltered AAI dataset. (**Right**) – AAI dataset filtered by CLO equal to 0 (i.e. cloud free pixels). NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0. Red circle shows the same area discussed in section 3.1.2 with low vegetation cover (see Figures 8 and 9)*

Figure 28 is a zoomed in map of the unfiltered AAI dataset compared to the AAI dataset where CLO not equal to 0 has been removed, meaning only cloud-free gridboxes were kept. This figure shows the FoO of AAI greater than 0.5. It can be seen that in both figures, there are frequent observations of AAI greater than 0.5. Notably, the same area in the Arctic Archipelago (circled in red) has recurring observations of AAI greater than 0.5. Since the right panel shows the AAI data where only cloud-free scenes have been kept, it is unlikely that these AAI observations are showing clouds. This is significant because the TROPOMI AAI can be calculated in the presence of clouds, but clouds can also block absorbing aerosols that are below it. So, the AAI data left are not obscured by clouds and are therefore likely only absorbing aerosols, thus showing the presence of MDAs. This further supports the results found in section 3.1.2.

4. Conclusions

4.1 RQ1

The results in section 3.1 demonstrate the distributions of aerosols in northern Canada, according to TROPOMI's aerosol products. These distributions are not as frequent or absorbing as other areas (like BC or Alberta) but there is still evidence in these maps that there is aerosol activity there. Some of the distributions are in line with already researched areas, but in some cases, there is evidence that there may be dust activity in other parts of northern Canada. So, according to TROPOMI, the distribution of dust in northern Canada appears to suggest that there is aerosol activity in unexplored or unidentified areas.

When filtering by ALH, there continues to be AAI values of interest in northern Canada. As ALH decreased (i.e. got closer to the surface) it was observed that the corresponding AAI values became less frequent. However, there was still some evidence of aerosol activity in northern Canada at an ALH below 1000 m. This indicates that those aerosols are close to the Earth's surface and therefore are unlikely to have been transported across any large distance. So, the ALH filtering implies that those areas with frequent aerosol activity and that are in close proximity to the Earth's surface are more likely to be aerosols originating in those areas, rather than aerosols that have been transported from somewhere else.

4.2 RQ2

4.2.1 MODIS

The correlation plots and statistical information in section 3.2.1 showed that MODIS and TROPOMI agreed as the number of minimum observations used in the correlation calculation (n) increased. The results also showed that strong correlations (i.e. R^2 greater than 0.5) became more frequent as n increased. When filtering by positive TROPOMI AAI values, the correlations became

even stronger. This observation is to be expected because MODIS rejects radiance data when it is contaminated by high cloud coverage and radiances have been analyzed for dust-specific aerosols. So, MODIS DOD data and TROPOMI AAI data demonstrate stronger correlations with positive AAI values.

4.2.2 OMI

Similar to the results found for MODIS-TROPOMI correlations, OMI and TROPOMI showed overall agreement, based on the results in section 3.2.2. It was found that the two instruments agreed for the majority, with positive correlations accounting for ~70% of all gridboxes. There were some phenomena of disagreement observed that were unexpected. While the disagreement in some areas was unexpected and unexplainable within the scope of this study, most of the gridboxes showed agreement between the two instruments. It can therefore be said that OMI and TROPOMI demonstrate mostly positive correlations and therefore the two instruments moderately agree.

4.3 RQ3

4.3.1 CO

The CO product removed certain instances of observed aerosols distributions when the AAI data was filtered by it. The results showed that as the threshold for CO increased, the significant hotspot areas of AAI became less frequent (i.e. the banding at 70N or southwest Canada). This result is indicative of absorbing aerosols being coincident with high concentrations of CO, which is in turn associated with aerosols that are not of interest in this study (e.g., soot, co-emitted with CO in fires). In essence, the effect of the CO product is that it aids in removing aerosol hotspots that are unlikely to be MDAs, which are the interest of this study. It was found that FoO

of AAI with CO filtering persists in the same areas as found with RQ1, therefore solidifying the idea that MDAs exist in areas previously unstudied in northern Canada (Figure 8).

When analysing the impacts of CO filtering on the correlation to other notable instruments, it was seen that CO filtering did not increase the correlation between MODIS and TROPOMI as anticipated. While CO is a good indicator of fires, other fire-related aerosols could still be present in the AAI data, leading to a minimal increase in correlation with DOD.

4.3.2 CLO

The CLO product removed gridboxes from the AAI data that were likely to be measurements of clouds mixed with MDAs, or pure clouds. As the threshold for CLO increased, many of the hotspot frequencies decreased. However, it was also observed that many of the high latitude regions retained their FoOs, even when only cloud-free scenes were kept. This indicates that the aerosols observed in those locations are not clouds and are therefore likely to be other, non-scattering aerosols (like MDAs).This filtering also showed the same frequent observations of AAI values above 0.5 in the same Arctic Archipelago regions (Figure 8) as found in RQ1, therefore continuing to support the likelihood of the aerosols being MDAs and not clouds or CO-related aerosols.

Examining the correlation of the filtered AAI data to MODIS DOD revealed that the CLO data filtering does not increase the correlation between the two instruments. Similar to CO filtering, CLO filtering might not be very effective in isolating mixed or obscured aerosols. The retrieval process for AAI might not directly correspond to the parameters measured by CLO, leading to a limited impact on the overall correlation with MODIS DOD.

4.4 Relevance and Future Work

The results in this study highlight the distributions of aerosols in northern Canada. The study used TROPOMI data to understand aerosol patterns in high latitudes. It was found that, according to TROPOMI, there are distributions of absorbing aerosols that appear to be close to the surface of the Earth in northern Canada (Figure 8), indicating possible MDA sources in those areas. The information on aerosol distributions found in this study paves the way for a few avenues of future research. Efforts could be directed towards verifying and characterizing these potential MDA sources. This might involve deploying instruments for ground-based measurements within the identified areas to directly confirm the presence and characteristics of the sources, such as dust emission and composition. Developing atmospheric models to include dust emission from these regions could also offer valuable insights into the emission and deposition patterns of these aerosols, therefore accurately assessing the potential climate impacts discussed in section 1. By building upon the current study and pursuing these suggested avenues of future research, there is a possibility of gaining a more comprehensive understanding of aerosol distributions and their potential climate and air quality impacts in northern Canada. This knowledge could ultimately inform strategies for environmental monitoring and climate change prevention practices.

Appendices

Appendix A – Section 1

No appendix materials from Section 1 (Introduction) were included.

Appendix B – Section 2

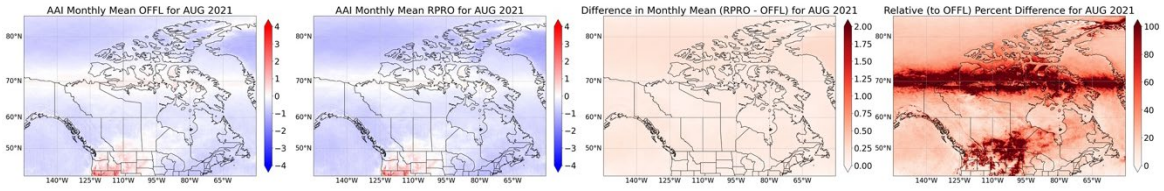

Figure 29: Differences in RPRO TROPOMI AAI data and OFFL TROPOMI AAI data for August 2021. (Left) – Monthly mean AAI for OFFL. (Middle left) – Monthly mean AAI for RPRO. (Middle right) - Difference in monthly mean AAI RPRO data and OFFL data. (Right) – Relative percent difference (to OFFL).

Appendix C – Section 3

3.1

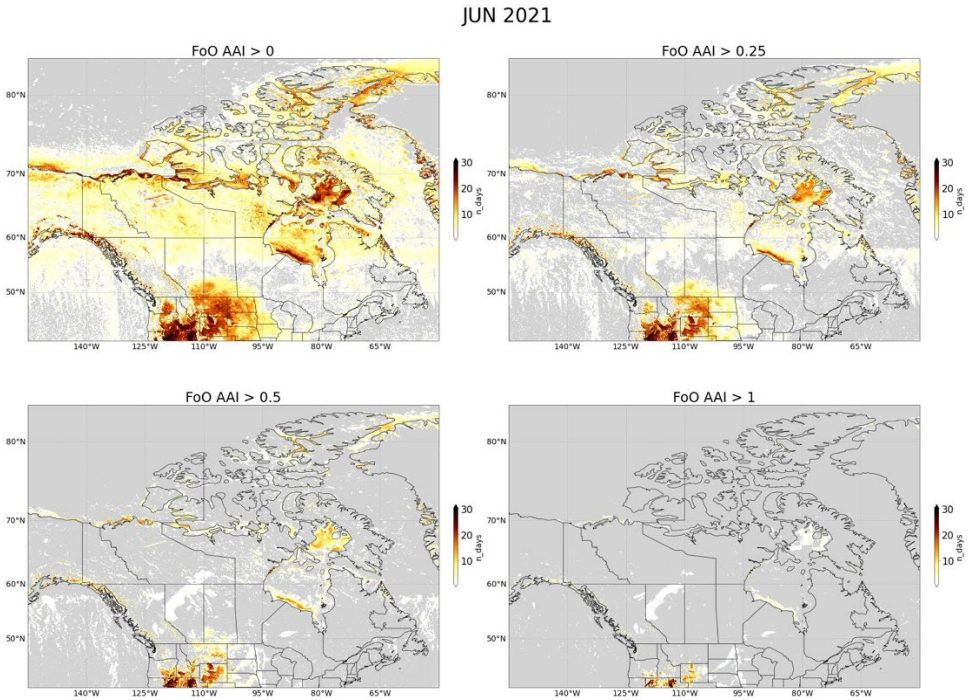

Figure 30: FoO of AAI at increasing thresholds over June 2021. Grey coloring indicates gridboxes either with NaN data or gridboxes with AAI below the given threshold (if n_days = 0, it was set to NaN). The minimum n_days is 1 and the maximum n_days is 30.

