In this well-written manuscript, the authors offer several interesting comparisons between models (MERRA2 and FLEXPART) and observations (CALIPSO). But because they confine themselves to “qualitative comparisons”, these comparisons are somewhat less informative than I had hoped when I first agreed to review the paper. I was particularly disappointed by the lack of a quantitative assessment of CALIPSO’s propensity to misclassify diamond dust as mineral dust. Nevertheless, I believe the authors’ results will be helpful in understanding the limits of both measurements and models in the Arctic environment. That being the case, I’m happy to recommend publication once the authors address the comments embedded in the annotated version of their manuscript appended below.
Comparisons between the distributions of dust and combustion aerosols in MERRA-2, FLEXPART, and CALIPSO and implications for deposition freezing over wintertime Siberia

Lauren M. Zamora¹,², Ralph A. Kahn², Nikolaos Evangeliou³, and Christine D. Groot Zwaaftink³

¹Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland, U.S.A.
²NASA Goddard Space Flight Center, Greenbelt, Maryland, U.S.A.
³Norwegian Institute for Air Research (NILU), Kjeller, Norway

Correspondence to: Lauren M. Zamora (lauren.m.zamora@nasa.gov)

Abstract. Aerosol distributions have a potentially large influence on climate-relevant cloud properties but can be difficult to observe over the Arctic given pervasive cloudiness, long polar nights, data paucity over remote regions, and periodic diamond dust events that satellites can misclassify as aerosol. We compared Arctic 2008-2015 mineral dust and combustion aerosol distributions from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite, the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) reanalysis products, and the FLEXible PARticle (FLEXPART) model. Based on Atmospheric Infrared Sounder (AIRS) satellite meteorological data, diamond dust may occur up to 60% of the time in winter, but it hardly ever occurs in summer. In its absence, MERRA-2 and FLEXPART each predict the vertical distribution of combustion aerosols with relatively high confidence, as does FLEXPART for mineral dust. Comparisons to ground and satellite data suggest that MERRA-2 Arctic dust concentrations can be improved by the addition of local dust sources. Apparent false negative rates compared to lidar were substantially higher in conditions favouring diamond dust formation for both MERRA-2 and FLEXPART, as would be expected if CALIPSO were misclassifying diamond dust as mineral dust aerosols. All three products predicted that wintertime dust and combustion aerosols occur most frequently over the same Siberian regions where diamond dust is most common in the winter. This suggests that aerosol impacts on ice phase processes may be particularly high over Siberia, although further wintertime model validation with non-CALIPSO observations is needed. This assessment paves the way for applying the model-based aerosol simulations to a range of regional-scale Arctic aerosol-cloud interaction studies.

1 Introduction

Aerosols have a potentially large influence on Arctic climate-relevant cloud properties such as cloud cover, cloud phase, and cloud particle size (Alterskjær et al., 2010; Gagné et al., 2017; Morrison et al., 2012; Schmale et al., 2021; Shindell and...
Faluvegi, 2009; Willis et al., 2018; Zamora et al., 2018). Mineral dust, for example, is thought to be a particularly efficient ice nucleating particle (INP) source, leading to enhanced freezing of liquid aerosol particles or potentially to depositional growth under favorable environmental conditions (Kanji et al., 2017). Marine aerosols are a source of both cloud condensation nuclei (CCN) and INPs (Willis et al., 2018). Combustion (anthropogenic pollution + biomass burning) aerosols are less efficient INPs (Kanji et al., 2017), but can readily form CCN, as can sulphate aerosols.

However, understanding the effects of aerosols on Arctic clouds is limited in large part by uncertainties in the distributions of aerosols and the contributions of each aerosol type to the CCN and INP budgets. Ground-based aerosol observations are sparse, and aerosol data from passive satellite instruments are unavailable during polar night (Duncan et al., 2020). Long-term, vertically-resolved remote sensing information on Arctic aerosols, including aerosol subtype distributions from dust, smoke, and other sources are available from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite (Di Pierro et al., 2013; Winker et al., 2013; Kim et al., 2018; Yang et al., 2021; Chen et al., 2021). However, CALIPSO can miss aerosols that are a) dilute (Kacenelenbogen et al., 2014; Zamora et al., 2017; Di Pierro et al., 2013; Winker et al., 2013; Rogers et al., 2014), b) very small (Hallen and Philbrick, 2018), including highly numerous marine biogenic aerosols (Burkart et al., 2017), c) in the 200 m immediately above the surface where local marine and terrestrial emission concentrations are highest, and d) below clouds. This latter issue is particularly challenging, as clouds occur in the Arctic at least half the time during winter, and up to 80% of the time over open ocean in summer (Zygmuntowska et al., 2012). Moreover, CALIPSO does not measure lidar ratios, and so periodic Arctic diamond dust events (i.e., ice crystal precipitation in clear sky conditions) can sometimes be misclassified as mineral dust aerosol, leading to overestimates of dust aerosol presence (Di Biagio et al., 2018).

Models are therefore critical for providing aerosol estimates below and within clouds at all vertical levels. However, most aerosol models are poorly validated at high altitudes and latitudes (Arnold et al., 2016; Eckhardt et al., 2015). These combined remote sensing and model issues lead to large uncertainties in predictions of Arctic aerosol types, levels, and their resulting cloud impacts now and in the future.

