Supplementary information for

Data-driven modeling of environmental factors influencing Arctic methanesulfonic acid aerosol concentrations

Jakob Boyd Pernov[1,a], William H. Aeberhard[2], Michele Volpi[2], Eliza Harris[2,b], Benjamin Hohermuth[3], Sakiko Ishino[4], Ragnhild B. Skeie[5], Stephan Henne[6], Ulas Im[7], Patricia K. Quinn[8], Lucia M. Upchurch[8,9], and Julia Schmale[1]

[1]Extreme Environments Research Laboratory, École Polytechnique Fédérale de Lausanne, Sion, Switzerland
[2]Swiss Data Science Center, ETH Zurich and École Polytechnique Fédérale de Lausanne, Switzerland
[3]Schroders Capital ILS, Zurich, Switzerland
[4]Institute of Nature and Environmental Technology, Kanazawa University, Kanazawa, Japan
[5]CICERO, Center for International Climate Research, Oslo, Norway
[6]Empa, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland
[7]Department of Environmental Science/Interdisciplinary Centre for Climate Change, Aarhus University, Roskilde, Denmark
[8]Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration, Seattle, WA, USA
[9]Cooperative Institute for Climate, Ocean, and Ecosystem Studies, University of Washington, Seattle, WA, USA
[a]Now at: School of Earth and Atmospheric Sciences, Queensland University of Technology, Brisbane, Australia.
[b]Now at: Climate and Environmental Physics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland

Correspondence to: Jakob Boyd Pernov (Jakob.pernov@epfl.ch) and Julia Schmale (julia.schmale@epfl.ch)

1. In situ aerosol observations

At Alert, Canada (82.5° N, 62.4° W, 210 m above sea level (asl)), aerosols, with no upper size limits, were sampled onto filters (20x25 cm Whatman 41) using a high-volume sampler with a nominal duration of seven days. MSA was extracted using deionized water and sonication (>97 % extraction efficiency) (Li and Barrie, 1993). MSA was quantified using a Dionex 4500i ion chromatograph with a 200 μL injection loop, AS4A column, conductivity detector, and an eluent of 5 mM $Na_2B_4O_7$ at 2.0 mL min^{-1}. In between analyses, the column was flushed with 28 mM $Na_2B_4O_7$ and then re-equilibrated with 5 mM $Na_2B_4O_7$. A micromembrane suppressor (H_2SO_4) was included to reduce the baseline conductivity and therefore background noise. The step function for the eluent concentration during analysis was incorporated to avoid interferences from the later elution of stronger anions (Li and Barrie, 1993). The analytical precision and accuracy are listed at 5 and 2 %, respectively, the uncertainty is estimated to be < 13% (Barrie et al., 1989), and the detection limits range from 0.03 to 0.4 ng m^{-3} (Sharma et al., 2012, 2019).

Gruvebadet Observatory, in Ny-Ålesund, Norway on the Svalbard Archipelago (78.9° N, 11.9° E, 50 m asl) has collected aerosol chemical composition since 2010 from March to September with a time resolution of 24 hours. The Thule High Arctic Atmospheric Observatory (THAAO) in Pituffik, Greenland (hereafter referred to as Pituffik/Thule) (76.5° N, 68.8° W, 220 m asl) began sampling aerosol chemical composition in 2010 with a time resolution of 48 hours. For both stations, dominant wind directions indicate that local emissions are not influencing either station (Maturilli et al., 2013; Muscari et al., 2014).

For both Gruvebadet and Pituffik/Thule, the same sampling and analytical methodology was applied. Aerosols of less than 10 micrometers diameter (PM$_{10}$) were sampled onto 47 mm diameter Teflon filters with 2 μm nominal porosity (99 % capture efficiency for 0.3 μm diameter particles) at a flow rate of 2.3 m^3 h^{-1} using a TECORA Skypost sequential sampler
40 following UNI-EN1234 (Becagli et al., 2019). The filters were stored in plastic Petri dishes, frozen, and shipped to Italy for extraction and analysis. Filter quarters were extracted using 10 mL of Milli-Q water (>18 MΩ) and sonicated for 20 min. Three chromatographic systems were used: one system for cations, one for inorganic anions and oxalate, and one for fluoride, MSA, and other low molecular weight organic acids (e.g., acetate, glycolate, propionate, formate, and pyruvate). MSA concentrations were determined by injecting 1 mL (Gilson 222 autosampler) into a Dionex Thermo-Fischer DX600 ion chromatograph
45 utilizing a Thermo-Fischer Dionex TAC-2 pre-concentration column with a 50 mL dead-volume and a Thermo-Fischer Dionex AS11 separation column with a gradient elution of Na$_2$B$_4$O$_7$ from 0.075 mM to 2.5 mM and electrochemical suppression. This allows for the complete separation of MSA from pyruvate (peak resolution = 0.9). A 45 mM Na$_2$B4O$_7$ cleaning solution was injected prior to each analysis to ensure reproducible results (Becagli et al., 2011). Field blanks were always below the detection limit of 0.1 μg L^{-1} (Becagli et al., 2011). The MSA analysis was reproducible with a relative standard deviation of <
50 5% (Becagli et al., 2016, 2019).