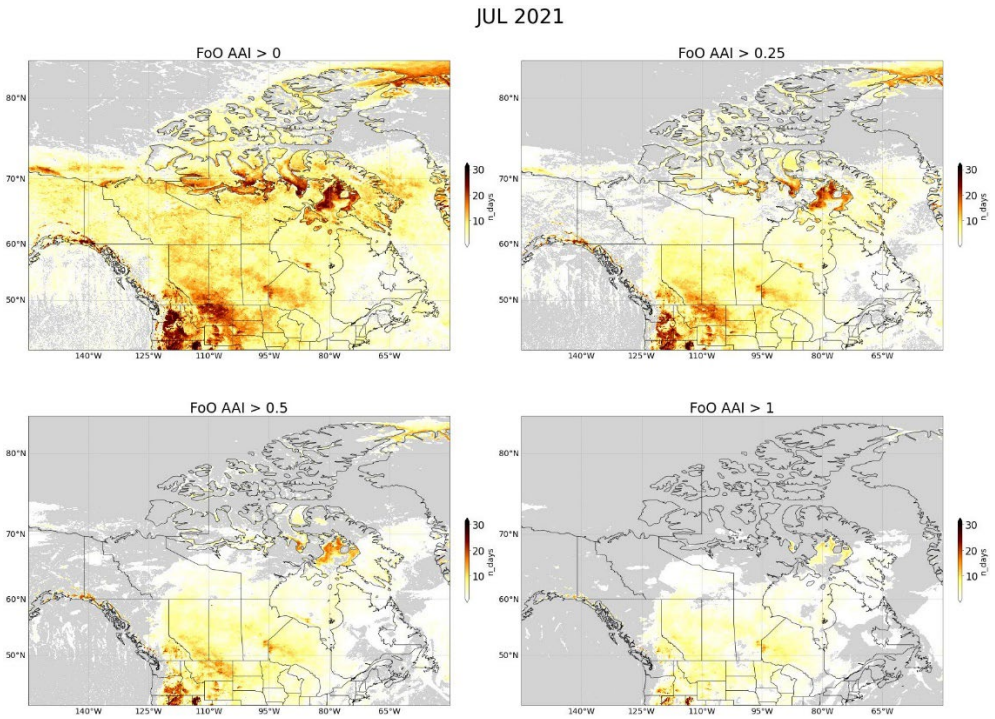

Figure 31: FoO of AAI at increasing thresholds over July 2021. Grey coloring indicates gridboxes either with NaN data or gridboxes with AAI below the given threshold (if n_days = 0, it was set to NaN). The minimum n_days is 1 and the maximum n_days is 31.

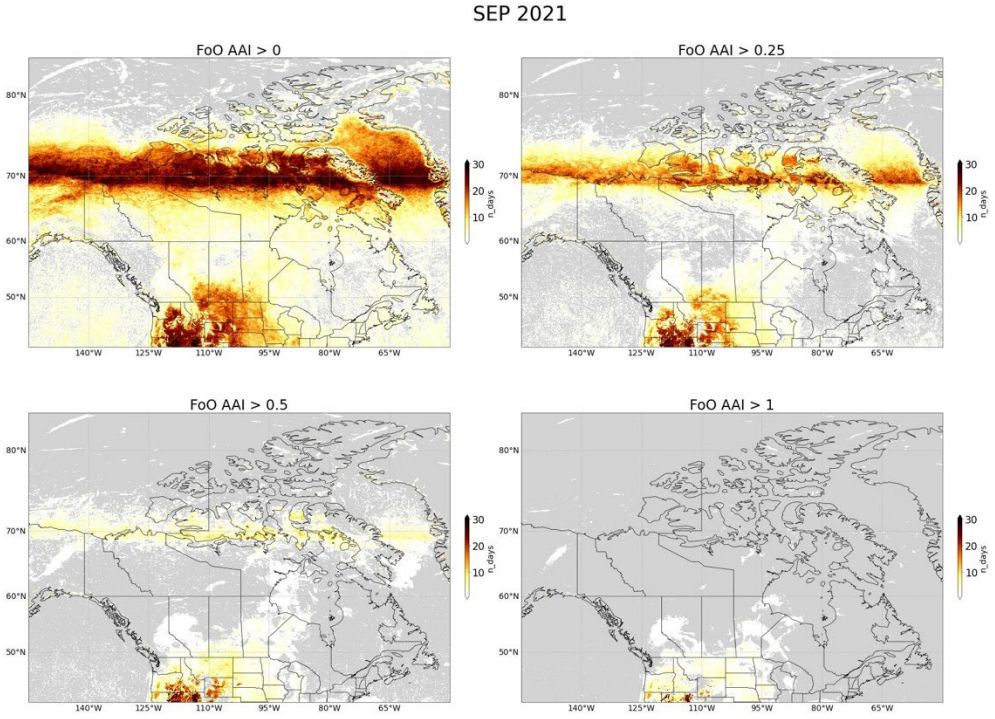

Figure 32: FoO of AAI at increasing thresholds over August 2021. Grey coloring indicates gridboxes either with NaN data or gridboxes with AAI below the given threshold (if n_days = 0, it was set to NaN). The minimum n_days is 1 and the maximum n_days is 31.

3.1.1

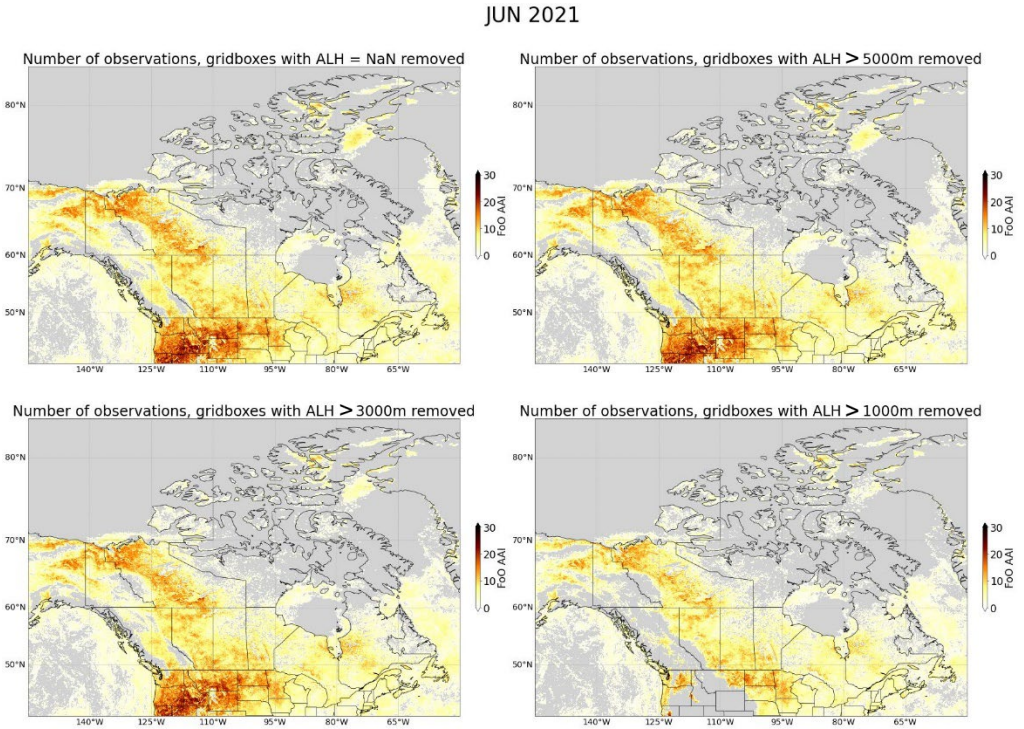

Figure 33: The changes in n_obs for July 2021 as the AAI data is filtered by ALH data at different thresholds. Gridboxes that did not have an ALH value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum FoO AAI is 1 and the maximum is 30.

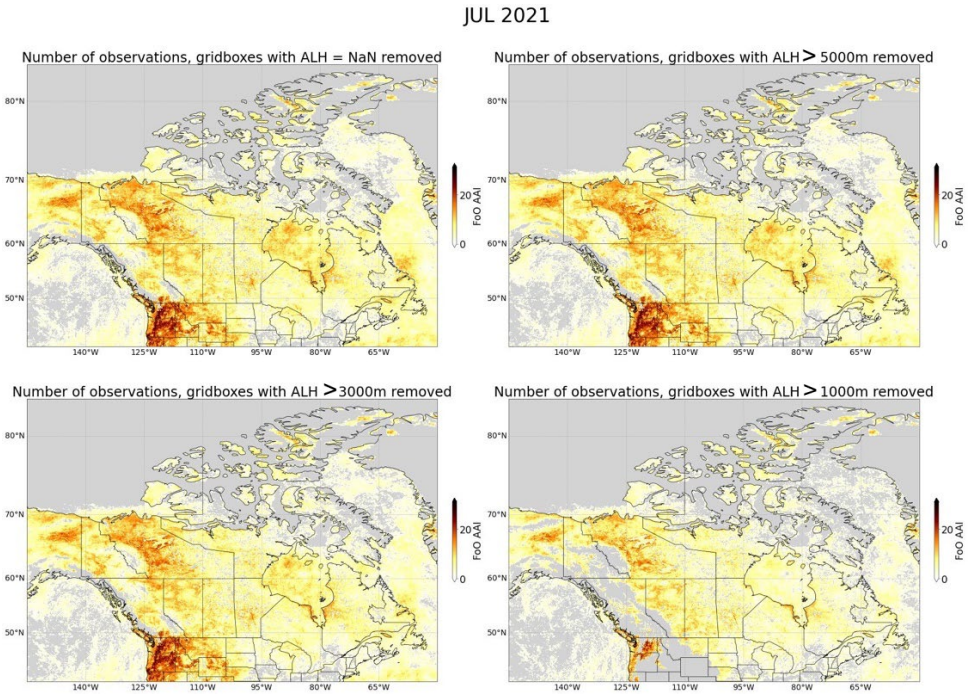

Figure 34: The changes in n_obs for July 2021 as the AAI data is filtered by ALH data at different thresholds. Gridboxes that did not have an ALH value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum FoO AAI is 1 and the maximum is 31.