In this paper, we aim to take advantage of the complementary information from model and reanalysis products and satellite data to a) understand the strengths and limitations of the model and reanalysis products and b) better identify those Arctic regions with the highest certainty and uncertainty in the distributions of dust and combustion aerosols. This information will enable improvements in the model and reanalysis products and will facilitate a range of Arctic aerosol-cloud interaction studies going forward, including targeted suborbital measurements. As one example application, we use the overlap between meteorological conditions conducive to deposition nucleation and elevated aerosol presence in CALIPSO, the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2), and the FLEXPART (FLEXible PARTicle) Lagrangian particle dispersion model to show that this process may be more common during winter than previously thought.
2 Methods

This study focuses on the Arctic areas poleward of 60 °N between 2008 to 2015. In some cases, data were separated into terrestrial and oceanic regions, as defined by the ETOPO1 1 Arc-Minute Global Relief Model (Amante and Eakins, 2009)). We assess the commonalities and differences in dust and combustion aerosol distributions between CALIPSO, MERRA-2 reanalysis aerosol products, and aerosols in the FLEXPART model.

2.1 Data sources

2.1.1 CALIPSO data

Aerosol data, including aerosol layer base and top elevation, cloud-aerosol discrimination (CAD) score, and aerosol type, were obtained from the CALIPSO Lidar Level 2 5-km Merged Layer Data, V4-20 dataset (Winker, 2018), as were data on solar zenith angle (SZA). CALIPSO reports an aerosol vertical resolution of 30 m up to ~8.2 km asl, 60 m above that up to 20.2 km asl. Averaged aerosol horizontal resolution ranges between 5 and 80 km, with higher resolution at higher aerosol concentrations, when signal-to-noise ratios are better (Vaughan et al., 2009). We focused on data above 200 m from the surface, to reduce the influence of blowing snow and ground contamination of the CALIPSO lidar on the results.

For this analysis, CALIPSO aerosol data were used only from cloud-free profiles where CAD scores ranged between -100 and -30, to exclude clouds and very low-confidence aerosol layers. CALIPSO aerosol layer properties have higher uncertainties during the daytime, especially over bright sea ice. As mentioned previously, CALIPSO may also miss dilute Arctic aerosols, even at night, when lidar sensitivity is higher (Zamora et al., 2017). Moreover, aerosol subtype designation (Omar et al., 2009; Kim et al., 2018) can be subject to other errors as well. In general, aerosol type is more difficult to discern in aerosol mixtures (Zeng et al., 2021). Marine aerosols are not allowed in over-land retrievals, but may still be present, at least near coastal locations (Kanitz et al., 2014). Also, desiccated marine aerosols might be misclassified as dust or polluted aerosol at relative humidity below 60-70% (Ferrare et al., 2020), and pollution aerosols can be misclassified as marine aerosols (Di Biagio et al., 2018).

Evaluations of the previous CALIPSO aerosol type version 3 dataset mainly from subarctic data indicate agreement with aerosol type estimates from other data sources most of the time, with best results for mineral dust aerosols (Burton et al., 2013; Papagiannopoulos et al., 2016; Mielonen et al., 2013). However, in the Arctic, diamond dust can be misattributed to the CALIPSO dust aerosol subtype (Di Biagio et al., 2018), and so CALIPSO dust subtyping is likely less certain over this region. It also appears that polluted continental aerosols over the Arctic Ocean may be misattributed to clean marine conditions (Rogers et al., 2014; Di Biagio et al., 2018), although this may not matter at all locations, as the seasonality of the clean marine aerosol subtype seems to be in agreement with long-term observations at Svalbard (Di Biagio et al., 2018). In this study, we used the updated version 4 products for aerosol type classification (Kim et al., 2018), which may result in fewer uncertainties in
CALIPSO aerosol type compared to previous studies. However, large-scale evaluations of CALIPSO version 4 aerosol subtype have not yet been conducted to our knowledge, and even for version 3, such evaluations are limited.

2.1.2 AIRS data

CALIPSO profiles were matched with concurrent temperature (T) and relative humidity (RH) data from the Atmospheric Infrared Sounder (AIRS) L3 Daily Standard Physical Retrieval (AIRS+AMSU) 1° x 1° V007 (AIRX3STD 007) product (AIRS project, 2019) from both the ascending and descending orbits. These data are available in 12-hourly time slots for every 1° x 1° at pressure levels 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150, and 100 mb. The AIRS L3 data are useful under most conditions in the Arctic troposphere, but have fewer errors when there is no heavy precipitation. Data with RH values > 200% were discarded, following AIRS team recommendations (Tian et al., 2020). Seasonally-averaged AIRS data during the study period were obtained separately using the Giovanni online data system (NASA GES DISC, Acker and Leptoukh (2007). For the seasonal data we used the ascending and descending orbits in the AIRX3STM 007 product.

Table 1: Conditions under which diamond dust can form over the Arctic. Heterogeneous ice formation also requires the presence of ice nuclei.

<table>
<thead>
<tr>
<th>Formation mechanism</th>
<th>T (°C)</th>
<th>RHi (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneous formation&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&lt; -35 °C</td>
<td>&gt; 140%</td>
</tr>
<tr>
<td>Heterogeneous formation: Deposition ice nucleation&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt; 0 °C</td>
<td>&gt; 100%</td>
</tr>
<tr>
<td>Heterogeneous formation: homogeneous freezing of preactivated water-containing aerosol pores&lt;sup&gt;c,d&lt;/sup&gt;</td>
<td>&lt; -38 °C</td>
<td>&lt; 100%</td>
</tr>
</tbody>
</table>

<sup>a</sup> Koop et al. (2000); <sup>b</sup> Kanji et al. (2017); <sup>c, d</sup> (Marcolli, 2014, 2017)

Ice crystals, including diamond dust (Table 1), form and grow in the atmosphere differently depending on ambient ice nuclei, temperature, and moisture levels. To identify locations where diamond dust is most likely to form in the presence of aerosols, we follow a similar approach as in Sakai et al. (2003), based on the locations where the equilibrium relative humidity over ice is exceeded (i.e., where ambient values of RH with respect to ice (RHi) are supersaturated, or > 100%). RHi values were calculated from AIRS T and RH values following Murphy and Koop (2005). First, saturation vapor pressure over liquid water (e_s) and saturation vapor pressure with respect to ice (e_{si}) were calculated from the AIRS T values based on the Murphy
and Koop equations 10 and 7, respectively. These equations are valid for $123 < T < 332$ K and down to 110 K, respectively. Then we estimated RH by multiplying this ratio by the relative humidity, following their equation 11:

$$RH_R = RH \frac{\sigma_L}{\sigma_{si}}$$

This approach could underestimate locations where diamond dust occurs, as it does not include, for example, locations with homogeneous freezing of preactivated aerosol pore water at temperatures $<-38$ °C (Marcolli, 2014) (Table 1). Locations where diamond dust forms from small-scale meteorological variations in supersaturation from factors such as vertical velocity (Korolev and Mazin, 2003) and radiative cooling (Zeng, 2018) will also be underestimated with this method.

### 2.1.3 MERRA-2 output

Mineral dust, black carbon (BC), and organic carbon (OC) aerosol concentrations and model mid-layer height output were obtained from MERRA-2 (Global Modeling and Assimilation Office (GMAO), 2015a, b), which has 3-hourly, 0.5 ° x 0.625 ° horizontal resolution. We obtained output at 72 different model levels above the surface, focusing mainly on the lower 29 levels (up to $\sim$10.5 km asl). MERRA-2 aerosol emissions datasets are described in Table 1 of Randles et al. (2017), and do not include local Arctic mineral dust sources. Aerosols are assumed to be externally mixed with different components (e.g., BC, OC, sulfate aerosols) each contributing to total aerosol load. Aerosol loss processes include dry and wet deposition (Randles et al., 2017), with precipitation-induced aerosol deposition based on merged precipitation observations and model products in MERRA-2 (Reichle et al., 2017). Although MERRA-2 assimilates aerosol data when available, its aerosol output is most likely model-driven during polar night. During daytime, some limited aerosol data may be assimilated in non-cloudy conditions, which during the time period of this study includes AERONET, MISR, and MODIS data (Randles et al., 2017).

Mineral dust mass is modelled in five particle size bins (0.1-1, 1-1.8, 1.8-3, 3-6, and 6-10 μm) (Colarco et al., 2010). Dust emissions are wind-driven for each size bin (Randles et al., 2017), parameterized following Marticorena and Bergametti (1995). We used dust from the five size classes and grouped them into two bins: submicron dust (bin 1) and dust from 1-10 microns (bins 2-5).

BC and OC are represented with two (hydrophobic and hydrophilic) mass tracers, which we added together. BC and OC emissions are based on biomass burning and anthropogenic emissions, including from ships (Randles et al., 2017). For reference, the single scattering albedos of BC and OC in MERRA-2 depend on RH, but for BC is in the range of 0.3-0.4 at $\sim$500nm (A. Darmenov, pers. comm.). These simulations did not include brown carbon (BrC), which is expected to more or less track the modelled OC from fires. Aerosol sulfate ($SO_4^{2-}$) data were available in MERRA-2 and contribute to combustion plumes from fires and Arctic haze. For example, within MERRA-2, the total aerosol optical depth (AOD) of a fresh smoke plume might roughly be driven by $\sim$90% OC, with the remaining $\sim$10% partitioned between BC and $SO_4^{2-}$, with the portions changing over time as the plume ages. However, $SO_4^{2-}$ aerosol data were not used in this study as additional tracers for combustion aerosols because there are $SO_4^{2-}$ contributions from other sources such as marine and volcanic emissions.
Details of the MERRA-2 Arctic aerosol distributions have been discussed previously (Wu et al., 2020; Lee et al., 2020; Sitnov et al., 2020). Only a few studies have evaluated MERRA-2 aerosols over the Arctic. MERRA-2 BC and OC aerosols tend to be a bit high compared to observed aerosol concentrations, although they tend to follow the qualitative trends (Vinogradova et al., 2020; Zhuravleva et al., 2020). One study found that dust optical depth and dust extinction were similar or a bit elevated compared to that of CALIPSO, but with large discrepancies in absolute concentrations (both over- and underpredicting concentrations) compared to two ground sites (Wu et al., 2020).

2.1.4 FLEXPART output

Separate simulations of mineral dust, BC, and OC were conducted using the FLEXPART version 10.4 Lagrangian particle dispersion model (Pisso et al., 2019). In the simulations presented here, the model was forced by ERA-Interim meteorological fields from the European Centre for Medium-Range Weather Forecasts (ECMWF) at 1° x 1° spatial and 3-hourly temporal resolution. In addition to dry and wet deposition, FLEXPART accounts for turbulence (Cassiani et al., 2014), unresolved mesoscale motions (Stohl et al., 2005) and includes a deep convection scheme (Forster et al., 2007). Gravitational settling, dry deposition and in-cloud and below-cloud scavenging are also included (Grythe et al., 2017). The resulting daily output has 1° x 1° horizontal resolution with upper vertical layer boundaries at 10, 100, 250, 500, 750, 1000, 1500, 2000, 4000, 6000, 8000, 10,000, 15,000, and 20,000 m agl. For comparison with the other datasets, we converted the FLEXPART output to km asl using surface elevation data from ETOPO1 bedrock GMT4 data (Amante and Eakins, 2009).

Emissions of mineral dust include local Arctic sources and were calculated with the FLEXDUST emission model (Groot Zwaaftink et al., 2016). Dust aerosols were split in 10 size classes in FLEXPART: 0.2, 0.5, 1, 1.5, 2.5, 5, 7.5, 12.5, 15, 20 μm. Emitted dust is assumed to follow the Kok (2011) size distribution. Dust from the different size classes were then grouped into two bins: submicron dust (bins 1-3) and dust from 1-10 μm (bins 4-7) for further analysis. We did not analyse dust with sizes > 10 μm for easier comparison with MERRA-2, which only assessed particles up to this size.