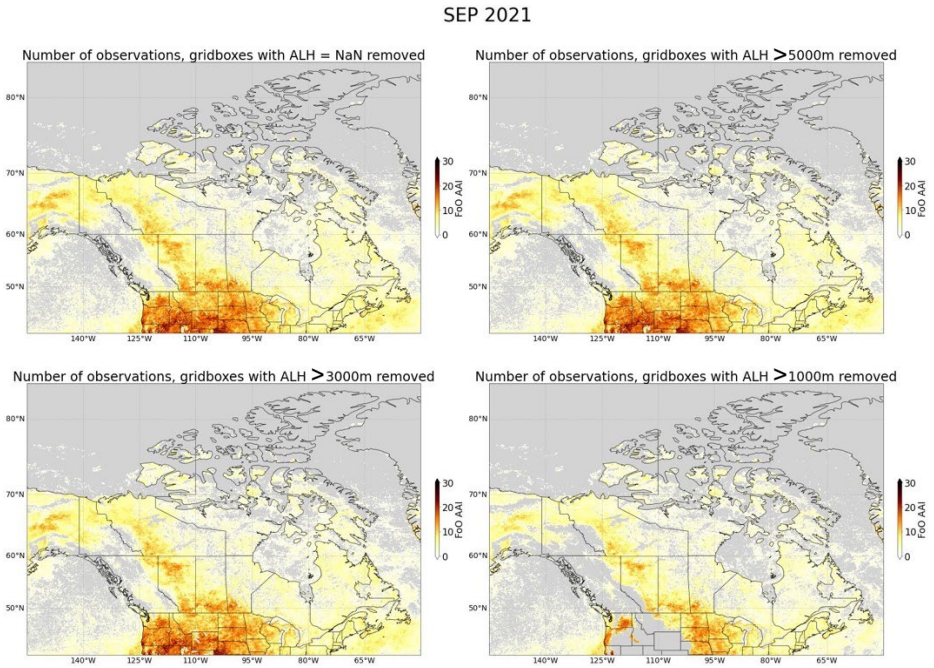

Figure 35: The changes in n_obs for September 2021 as the AAI data is filtered by ALH data at different thresholds. Gridboxes that did not have an ALH value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum FoO AAI is 1 and the maximum is 30.

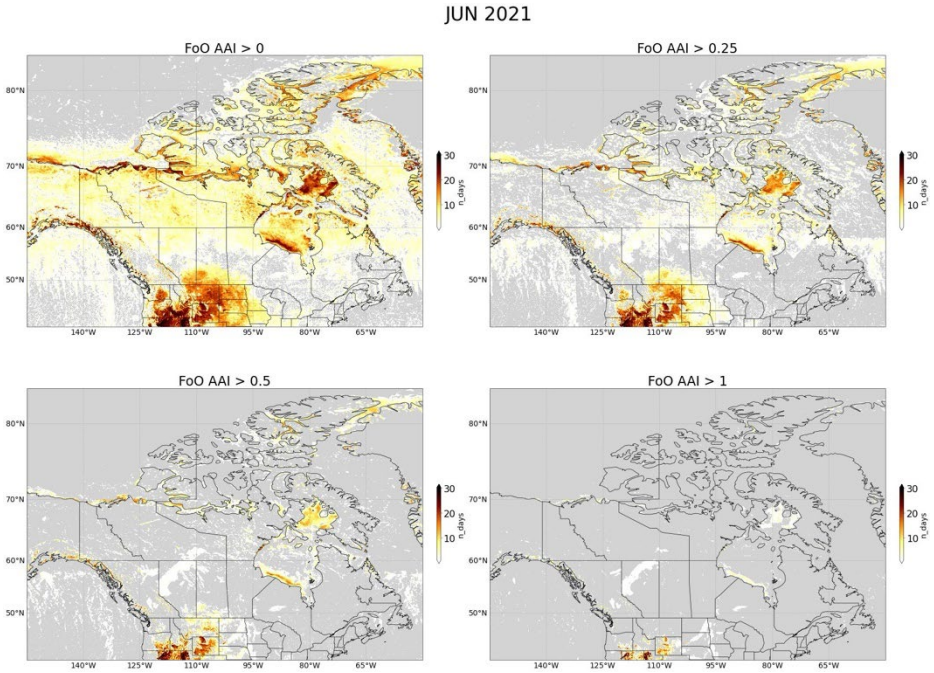

Figure 36: FoO of AAI greater than 0.5 for June 2021. Each panel is a different AAI dataset that has been filtered by ALH. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.

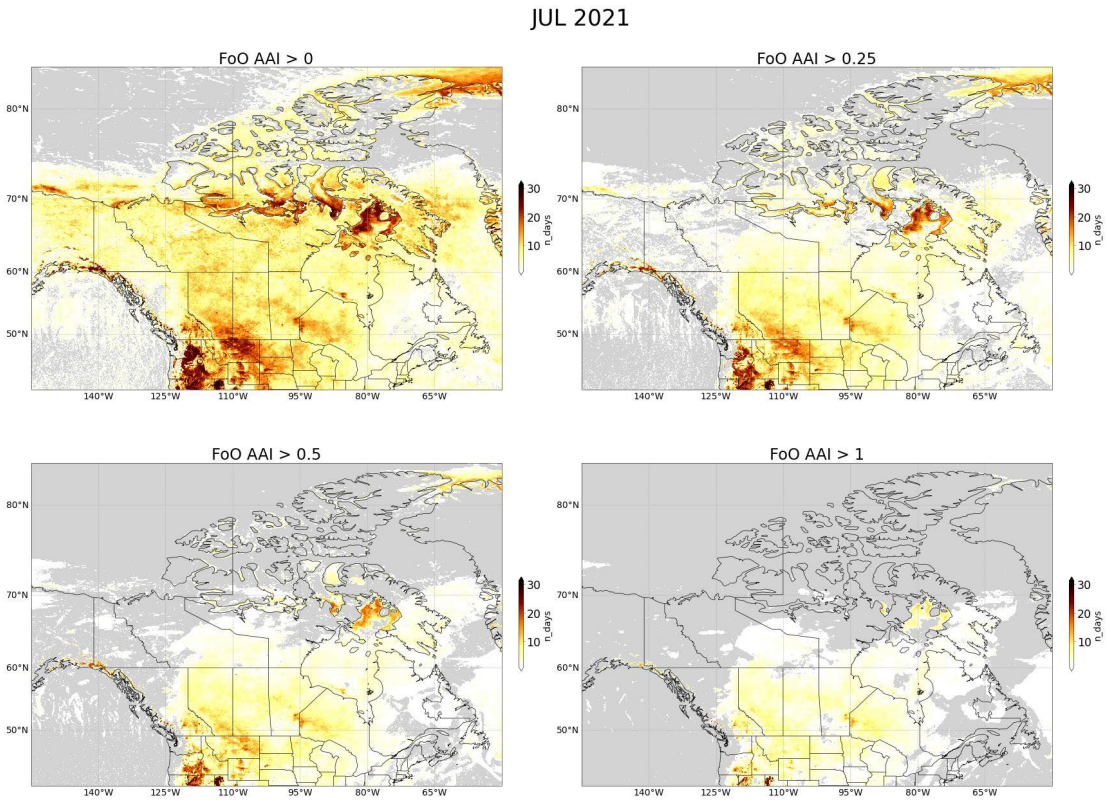

Figure 37: FoO of AAI greater than 0.5 for July 2021. Each panel is a different AAI dataset that has been filtered by ALH. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.

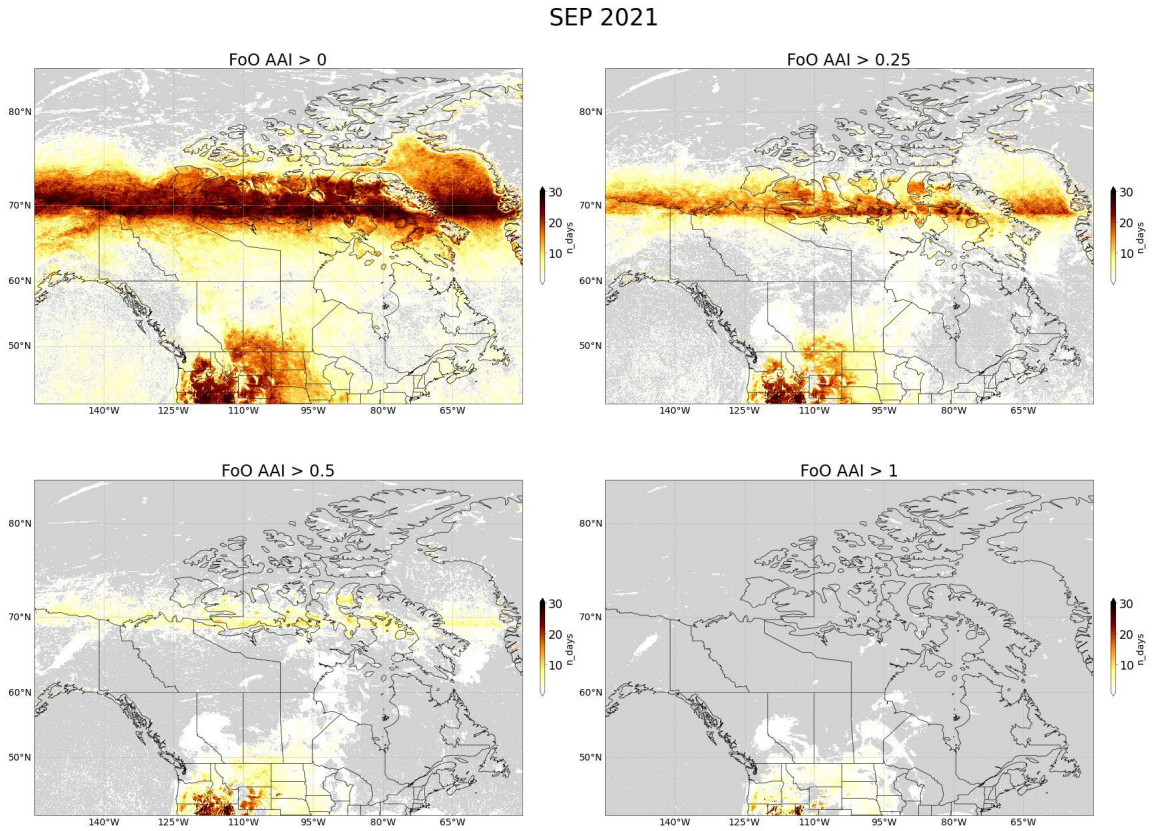

Figure 38: FoO of AAI greater than 0.5 for September 2021. Each panel is a different AAI dataset that has been filtered by ALH. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.