As with MERRA-2, brown carbon was not modelled explicitly, and OC and BC are used as the main proxies for combustion aerosols. BC and OC were run separately, and do not chemically age over time or interact in the model. They were also assumed to be hydrophilic. BC and OC concentrations were calculated from both anthropogenic emissions (using ECLIPSEv6b) and biomass burning (GFED4.1s (Giglio et al., 2013)), following Klimont et al. (2017) but with updated emissions factors (Z. Klimont, pers. comm.). The tracking of BC and OC particles includes gravitational settling for all spherical particles, and BC and OC aerosols have assumed mean diameters of 0.25 μm, a logarithmic standard deviation of 0.3, and a particle density of 1500 kg m⁻³ (Long et al., 2013). The BC and OC emissions datasets may not include some local sources of combustion aerosols.

Details on FLEXPART Arctic aerosol distributions have been discussed previously (Groot Zwaaftink et al., 2016; Stohl, 2006; Eckhardt et al., 2015) and are further evaluated in section 3 below. For now we just note that smoke and pollution transport in FLEXPART have been well validated over the Arctic, and various observations suggest that FLEXPART BC can...
proxy strong, CALIPSO-detectable aerosol layers (Damoah et al., 2004; Eckhardt et al., 2015; Forster et al., 2001; Paris et al., 2009; Sodemann et al., 2011; Stohl et al., 2002, 2003, 2015; Zamora et al., 2017, 2018). In contrast, mineral dust aerosol validation data in the high Arctic are rare, and prior analysis suggests somewhat higher uncertainty in FLEXPART dust aerosols (Groot Zwaaftink et al., 2017, 2016).

2.2 Comparisons between MERRA-2, FLEXPART, and CALIPSO

In order to meet our goal of better understanding the strengths and limitations of MERRA-2 and FLEXPART dust and combustion aerosol products, we focused primarily on environmental conditions where the CALIPSO satellite product does best. These conditions include cloud-free, nighttime cases when diamond dust does not occur, at altitudes > 200 m over the surface. Note that marine aerosols are another very important aerosol source over the Arctic (Schmale et al., 2021). They are not included in this assessment because CALIPSO has uncertainties in detecting near-surface aerosols and small biogenic marine aerosols, as discussed in section 2.1.1.

Table 2: Definitions for CALIPSO, MERRA-2, or FLEXPART concentrated aerosol layers. The MERRA-2 and FLEXPART aerosol layer definitions are based on a range of potential minimum model concentrations, above which CALIPSO aerosols are assumed to be detectable, and the CALIPSO layers are based on a range of potential minimum fraction of the model bin filled with a CALIPSO aerosol layer. The sensitivity of the results to these ranges was tested using a Latin hypercube approach at a later step.

<table>
<thead>
<tr>
<th>CALIPSO</th>
<th>MERRA-2</th>
<th>FLEXPART</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Combustion aerosols</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CALIPSO aerosol layer of type 3 (polluted continental/smoke), 5 (polluted dust), or 6 (elevated smoke) covers 100% of the model vertical bin</td>
<td>&gt;10 to &gt;90% coverage of the model vertical bin</td>
<td>BC: &gt;41 to &gt;100 ng m(^{-3}); OC: &gt;173 to &gt;968 ng m(^{-3})</td>
</tr>
<tr>
<td><strong>Dust aerosols</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CALIPSO aerosol layer of type 2 (dust), 5 (polluted dust), or 7 (dusty marine) covers 100% of the model vertical bin</td>
<td>&gt;10 to &gt;90% coverage of the model vertical bin</td>
<td>Submicron: &gt;0.50 to &gt;1.35 µg m(^{-3}); Supermicron: &gt;2.23 to &gt;6.36 µg m(^{-3})</td>
</tr>
</tbody>
</table>

\(^a\) In cloud-free profiles where Cloud and Aerosol Discrimination (CAD) scores ranged between -100 and -30. Polluted dust can be classified as either combustion or dust.

\(^b\) Based on the 2008 67-92.5% quantile of aerosol concentrations from 0.25-6 km.
Modelled aerosol concentrations are not directly comparable to CALIPSO direct backscatter and polarization observations and CALIPSO-inferred aerosol property information. To derive a parameter such as aerosol concentration from CALIPSO, one would need to make speculative assumptions about the lidar ratio and the extinction cross section of particles, which is beyond the scope of this paper. Therefore, we focused on qualitative comparisons between dust and combustion aerosol distributions among MERRA-2, FLEXPART, and CALIPSO.

First, any aerosol layer with a CALIPSO aerosol classification of either “dust,” “polluted dust,” or “dusty marine” was included in a larger dust group for comparison to other datasets (Table 2). Any aerosol layer with a CALIPSO aerosol classification of either “polluted continental/smoke”, “polluted dust,” “elevated smoke” was included in a larger combustion aerosol group. “Polluted dust” was included in both the larger dust and pollution groupings. Qualitative aerosol distribution comparisons were based on how often CALIPSO combustion or mineral dust aerosol layers were present in cloud-free conditions compared to the most concentrated modelled aerosol layers. Concentrated modelled aerosol layers were defined using a range of dust and combustion aerosol quantile values (Table 2). Because this analysis focuses on aerosol concentration quantile values rather than on absolute concentrations, the aerosol concentration thresholds for determining a “concentrated” aerosol layer are different between MERRA-2 and FLEXPART (see Table 2). This approach has the benefit of enabling us to compare similar numbers of aerosol events with CALIPSO for both models without assuming a priori which one of the two models is more accurate in terms of absolute aerosol concentrations. Avoiding that assumption is important, particularly because of the paucity of in situ validation data for the models over the Arctic at higher latitudes and altitudes.