3.2

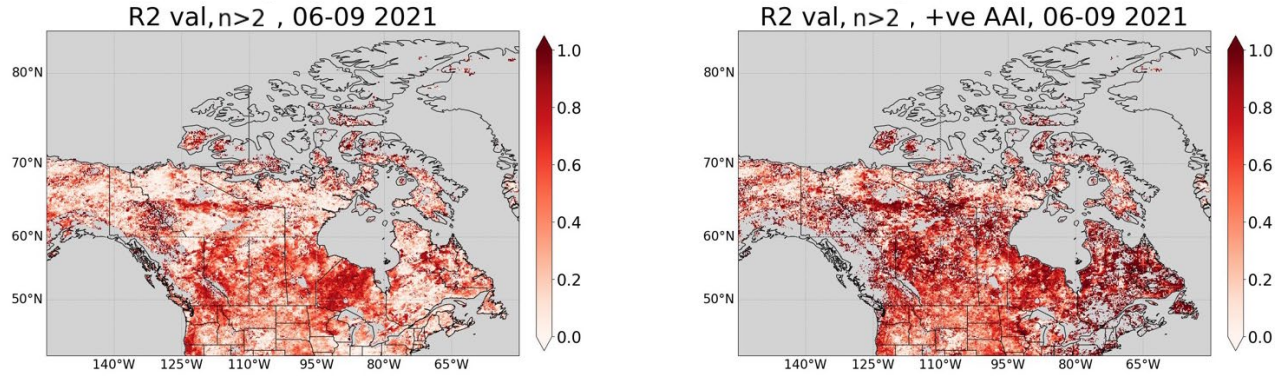

Figure 39: R2 for all AAI (left) compared to positive AAI (right). White coloring indicates an R^2 at or close to 0 and red coloring indicates R^2 at or close to 1.

3.3.1

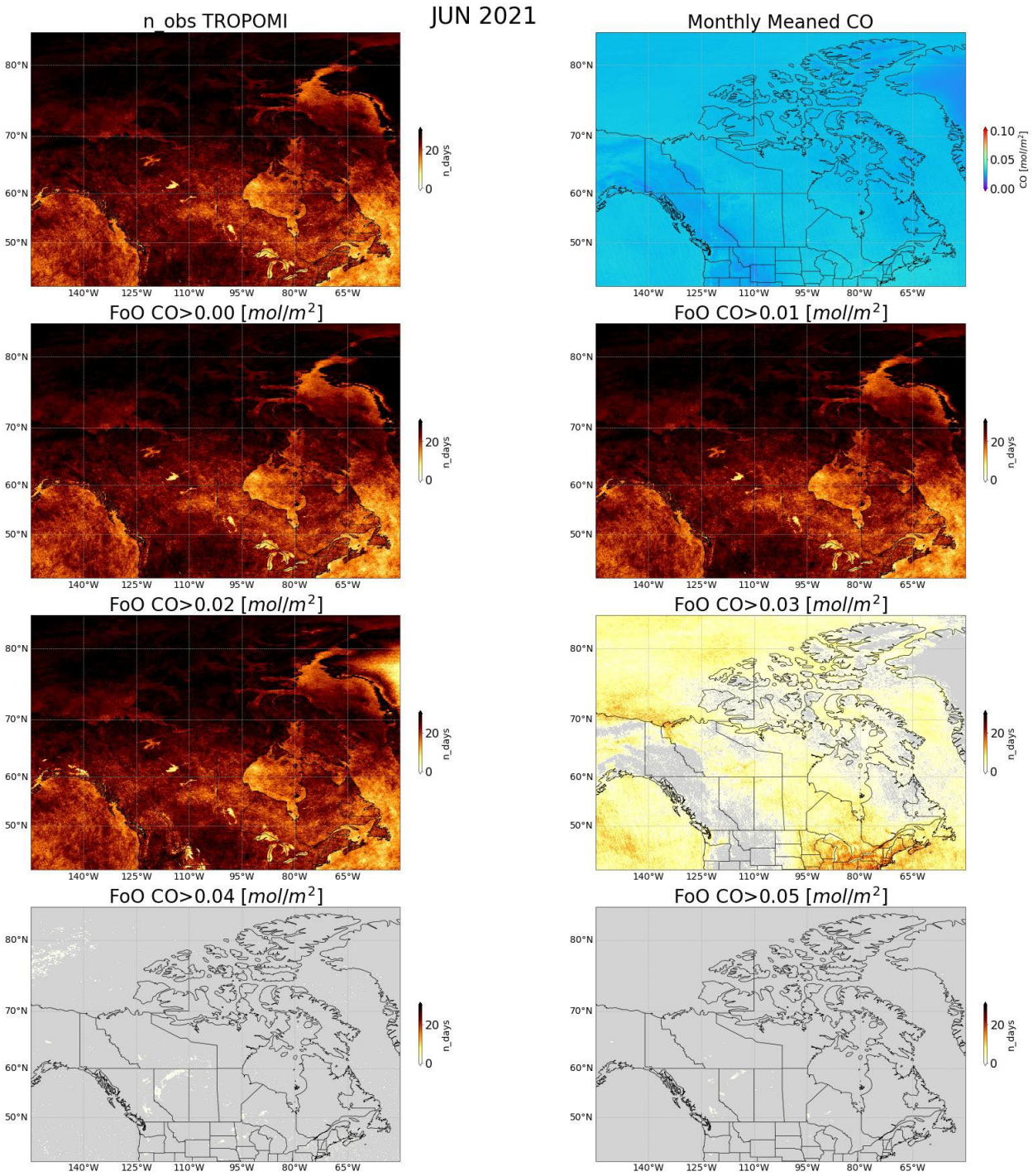

Figure 40: FoO maps showing the distribution of CO column density (measured in x 10^{18} [mol/m^2]) for June 2021. (Top left) – n_obs for the TROPOMI CO data. (Rows 2:4) - FoO of CO at different thresholds. (Top right) - Monthly meaned CO. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.