Because CALIPSO has much finer vertical resolution than either MERRA-2 or FLEXPART, each CALIPSO profile was analysed within vertical bins comparable to either MERRA-2 or FLEXPART model vertical layers. The quality of the MERRA-2 and FLEXPART aerosol spatial distributions relative to those of CALIPSO was then assessed based on false negative (FN) and false positive (FP) rates. The false negative (FN) rate is defined as:

$$FN\ rate = \frac{n_{FN}}{n_p} \times 100\%$$

where $n_p$ is the total number of CALIPSO observations with aerosols of that subtype (dust or combustion), and $n_{FN}$ is subset of those data where modelled aerosol concentrations are less than the threshold in Table 2 for that aerosol subtype. The false positive (FP) rate is defined as:

$$FP\ rate = \frac{n_{FP}}{n_c} \times 100\%$$

where $n_c$ is the number of CALIPSO observations with no aerosol of that subtype and $n_{FP}$ is the subset of those data where modelled aerosol concentrations exceed the Table 2 threshold for that aerosol subtype. The FN and FP rates were calculated separately in each model bin for dust and combustion aerosols and are presented as averages across the 8-year study period. These rates tell us the likelihood of incorrectly identifying clean or aerosol-laden environments based on the definitions in Table 2.

We address uncertainties in this approach as follows:
1) The modelled aerosol concentrations that could form an observable CALIPSO aerosol layer are highly uncertain. Therefore, as detailed in Table 2, we tested a range of threshold values for the assumed minimum aerosol concentration. As shown in Table 2, we did the same for the fraction of a model vertical bin that CALIPSO must fill to be comparable to model aerosol concentrations. The sensitivity of the FN and FP results to these combined assumptions were quantified using a Latin hypercube test (n = 100), following the procedure in Zamora and Kahn (2020).

2) Because the lidar signal-to-noise ratio is smaller during daytime, FN and FP rates were analysed separately during daytime (defined as when the solar zenith angle (SZA) is < 90°). They were also analysed separately when RH is > 100% to assess and reduce potential aerosol type errors from diamond dust.

3) Modelled aerosol layer presence was based on just a few aerosol species (i.e., BC and OC for combustion aerosol layers and super- and submicron dust for dust aerosol layers), but CALIPSO aerosol layer mixtures contain additional species that may vary relative to these constituents. Therefore, if for example, non-carbonaceous constituents in a plume observed by CALIPSO are present at high ratios relative to the modelled BC and OC species, that could lead to higher FN rates for combustion aerosols. Conversely, if the ratios of non-carbonaceous constituents are low relative to the modelled species, that could lead to higher apparent FP rates for combustion aerosols. Similar trends would be expected for dust aerosols, if carbonaceous, biogenic, sulfate or maritime aerosols were elevated or reduced, respectively. However, we are not aware of any evidence to suggest that these uncertainties would lead to systematic biases in FN and FP rates, except near the surface over open ocean where marine particles are more common.

3 Results and Discussion

3.1 Lower troposphere wintertime aerosol distributions are elevated over Siberia

Figure 1 shows the average distribution of MERRA-2 and FLEXPART submicron dust, BC, and OC aerosols at different altitude levels below 4 km at the same locations where CALIPSO cloud-free profiles were available. FLEXPART predicts higher mineral dust concentrations than MERRA-2, which is consistent with its inclusion of local dust sources. FLEXPART BC levels were also elevated relative to MERRA-2, but MERRA-2 had slightly higher OC concentrations over North America. These differences in dust and combustion aerosols may be influenced by different assumptions in emitted particle size distribution and subsequent deposition.

CALIPSO aerosol presence in cloud-free profiles at the same altitude levels as the FLEXPART and MERRA-2 results is also shown in Figure 1. CALIPSO aerosol presence is not directly comparable to MERRA-2 and FLEXPART aerosol concentrations, which for example, can be impacted by high concentrations during infrequent events. However, the general regional trends still provide information on where dust and combustion aerosols are most common. For example, based on
Figure 1: Wintertime occurrence (%) of CALIPSO dust and combustion aerosols in cloud-free profiles below 4 km at two different altitude levels, and the corresponding distributions of MERRA-2 and FLEXPART submicron mineral dust, black carbon, and organic carbon. There is agreement that dust (an efficient source of ice nucleating particles) is high over Siberia. Data at higher altitudes is shown in Fig. S1.
Figure 1, CALIPSO dust aerosols are commonly observed below 4 km over Siberia during wintertime. High average dust concentrations are also predicted by MERRA-2 and FLEXPART in this region, providing confidence in elevated dust levels in the region. In contrast, there is some indication from CALIPSO and FLEXPART that dust sources over the western hemisphere may also be slightly elevated, but below 2 km, there was disagreement on whether these sources are more elevated over the Canadian archipelago (FLEXPART) or over the Labrador Sea (CALIPSO). As there are no known major dust sources over this region, there are higher uncertainties in dust distributions over North America.

CALIPSO, MERRA-2 and FLEXPART also agree that combustion aerosols are elevated over the European and Asian portions of the Arctic, in line with other studies (e.g., Eckhardt et al. (2015); Di Pierro et al. (2013)), and consistent with smoke sources from these locations. However, combustion aerosol layer distributions below are more sharply reduced over oceanic areas in CALIPSO than is predicted in the models. We suspect that this observation is caused by a CALIPSO aerosol subtyping artifact, as a) there are no known dramatic precipitation differences over the land vs. ocean over this large area to explain this phenomenon, b) aerosols are treated differently in the CALIPSO aerosol subtyping algorithm if they are taken over land and ocean (Kim et al., 2018), and c) others have found that polluted continental aerosols over the Arctic Ocean can be misattributed to clean marine conditions (Rogers et al., 2014; Di Biagio et al., 2018). The dust and combustion aerosol trends seen above 4 km are more difficult to discern (Fig. S1), given the inability of CALIPSO to detect very dilute aerosols.

3.2 There is overlap between the distribution of aerosols and the meteorological conditions conducive to diamond dust formation

We next assessed the locations where diamond dust occurs, as these are known to lead to overestimation of dust in CALIPSO (Di Biagio et al., 2018). Table 1 shows the pathways through which diamond dust is most likely to form. Based on Table 1, the homogeneous and deposition ice nucleation pathways for diamond dust formation are most likely to occur when RH$_i$ is > 100%. In line with previous studies (Intrieri and Shupe, 2004; Maxwell, 1982), we do not expect diamond dust formed from the Table 1 pathways to be important during the summer (Figure 2b,d). There are some minor differences in RH$_i$ above 4 km during the summer between ocean and land (Fig. 2b), which are likely due to stronger vertical mixing between more moist surface air and colder stratospheric air over land versus ocean in the summer compared to winter (Stohl, 2006), but generally RH$_i$ conditions are well below 100% during summer at most locations and altitudes.

However, conditions favourable for diamond dust formation from these pathways do occur frequently in some locations during the winter. Estimate that wintertime diamond dust has the potential to confound the dust-CALIPSO comparisons up to 60% of the time (Fig. 2a), especially in low-lying areas of the Arctic Archipelago and the Siberian interior and coast (Fig. 2c). These are extremely cold Arctic locations that also routinely experience moisture transport events (Dufour et al., 2016; Graham et al., 2017), and which also are undersampled by ground-based observations in the winter when diamond dust is most likely to occur. Diamond dust formation may also occur in the frigid conditions near the wintertime tropopause (Fig. 2b), and is particularly likely to occur very near the surface within the stable boundary layer (Intrieri and Shupe, 2004).
Diamond dust is unlikely to occur during the summer and is more likely to occur during winter, especially over the Canadian archipelago and over Siberia, based on seasonally averaged RH$_i$ distributions for the winter (DJF) and summer (JJA) between 2008-2016. On the left are the vertical average RH$_i$ profiles over the Arctic Ocean (solid lines, including areas covered by sea ice) and land regions (dashed lines) for a) winter, and b) summer. For reference, also shown are the percentages of these regions in which RH$_i$ is > 100% (red lines). On the right are RH$_i$ distributions at the 925 mb isobar for c) winter, and d) summer. Also shown in dark grey in c) is the SHEBA cruise track and the region surrounding it. Diamond dust is most likely to occur where RH$_i$ values are > 100% (red regions).

Interestingly, around 1-1.5 km asl over Siberia there is overlap between where MERRA-2, and FLEXPART predict higher mineral dust and combustion aerosols and where high RH$_i$ values also appear. We have some confidence that the dust levels really are elevated in this region because although CALIPSO may misattribute diamond dust to mineral dust at times, MERRA-2 and FLEXPART do not have this source of error. Therefore, based on the data shown in Figures 1 and 2c, we conclude that Arctic aerosol impacts on ice and mixed phase processes are particularly high over Siberia during winter.

Conditions are also favorable for diamond dust formation over the Canadian archipelago during winter. However, the presence of mineral dust in this region in FLEXPART (which includes local dust sources) but not in MERRA-2 (which does not include local dust sources) combined with the uncertainties in dust presence from CALIPSO under conditions favoring diamond dust formation make it harder to determine whether mineral dust aerosols could have a disproportionate impact on ice and mixed phase processes over this region, as is inferred over Siberia.

The locations where we estimate diamond dust formation occurs are consistent with ship-based observations. During the Surface Heat Budget of Arctic Ocean (SHEBA) ship campaign (1997-1998 in the Beaufort Sea), diamond dust was observed 23% of the time between December and February, mainly near the ocean surface (Intrieri and Shupe, 2004). The first wintertime AIRS data over that region were taken starting December 2002, and so are not directly comparable with the SHEBA data. However, based on AIRS data taken during the 2002-2003 winter at 1000 mb within a similar area (between 74.5 and 80.5 °N and -142 and -168 °E, Fig. 2), RH$_i$ exceeded 100% about 26% of the time, which is comparable with how often diamond dust was observed during SHEBA five years prior. Conditions conducive to diamond dust formation are much less
likely to occur in the mid-troposphere in this region (e.g., see distributions in Fig. 2), also in line with SHEBA observations of diamond dust forming predominantly near the surface.

It is important to note that the locations where AIRS RH$_i$ values are > 100% may not capture every instance of diamond dust formation. Diamond dust could occur in locations that are estimated to be subsaturated with respect to ice if, for example, ice particles fall through subsaturated layers, or if the AIRS resolution (reported at 1° in the current study) is too coarse to observe smaller-scale supersaturations (Sakai et al., 2003). As previously mentioned, diamond dust may also occur in conditions favourable for freezing of preactivated water-containing aerosol pores at RH$_i$ values < 100% (third line mechanism, Table 1). Such conditions are presumed less likely to occur during the summer due to the warmer average temperatures, but their impact during colder periods of the year is currently not well known.

### 3.3 Model aerosol validation

Next, we assess how well aerosol distributions in MERRA-2 and FLEXPART compare to the CALIPSO aerosol distributions. Figure 3 shows the MERRA-2 and FLEXPART Arctic area-averaged false negative (FN) and false positive (FP) rates during night and day. We focus the subsequent discussion on times and locations when FP and FN rates are most trustworthy, i.e., during nighttime when RH$_i$ < 100% (grey background, solid lines). We do not focus on daytime data because CALIPSO detects fewer aerosol layers during the day due to reduced sensitivity, which could cause FP rates to be biased high.