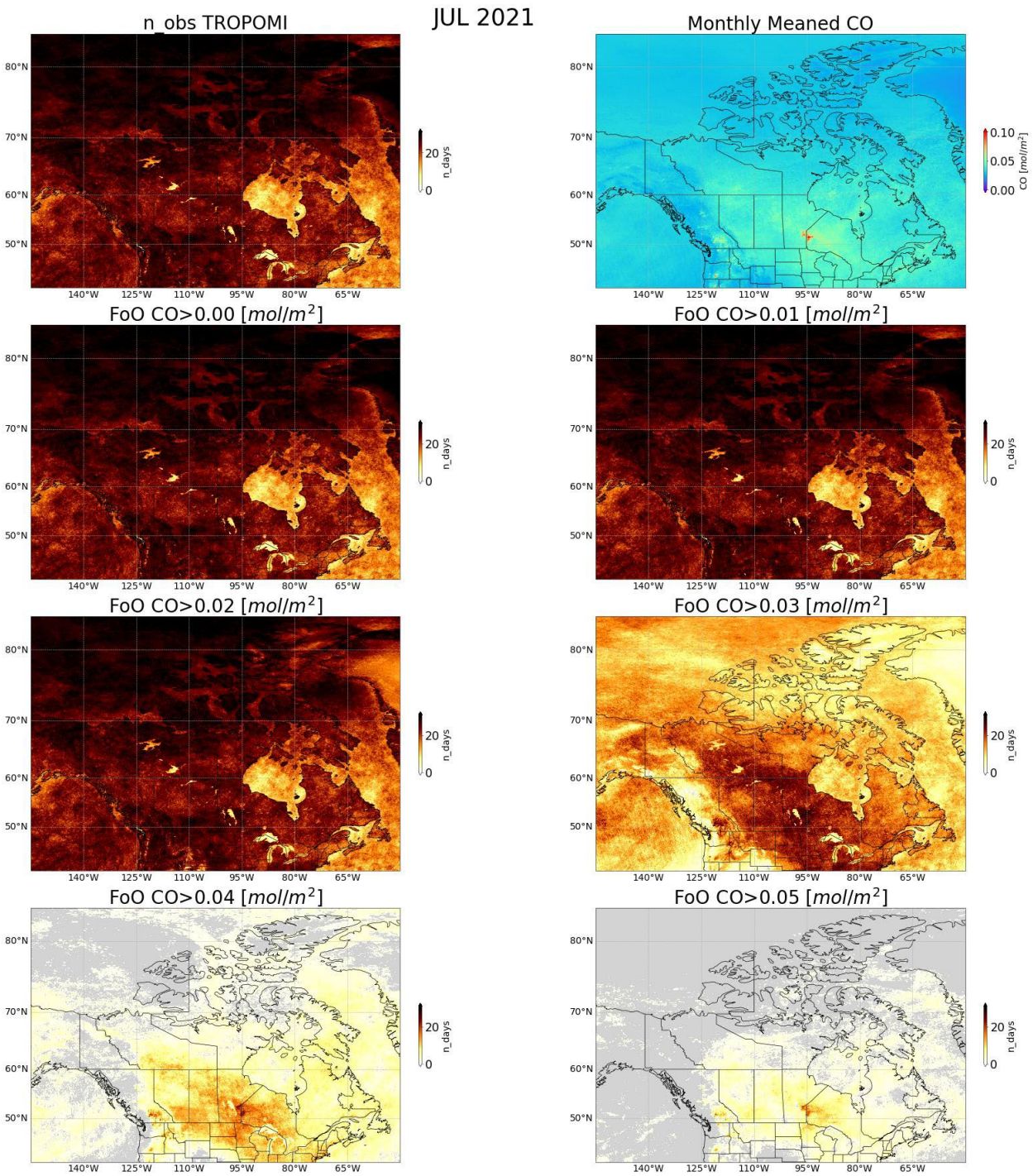

Figure 41: FoO maps showing the distribution of CO column density (measured in x 10^{18} [mol/m^2]) for July 2021. (Top left) – n_obs for the TROPOMI CO data. (Rows 2:4) - FoO of CO at different thresholds. (Top right) - Monthly meaned CO. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.

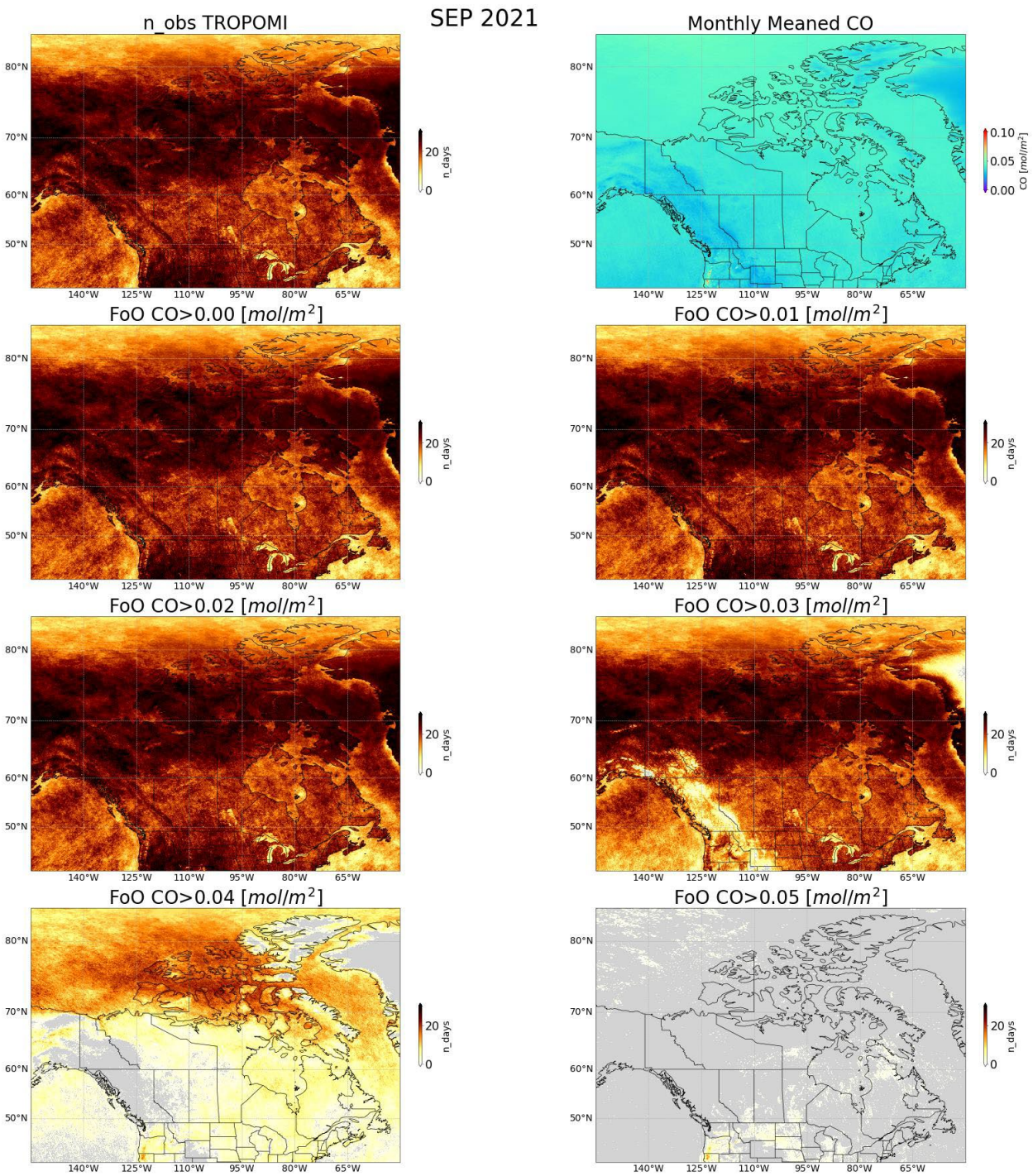

Figure 42: FoO maps showing the distribution of CO column density (measured in x 10^{18} [mol/m^2]) for September 2021. (Top left) – n_obs for the TROPOMI CO data. (Rows 2:4) - FoO of CO at different thresholds. (Top right) - Monthly meaned CO. NaN data is represented by grey coloring. Here, NaN data is either bad quality or gridboxes where n_days was 0.

3.3.2

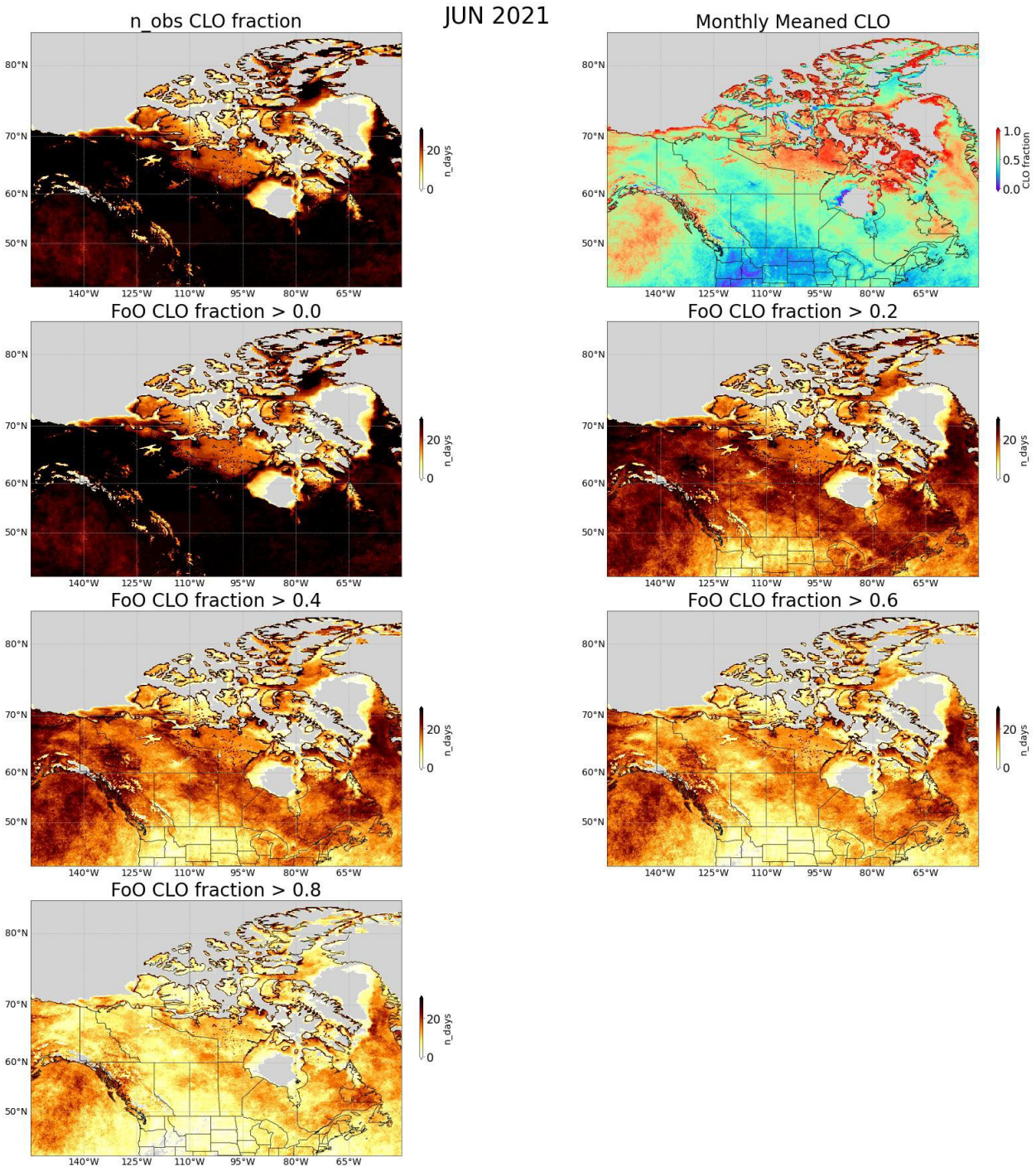

Figure 43: The changes in n_obs for June 2021 as the AAI data is filtered by CLO data at different thresholds. (Top left) – n_obs for the unfiltered dataset. (Rest of panels) – n_obs for the AAI datasets that have been filtered by different values of CLO. Gridboxes that did not have a CLO value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum n_days is 1 and the maximum is 30.