These rates can also be biased high when RH$_i$ is > 100% (dashed lines in Fig. 3), because CALIPSO can misclassify diamond dust as mineral dust.

The peak vertical distribution of MERRA-2 dust concentration is inconsistent with where CALIPSO dust aerosols were observed in the vertical column. Based on the FN rate (top left plot in Fig. 3), MERRA-2 dust levels were dilute (defined as being below the 67th % quantile of modeled levels) nearly all of the time when CALIPSO observed dust, with the discrepancy higher closer to the surface. For comparison, based on the MERRA-2 dust FP rate (Fig. 3), dust-free CALIPSO observations were associated with the more concentrated MERRA-2 dust levels ~17% of the time at all altitudes, suggesting that most of the time MERRA-2 can predict dust-free concentrations, independent of altitude. We interpret this to mean that lower troposphere dust sources are underestimated in MERRA-2. FLEXPART, in contrast, showed no strong vertical bias in dust aerosols, with FP and FN rates of < 20% consistently throughout the vertical column, even in conditions conducive to diamond dust formation (Fig. 3).

Combustion aerosol FP and FN rates were consistently < 3% and < 20% respectively throughout the vertical column, based on the 67-92.5% quantile of BC and OC concentrations in each model (Table 2). The low FP rates suggest that most of the time, dilute aerosol concentrations considered too low to be detected by CALIPSO really did correspond to a lack of CALIPSO aerosol identifications of that type. The low FN rates suggest that concentrated combustion aerosol levels in MERRA-2 and FLEXPART corresponded with actual CALIPSO combustion aerosol observations. From that information,
Figure 3: Arctic nighttime area-averaged false negative (FN) and false positive (FP) rates in MERRA-2 and FLEXPART on cloud-free days. Data are shown for nighttime and daytime conditions (grey and white backgrounds, respectively). The sensitivity ranges in estimated FN and FP rates (the pastel colours indicating interquartile range of the Latin hypercube analysis) are calculated from the combined ranges of assumptions shown in Table 2 for RH$_i$ conditions < 100%. The median values for FN and FP rates in RH$_i$ conditions < 100% and > 100% (night only) are shown by the solid and dashed lines, respectively. Model bins with < 20 observations were excluded in the averaging, which excludes < 9% of total model bins. Both models can predict the vertical distribution of combustion aerosols (blue) with confidence, and FLEXPART can predict dust aerosols (red) most of the time, with FP and FN rates of < 10%, even in conditions conducive to diamond dust formation (dashed lines). Vertical distributions of dust in MERRA-2 are less trustworthy near the surface, overpredicting clean cases. This finding suggests that near-surface dust sources may be underestimated.
Figure 4: MERRA-2 and FLEXPART dust and combustion aerosol false negative and false positive rates (%) below 4 km relative to CALIPSO nighttime aerosol layer presence in conditions where diamond dust formation is unlikely (RH_i is < 100%). The FN and FP rates help indicate where and how often aerosols are under- and overpredicted. Locations with < 15 observations are excluded in the above plots. Data from higher altitudes is shown in Fig. S2.
and the fact that there was no bias in the modeled vertical distributions of these aerosols relative to CALIPSO, we conclude that MERRA-2 and FLEXPART are both able to predict the vertical distribution of combustion aerosols with moderate confidence.

Figure 4 shows the distribution of false negative and false positive rates below 4 km, where the majority of aerosols are located in the vertical column (Di Pierro et al., 2013). There are more observations where no dust or combustion aerosols were observed in CALIPSO than when they were observed, and so there are more locations where FP rates could be determined than where FN rates could be determined. For FN rates, dust and combustion aerosols are generally distributed closer to the surface, which is why there are more CALIPSO FN data below 2 km than above 2 km over marine areas.

As shown in Figure 4, FP rates are generally small over most locations, but they indicate that high dust concentrations in MERRA-2 may be predicted too often over the eastern hemisphere. The red colors in the MERRA-2 dust FN rates indicate that the times and locations when MERRA-2 predicted high dust aerosols were not generally located at the same time and place as when CALIPSO observed dust, independent of spatial location. In FLEXPART, the lower FN rates indicated that the distributions of modeled high dust concentrations were much more likely to match with where CALIPSO observed dust aerosols. However, FLEXPART was more likely to miss dust events from 0.2 to 2 km over the Greenland, Norwegian and Barents Seas relative to other Arctic areas. Lastly, combustion aerosol FN rates in both MERRA-2 and FLEXPART seemed to be better predicted at latitudes below 70°N. We hypothesize that this is because the air at lower altitudes is closer to the emission sources (Stohl, 2006), and thus are subject to fewer cumulative uncertainties due to dispersion and deposition. In addition, for MERRA-2, the satellite data feeding into the reanalysis product (which do not include CALIPSO) are more abundant at those lower latitudes, particularly during winter.

### 3.4 The importance of local dust

There are only a few locations where Arctic dust aerosol concentrations have been directly measured over long time periods, but we know that local dust emissions can be important aerosol sources at specific sites, such as near receding glaciers (Bullard and Mockford, 2018; Gassó et al., 2018). In Figure 5, we compare long-term ground-based dust observations at Stórhöfði, Iceland (a site of large local dust emissions) from Prospero et al. (2012) to dust concentrations during the study period from MERRA-2 and FLEXPART. Although FLEXPART tended to somewhat overestimate dust concentrations at this site, MERRA-2 substantially underestimated observed mineral dust concentrations, particularly in the spring when local emissions are highest. Wu et al. (2020) showed similar data at Stórhöfði (Heimaey) for MERRA-2, and also showed that ground dust observations at Alert, Greenland were also underpredicted. However, they also found that MERRA-2 tended to overpredict dust aerosol optical depth in many remote regions of the Arctic compared to Arctic satellite dust optical depths from CALIPSO. Thus, the underprediction of dust at the Icelandic and Greenland sites despite a potential overprediction bias at most other sites underscores a need to improve the relative spatial distributions of dust in the reanalysis product by adding local dust emission sources.
Figure 5: We compared long-term ground-based mineral dust observations at an Icelandic site with large local dust emissions (Stórhöfði, Heimaey) from Prospero et al. (2012) to dust concentrations during the 2008-2015 study period from MERRA-2 and FLEXPART. The whiskers show mean and standard deviation. MERRA-2 (brown) substantially underestimated observed mineral dust concentrations compared to observations (black), particularly in the spring when local emissions are highest. Adding local dust emission sources into MERRA-2 should improve the product.

4 Conclusions

Aerosol models are often used in aerosol-cloud interaction studies over the climatically sensitive Arctic region. Our goals in this study were to: 1) to better understand the strengths and limitations of the Arctic MERRA-2 and FLEXPART dust and combustion aerosol products and to suggest where they might be improved, and 2) to combine satellite, reanalysis, and model products (each with their own limitations and biases) to learn more about where there is high confidence in the concentrations of aerosols, and how that relates to meteorological conditions conducive to the formation of diamond dust.

To summarize the model validation portion of the study, we found that both MERRA-2 and FLEXPART can provide useful information for aerosol studies in the Arctic region at high altitudes, where clouds often make it difficult to observe aerosols and where in situ observations are scarce. They each predict the vertical distribution of combustion aerosols with relatively high confidence, despite some likely CALIPSO combustion aerosol subtyping artifacts over the Arctic Ocean. When using the MERRA-2 product to estimate combustion aerosols, it would be useful to keep in mind that the lack of satellite data for assimilation at high latitudes during winter led to small but measurably higher uncertainties in combustion aerosols poleward of ~70 °N than elsewhere; a similar bias was not observed in the FLEXPART model. Also, near-surface dust sources
in MERRA-2 are lacking, leading to underestimates in dust presence based on comparisons to CALIPSO and ground observations, especially in boreal spring. Adding local dust emission sources into MERRA-2 should improve the distribution of dust in this product. In contrast, FLEXPART, which does include local dust sources, is already able to predict the presence and absence of dust most of the time.

To summarize the second part of the study, we provide evidence that the impacts of mineral dust and combustion aerosols on Arctic ice particle formation may be higher over Siberia than over other Arctic regions. CALIPSO, MERRA-2, and FLEXPART show agreement that Siberia has high wintertime mineral dust and combustion aerosol levels. Moreover, this region is very likely to experience conditions conducive to diamond dust formation frequently; during winter, conditions favourable to diamond dust formation occurred up to 60% of the time in the altitude ranges we assessed. Such conditions hardly ever occurred during summer and are not expected to contribute to CALIPSO dust aerosol uncertainties much during this time. This finding suggests that aerosol effects on clouds in a warming Arctic are changing, as the temporal window and locations where diamond dust and heterogeneous freezing can occur is shrinking, even as local mineral dust sources may be increasing due to retreating glaciers (Prospero et al., 2012).

The study also helps us identify several areas where future research is likely to be particularly fruitful or helpful.

1. Models can help identify where changing temperatures, moisture fluxes and aerosol types and concentrations are most likely to impact future aerosol homogeneous and deposition ice nucleation pathways.
2. Non-CALIPSO platforms (such as Raman lidar) can help better distinguish diamond dust from mineral dust aerosol over Siberia during winter, which will enable better validation of modelled wintertime aerosols, and better overall assessment of mineral dust and combustion aerosol impacts on freezing processes over this region.
3. Aircraft, ground, and ship data can help validate models of aerosols from different sources over remote Arctic locations, and can better assess their contributions to INP budgets. Based on our analysis, some regions of particular interest would be places where local dust emissions are thought to be strongest.
4. Laboratory, satellite, and aircraft data can help better assess the extent to which diamond dust forms through pre-activation processes, which will help elucidate how often diamond dust and ice nucleation actually occur at locations with RHi values < 100%.

Data availability

See AIRS project (2019) for the AIRS data, Winker (2018) for the CALIPSO data, Amante and Eakins (2009) for the ETOPO data, and GMAO (Global Modeling and Assimilation Office (GMAO), 2015a, b) for the MERRA-2 output. Upon submission of this manuscript to ACPD, we will also submit the FLEXPART output for archiving at the NASA Earth Observing System.
Data and Information System (EOSDIS) data repository. If the manuscript is accepted at ACP, a link to those output will be made available in the final manuscript.

Author contributions

LZ and RK designed the experiments and LZ carried them out. NE and CGZ ran FLEXPART and provided simulation output. LZ prepared the manuscript with contributions from all co-authors.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center. Data from Figure 1 were produced with the Giovanni online data system, developed and maintained by the NASA GES DISC. LZ and RK would like to thank H. Bian, A. Darmenov, K.B. Huebert, J. Lee and the ARCSIX science writing team for helpful discussions. LZ and the contributions of RK were supported by the NASA Aerosol-Cloud Modeling and Analysis Program (Grant 80NSSC19K0978) under Richard Eckman. NE was supported by the COMBAT (Quantification of Global Ammonia Sources constrained by a Bayesian Inversion Technique) project funded by NFR’s ROMFORSK – Program for romforskning of the Research Council of Norway (Project ID: 275407).

References


Maxwell, J. B.: The Climate of the Canadian Arctic Islands and Adjacent Waters, Ministry of Supply and Services, Canada, 589 pp., 1982.