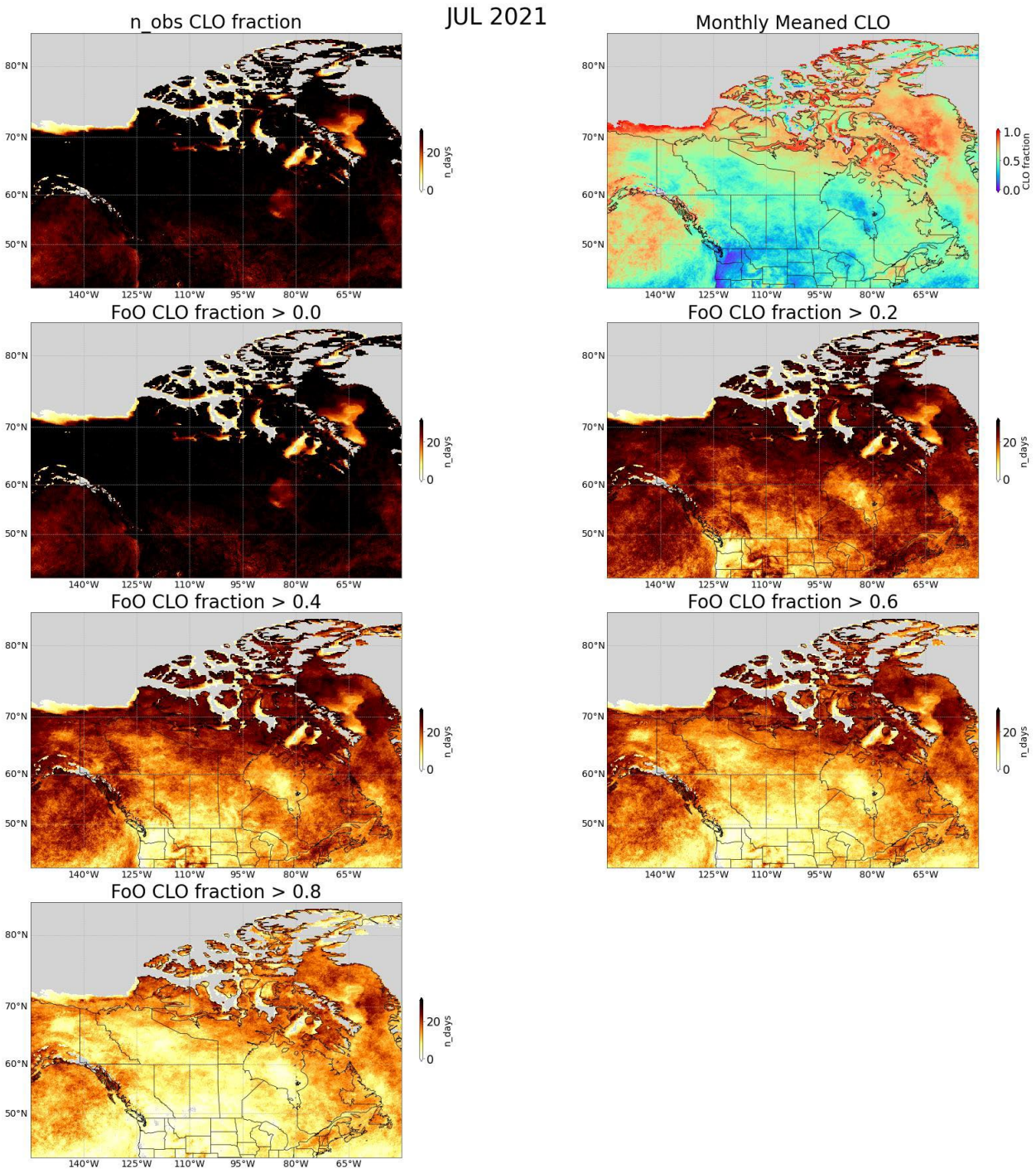

Figure 44: The changes in n_obs for July 2021 as the AAI data is filtered by CLO data at different thresholds. (Top left) – n_obs for the unfiltered dataset. (Rest of panels) – n_obs for the AAI datasets that have been filtered by different values of CLO. Gridboxes that did not have a CLO value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum n_days is 1 and the maximum is 31.

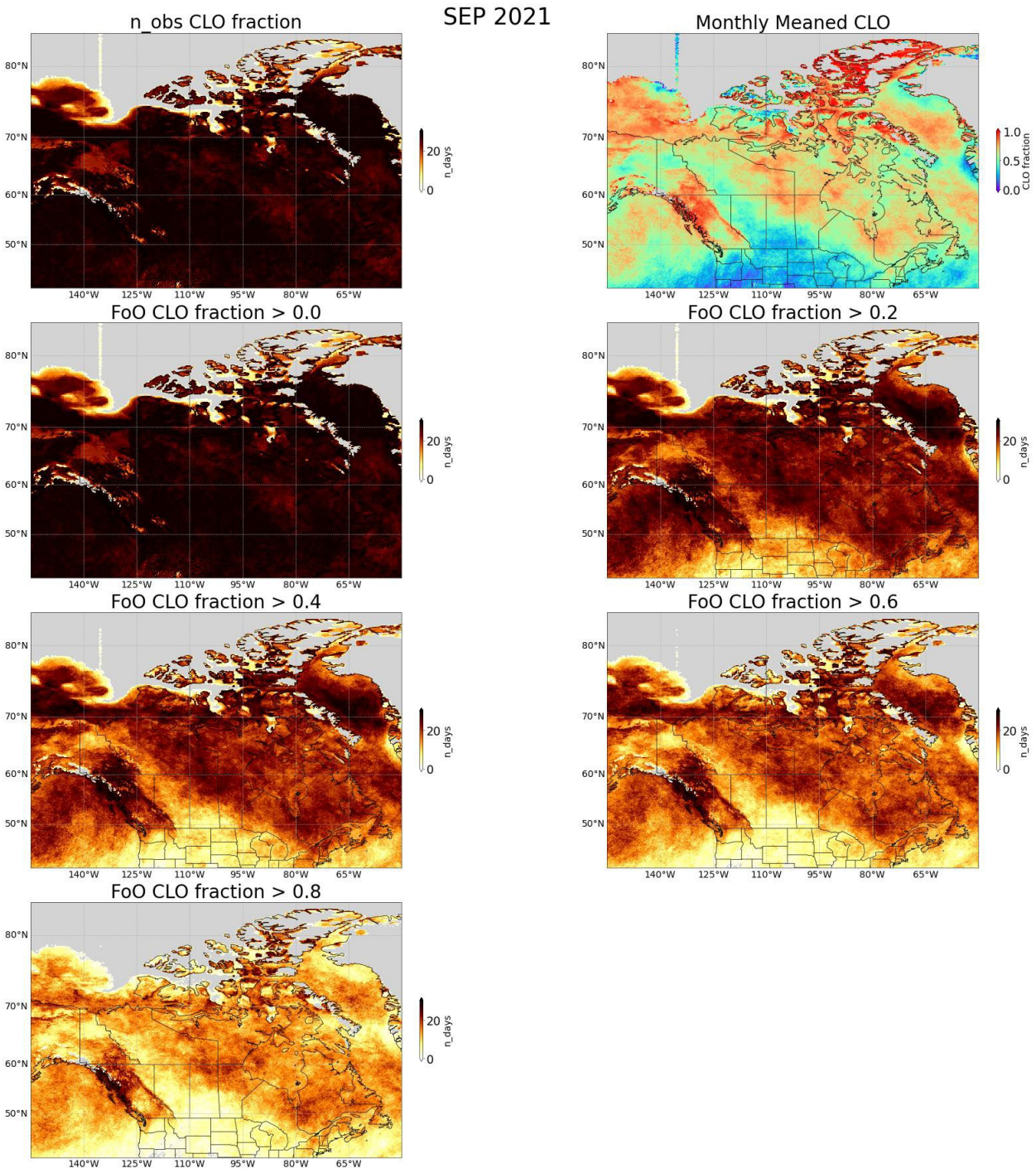

Figure 45: The changes in n_obs for September 2021 as the AAI data is filtered by CLO data at different thresholds. (Top left) – n_obs for the unfiltered dataset. (Rest of panels) – n_obs for the AAI datasets that have been filtered by different values of CLO. Gridboxes that did not have a CLO value below the desired threshold were set to NaN, indicated by the grey coloring. The minimum n_days is 1 and the maximum is 30..

References

Althaf, P., Kannemadugu, H. B., & Kumar, K. R. (2022). Hotspot analysis and long-term trends of absorbing aerosol index from dust emissions measured by the ozone monitoring instrument at different urban locations in India during 2005–2018. *Atmospheric Environment*, *272*, 118933. https://doi.org/10.1016/j.atmosenv.2021.118933

Amrhein, V., & Greenland, S. (2022). Discuss practical importance of results based on interval estimates and p-value functions, not only on point estimates and null p-values. *Journal of Information Technology,* 37(3), 316–320. https://doi.org/10.1177/02683962221105904

Anderson, T. L., Wu, Y., Chu, D. A., Schmid, B., Redemann, J., & Dubovik, O. (2005). Testing the MODIS satellite retrieval of aerosol fine-mode fraction. *Journal of Geophysical Research*., 110(D18). https://doi.org/10.1029/2005JD005978

Apituley, A, Pedergnana, M., Sneep, M., Veefkind, J. P., Loyola, D., Landgraf, J., & Borsdorff, T. (2022a). Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Carbon Monoxide. Royal Netherlands Meteorological Institute.

Apituley, A., Pedergnana, M., Sneep, M., Veefkind, J. P., Loyola, D., Sanders, B., & de Graaf, M. (2022b). Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Aerosol Layer Height. Royal Netherlands Meteorological Institute.

Apituley, A., Pedergnana, M., Sneep, M., Veefkind, J., Loyola, D., & Zweers, D. (2022c). Sentinel-5 precursor/TROPOMI Level 2 Product User Manual UV Aerosol Index. Royal Netherlands Meteorological Institute.

Babić, L., Braak, R., Dierssen, W., Kissi-Ameyaw, J., Kleipool, Q., Leloux, J., Loots, E., Ludewig, A., Rozemeijer, N., Smeets, J., & Vacanti, G. (2022). Algorithm theoretical basis document for the TROPOMI L01b data processor. Royal Netherlands Meteorological Institute.

Bellamy, D., Nadeau, D. F., & King, J. (2024). Forcing mechanisms of strong surface 1 winds in a dust storm-prone, high 2 latitude proglacial valley. Submitted to: *Journal of Applied Meteorology and Climatology*.

Bergstrom, R. W., Pilewskie, P., Russell, P. B., Redemann, J., Bond, T. C., Quinn, P. K., & Sierau, B. (n.d.). Spectral absorption properties of atmospheric aerosols. *Atmospheric Chemistry and Physics*, 7(23), 5937–5943. https://doi.org/10.5194/acp-7-5937-2007

Berman, J. J. (2016). *Data simplification taming information with open source tools*. Elsevier, MK Morgan Kaufmann.

Broomandi, P., Dabir, B., Bonakdarpour, B., Rashidi, Y., & Akherati, A. (2017). Simulation of mineral dust aerosols in southwestern Iran through numerical prediction models.

Environmental Progress & Sustainable Energy, *37*(4), 1380–1393. https://doi.org/10.1002/ep.12805

Boucher, O. (2015). *Atmospheric Aerosols Properties and Climate Impacts* (1st ed. 2015.). Springer Netherlands. https://doi.org/10.1007/978-94-017-9649-1

Bullard, J. E., Baddock, M., Bradwell, T., Crusius, J., Darlington, E., Gaiero, D., Gassó, S., Gisladottir, G., Hodgkins, R., McCulloch, R., McKenna-Neuman, C., Mockford, T., Stewart, H., & Thorsteinsson, T. (2016). High-latitude dust in the Earth System. *Reviews of Geophysics*, *54*(2), 447–485. https://doi.org/10.1002/2016rg000518

Buseck, P. R., Posfai, M., & Smith, J. V. (1999). Airborne minerals and related aerosol particles; effects on climate and the environment. *Proceedings of the National Academy of Sciences - PNAS*, 96(7), 3372–3379. https://doi.org/10.1073/pnas.96.7.3372

Chen, S., Tong, B., Dong, C., Wang, F., Chen, B., Cheng, C., Zhang, X., & Liu, D. (2020). Retrievals of aerosol layer height during dust events over the taklimakan and gobi desert. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 254, 107198-. https://doi.org/10.1016/j.jqsrt.2020.107198

Choi, Y., Hwang, J., Ok, J., Park, D. R., Su, H., Jiang, J. H., Huang, L., & Limpasuvan, T. (2020). Effect of arctic clouds on the ice-albedo feedback in Midsummer. *International Journal of Climatology*, *40*(10), 4707–4714. https://doi.org/10.1002/joc.6469

Cooper, M. J., Martin, R. V., Hammer, M. S., & McLinden, C. A. (2019). An observation-based correction for aerosol effects on nitrogen dioxide column retrievals using the absorbing aerosol index. *Geophysical Research Letters*, *46*(14), 8442–8452. https://doi.org/10.1029/2019gl083673

de Graaf, M. (2023). S5P Mission Performance Centre Aerosol Layer Height [L2__AER_LH] Readme. Royal Netherlands Meteorological Institute.

de Graaf, M., de Haan, J. F., & Sanders, A. F. J. (2022). TROPOMI ATBD of the Aerosol Layer Height. Royal Netherlands Meteorological Institute.

de Graaf, M., & Stammes, P. (2005). SCIAMACHY Absorbing Aerosol Index – calibration issues and global results from 2002–2004. *Atmospheric Chemistry and Physics*, 5(9), 2385–2394. https://doi.org/10.5194/acp-5-2385-2005

Faizi, N., & Alvi, Y. (2023). Biostatistics manual for health research : a practical guide to data analysis / Nafis Faizi, Yasir Alvi. *Elsevier Science & Technology,* 109-126.

Garimella, S., Huang, Y.-W., Seewald, J. S., & Cziczo, D. J. (2014). Cloud condensation nucleus activity comparison of dry- and wet-generated mineral dust aerosol: the significance of soluble material. *Atmospheric Chemistry and Physics*, 14(12), 6003–6019. https://doi.org/10.5194/acp-14-6003-2014

Ginoux, P., Prospero, J. M., Gill, T. E., Hsu, N. C., & Zhao, M. (2012). Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products. *Reviews of Geophysics (1985),* 50(3). https://doi.org/10.1029/2012RG000388

Google. (2024). [Google Maps image of the Arctic Archipelago showing the composite image of vegetation cover from multiple satellite sources]. Retrieved April 4th, 2024, from https://www.google.com/maps/

Griffin, D., Sioris, C., Chen, J., Dickson, N., Kovachik, A., de Graaf, M., Nanda, S., Veefkind, P., Dammers, E., McLinden, C. A., Makar, P., & Akingunola, A. (2020). The 2018 fire season in North America as seen by tropomi: Aerosol layer height intercomparisons and evaluation of model-derived Plume Heights. *Atmospheric Measurement Techniques*, *13*(3), 1427–1445. https://doi.org/10.5194/amt-13-1427-2020

Groot Zwaaftink, C. D., Grythe, H., Skov, H., & Stohl, A. (2016). Substantial contribution of northern high-latitude sources to mineral dust in the Arctic. *Journal of Geophysical Research. Atmospheres*, 121(22), 13,678-13,697. https://doi.org/10.1002/2016JD025482

Gupta, G., Ratnam, M. V., Madhavan, B. L., Prasad, P., & Narayanamurthy, C. S. (2021). Vertical and spatial distribution of elevated aerosol layers obtained using long-term ground-based and space-borne lidar observations. *Atmospheric Environment (1994)*, 246, 118172-. https://doi.org/10.1016/j.atmosenv.2020.118172

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science (American Association for the Advancement of Science)*, 342(6160), 850–853. https://doi.org/10.1126/science.1244693

Hsu, N. C., Jeong, M.-J., Bettenhausen, C., Sayer, A. M., Hansell, R., Seftor, C. S., Huang, J., & Tsay, S.-C. (2013). Enhanced Deep Blue aerosol retrieval algorithm: The second generation. *Journal of Geophysical Research. Atmospheres*, 118(16), 9296–9315. https://doi.org/10.1002/jgrd.50712

Huck, R., Bryant, R. G., & King, J. (2023). The (mis)identification of high-latitude dust events using remote sensing methods in the Yukon, Canada: a sub-daily variability analysis. *Atmospheric Chemistry and Physics*, 23(11), 6299–6318. https://doi.org/10.5194/acp-23-6299-2023

Ichoku, C., Kaufman, Y. J., Remer, L. A., & Levy, R. (2004). Global Aerosol Remote Sensing from Modis. *Advances in Space Research*, *34*(4), 820–827. https://doi.org/10.1016/j.asr.2003.07.071

Jain, P., Sharma, A. R., Acuna, D. C., Abatzoglou, J. T., Flannigan, M. (2024). Record-breaking fire weather in North America in 2021 was initiated by the Pacific northwest heat dome. *Commun Earth Environ* 5(202). https://doi.org/10.1038/s43247-024-01346-2

Khorram, S. (2012). *Remote Sensing*. Springer.

Knippertz, P., & Stuut, J.-B. W. (2014). Mineral dust a key player in the earth system. *Springer Netherlands*.

Kooreman, M. L., Stammes, P., Trees, V., Sneep, M., Tilstra, L. G., de Graaf, M., Stein Zweers, D. C., Wang, P., Tuinder, O. N., & Veefkind, J. P. (2020). Effects of clouds on the UV absorbing aerosol index from Tropomi. *Atmospheric Measurement Techniques*, *13*(12), 6407–6426. https://doi.org/10.5194/amt-13-6407-2020

Landgraf, J, Borsdorff, T., Langerock, B., & Keppens, A. (2023). S5P Mission Performance Centre Carbon Monoxide [L2__CO____] Readme. Royal Netherlands Meteorological Institute.

Landgraf, J., aan de Brugh, J., Scheepmaker, R., Borsdorff, T., Houweling, S., & Hasekamp, O. (2022). Algorithm Theoretical Baseline Document for Sentinel-5 Precursor: Carbon Monoxide Total Column Retrieval. Royal Netherlands Meteorological Institute.

Latsch, M., Richter, A., Eskes, H., Sneep, M., Wang, P., Veefkind, P., Lutz, R., Loyola, D., Argyrouli, A., Valks, P., Wagner, T., Sihler, H., van Roozendael, M., Theys, N., Yu, H., Siddans, R., & Burrows, J. P. (2022). Intercomparison of sentinel-5p tropomi cloud products for tropospheric trace gas retrievals. *Atmospheric Measurement Techniques*, *15*(21), 6257–6283. https://doi.org/10.5194/amt-15-6257-2022

Li, F., Zhang, X., Kondragunta, S., & Lu, X. (2020). An evaluation of Advanced Baseline Imager fire radiative power based wildfire emissions using carbon monoxide observed by the tropospheric monitoring instrument across the conterminous United States. *Environmental Research Letters*, *15*(9), 094049. https://doi.org/10.1088/1748-9326/ab9d3a

Li, Y., Mickley, L. J., & Kaplan, J. O. (2021). Response of dust emissions in southwestern North America to 21st century trends in climate, CO2 fertilization, and land use: implications for air quality. *Atmospheric Chemistry and Physics*, 21(1), 57–68. https://doi.org/10.5194/acp-21-57-2021

Loyola, D., Lutz, R., Argyrouli, A., & Spurr, R. (2021). S5P/TROPOMI ATBD Cloud Products. Royal Netherlands Meteorological Institute.

Lubin, D., & Massom, Robert. (2006). *Polar remote sensing* (1st ed.). Springer.

Lutz, R., Romahn, F., Compernolle, S., & Lambert, J. C. (2023). S5P Mission Performance Centre CLOUD [L2__CLOUD_] Readme. Royal Netherlands Meteorological Institute.

Ma, J., Shen, H., Cai, Y., Zhang, T., Su, J., Chen, W.-H., & Li, J. (2023). UCTNet with dual-flow architecture: Snow coverage mapping with sentinel-2 satellite imagery. *Remote Sensing*, *15*(17), 4213. https://doi.org/10.3390/rs15174213

Mazzola, M., Stone, R. S., Herber, A., Tomasi, C., Lupi, A., Vitale, V., Lanconelli, C., Toledano, C., Cachorro, V. E., O'Neill, N. T., Shiobara, M., Aaltonen, V., Stebel, K., Zielinski, T., Petelski, T., Ortiz de Galisteo, J. P., Torres, B., Berjon, A., Goloub, P., … Virkkula, A. (2012). Evaluation of sun photometer capabilities for retrievals of aerosol optical depth at high latitudes: The polar-AOD intercomparison campaigns. *Atmospheric Environment*, *52*, 4–17. https://doi.org/10.1016/j.atmosenv.2011.07.042

Meinander, O., Dagsson-Waldhauserova, P., Amosov, P., Aseyeva, E., Atkins, C., Baklanov, A., Baldo, C., Barr, S. L., Barzycka, B., Benning, L. G., Cvetkovic, B., Enchilik, P., Frolov, D., Gassó, S., Kandler, K., Kasimov, N., Kavan, J., King, J., Koroleva, T., … Vukovic Vimic, A. (2022). Newly identified climatically and environmentally significant high-latitude dust sources. *Atmospheric Chemistry and Physics*, *22*(17), 11889–11930. https://doi.org/10.5194/acp-22-11889-2022

Mushtaq, F., Farooq, M., Mukherjee, A. B., & Ghosh Nee Lala, M. (2023). Urban Air Quality Monitoring and Modelling Using Ground Monitoring, Remote Sensing, and GIS. In Geospatial Analytics for Environmental Pollution Modeling (pp. 213–247). *Springer*. https://doi.org/10.1007/978-3-031-45300-7_9

Park, R. J., Jacob, D. J., & Logan, J. A. (2007). Fire and biofuel contributions to annual mean aerosol mass concentrations in the United States. *Atmospheric Environment (1994),* 41(35), 7389–7400. https://doi.org/10.1016/j.atmosenv.2007.05.061

Parya, B., Dabir, B., Bonakdarpour, B., Rashidi, Y., & Akherati, A. (2018). Simulation of mineral dust aerosols in southwestern iran through numerical prediction models. *Environmental Progress & Sustainable Energy*, 37(4), 1380–1393. https://doi.org/10.1002/ep.12805

Penning de Vries, M.J.M., Beirle, S., & Wagner, T. (2009). UV aerosol indices from SCIAMACHY: introducing the SCattering Index (SCI). *Atmospheric Chemistry and Physics Discussions*, 9(3), 13569–13592.

Plant, G., Kort, E. A., Murray, L. T., Maasakkers, J. D., & Aben, I. (2022). Evaluating urban methane emissions from space using tropomi methane and carbon monoxide observations. *Remote Sensing of Environment*, *268*, 112756. https://doi.org/10.1016/j.rse.2021.112756

Pu, B., & Ginoux, P. (2016). The impact of the Pacific Decadal Oscillation on springtime dust activity in Syria. *Atmospheric Chemistry and Physics*, 16(21), 13431–13448. https://doi.org/10.5194/acp-16-13431-2016

Ranjbar, K., O'Neill, N. T., Ivanescu, L., King, J., & Hayes, P. L. (2021). Remote Sensing of a high-arctic, local dust event over Lake Hazen (Ellesmere Island, Nunavut, Canada). *Atmospheric Environment*, *246*, 118102. https://doi.org/10.1016/j.atmosenv.2020.118102

Romahn, F., Pedergnana, M., Loyola, D., Apituley, A., Sneep, M., & Veefkind, J. P. (2022). Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Cloud Properties. Royal Netherlands Meteorological Institute.

Shikwambana, L. (2022). Global Distribution of Clouds over Six Years: A Review Using Multiple Sensors and Reanalysis Data. *Atmosphere*, 13(9), 1514-. https://doi.org/10.3390/atmos13091514

Stammes, P., & Noordhoek, R. (2002). OMI Algorithm Theoretical Basis Document - Volume III - Clouds, Aerosols, and Surface UV Irradiance. Royal Netherlands Meteorological Institute.

Tang, F., Wang, W., Si, F., Zhou, H., Luo, Y., & Qian, Y. (2022). Successful derivation of absorbing aerosol index from the environmental trace gases monitoring instrument (EMI). *Remote Sensing*, *14*(16), 4105. https://doi.org/10.3390/rs14164105

Tang, F., Wang, W., Si, F., Zhou, H., Luo, Y., Yang, D., & Qian, Y. (2023). A statistical algorithm for retrieving background value of absorbing aerosol index based on Tropomi Measurements. *Advances in Atmospheric Sciences*, *40*(6), 1104–1116. https://doi.org/10.1007/s00376-022-2093-3

Thenkabail, P. S. (2014). Hyperspectral Remote Sensing of Vegetation and Agricultural Crops. *Photogrammetric Engineering and Remote Sensing*, *80*(8), 697–709.

Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., & Gleason, J. (1998). Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis. *Journal of Geophysical Research*., 103(D14), 17099–17110. https://doi.org/10.1029/98JD00900

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., & Levelt, P. (2007). Aerosols and surface UV products from Ozone Monitoring Instrument Observations: An overview. *Journal of Geophysical Research: Atmospheres*, *112*(D24). https://doi.org/10.1029/2007jd008809

Trees, V., Wang, P., & Stammes, P. (2021). *Restoring the Top-of-Atmosphere Reflectance during Solar Eclipses: A Proof of Concept with the UV Absorbing Aerosol Index Measured by Tropomi*. https://doi.org/10.5194/egusphere-egu21-3181

United Nations Environment Programme. (1970) *Global assessment of sand and dust storms*. UN Environment Document Repository Home. https://wedocs.unep.org/handle/20.500.11822/7681

Vaillant, N. M., Kolden C. A., Smith, A. M. S. (2016). Assessing Landscape Vulnerability to Wildfire USA. *Current Forestry Reports,* 2. https://doi.org/10.1007/s40725-016-0040-1

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., & Arellano, A. F. (2006). Interannual variability in global biomass burning emissions from 1997 to 2004. *Atmo. Chem. and Phys.*, 6(11), 3423–3441. https://doi.org/10.5194/acp-6-3423-2006

Wan, N., Xiong, X., Kluitenberg, G. J., Hutchinson, J. M. S., Aiken, R., Zhao, H., & Lin, X. (2023). Estimation of biomass burning emission of NO2 and CO from 2019–2020 Australia fires based on satellite observations. *Atmospheric Chemistry and Physics*, 23(1), 711–724. https://doi.org/10.5194/acp-23-711-2023

Wang, H., Wu, Q., Liu, H., Wang, Y., Cheng, H., Wang, R., Wang, L., Xiao, H., & Yang, X. (2018). Sensitivity of biogenic volatile organic compound emissions to leaf area index and land cover in Beijing. *Atmospheric Chemistry and Physics*, 18(13), 9583–9596. https://doi.org/10.5194/acp-18-9583-2018

Wiacek, A., T., Peter, and U. Lohmann (2010), The potential influence of Asian and African mineral dust on ice, mixed-phase and liquid water clouds, *Atm. Chem. Phys.,* 10, 8649–8667, https://doi:10.5194/acp-10-8649-2010.

Williamson, S. N., & Menounos, B. (2021). The influence of forest fire aerosol and air temperature on Glacier Albedo, western North America. *Remote Sensing of Environment*, *267*, 112732. https://doi.org/10.1016/j.rse.2021.112732

Wu, L., Hasekamp, O., van Diedenhoven, B., Cairns, B., Yorks, J. E., & Chowdhary, J. (2016). Passive remote sensing of aerosol layer height using near-uv multiangle polarization measurements. *Geophysical Research Letters*, *43*(16), 8783–8790. https://doi.org/10.1002/2016gl069848

Xie, Y., Xiong, X., Qu, J. J., Che, N., & Wang, L. (2006). Modis on-orbit spatial characterization results using ground measurements. *SPIE Proceedings*. https://doi.org/10.1117/12.680787

Yong-Sang, C., Hwang, J., Ok, J., Park, D. R., Su, H., Jiang, J. H., Huang, L., & Limpasuvan, T. (2020). Effect of Arctic clouds on the ice-albedo feedback in midsummer. *International Journal of Climatology*, 40(10), 4707–4714. https://doi.org/10.1002/joc.6469

Zhang, B., Wu, Y., Zhao, B., Chanussot, J., Hong, D., Yao, J., & Gao, L. (2022). Progress and challenges in Intelligent Remote Sensing Satellite Systems. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 1814–1822. https://doi.org/10.1109/jstars.2022.3148139

Zweers, S. (2022). TROPOMI ATBD of the UV aerosol index. Royal Netherlands Met. Institute.

Zweers, D., & Wagner, T. (2023). S5P Mission Performance Centre UV Aerosol Index [L2__AER_AI] Readme. Royal Netherlands Meteorological Institute.