The Barrow Atmospheric Baseline Observatory, located near Utqiaġvik, Alaska (referred to as Utqiaġvik/Barrow) (71.3° N, 156.6° W, 10 m asl), is part of the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL) Global Monitoring Division (GMD). The real-time wind direction was used to avoid local pollution by only sampling from the clean air sector (0°–129°) (Delene and Ogren, 2002; Kolesar et al., 2017; Quinn et al.,
55 2002). Aerosols were sampled onto a Berner-type multi-jet cascade impactor with 1 and 10 μm aerodynamic D$_{50}$ cutoff diameters (in this study only submicron chemical composition data was used). Depending on the time of year, sample durations (volumes) of 1 (43 m^3) to 5 (172 m^3) days were collected. Samples were sealed and shipped to NOAA's Pacific Marine Environmental Laboratory (PMEL) for chemical analysis by ion chromatography (Metrohm Compact IC 761). Filters were extracted in 1 mL of spectral grade methanol and then an additional 5 mL of distilled deionized water and sonicated for 30
60 min. A Phenomenex Star IonTM A300 anion column in front of a Metrosep ASUPP5 250/4.0 column with a 1.0 mM NaHCO$_3$ and 5.0 mM Na$_2$CO$_3$ eluent and a 70 mM H$_2$SO$_4$ regenerant was used MSA analysis. The column and eluent combination ensure the separation of MSA from other organic acids (i.e., pyruvic acid). The detection limit was 0.001 mg L^{-1} (Moffett et al., 2020). MSA concentrations are reported at 0 °C and 1013 mbar. The relative uncertainty is ±11% (95% confidence level) (Quinn et al., 2000, 2002, 2009).

65

2.   Models set up

Here we briefly discuss here some modeling choices and approaches we explored but did not retain as they generally performed worse than the RF and AM presented in the main body of the paper. Figure S1 below summarizes the performance comparison. First, we considered various CV schemes to assess out-of-sample prediction error in the training data. Specifically, we considered temporal versus completely random splits. While the latter is generally convenient from an implementation point of view, it did not estimate the prediction error we were interested in. The purpose was to assess the forecasting ability of the models in time, and thus random splits too often led to interpolation in time (with observations directly before and after) rather than extrapolation in time (observations only available before or only available in more distant time windows).

Second, we investigated different transformations of the target MSA concentrations. In conjunction with a mean squared error loss function, this can have a substantial impact on how well a model can predict high peaks versus low values on average. The natural logarithm transformation with a constant of 10^{-3} yielded the best predictive performance overall. The purpose of this constant is to avoid applying a logarithm on a number too close to zero, yielding a very low negative value which in turn can dominate the optimization in the training process. Third, we considered other baseline models in addition to the RF. Ridge regression (Ri in Figure S1) performed generally quite poorly, most likely because many feature effects are far from a linear function. Gradient-boosted regression trees (GB) were nearly on par with the RF, although the latter was generally better.

Finally, for AM we explored various out-of-sample prediction error criteria, specifically the Akaike Information Criterion (AIC in Figure S1) and the Bayesian Information Criterion (BIC). While these information-theoretic criteria have the advantage of being inexpensive to compute we systematically found that AIC led to including too many features in the forward stepwise selection (FSS) procedure (overfit), while the BIC tended to include too few. Both yielded poor generalizations as measured by the temporal cross-validation (CV) mean squared error (MSE). We thus opted to directly use CV MSE as the criterion in the FSS procedure.

3. Contribution of selected features to the model output (Partial effects) for Gruvebadet, Pituffik/Thule, AllStations, and AllStationsFull

In Sect. 3.5 of the main text, we discussed the contribution of the selected features to the model output for Alert and Utqiaġvik/Barrow in detail and summarized the remaining stations and merged datasets, while here we provide a more detailed description.

3.1. Gruvebadet

For Gruvebadet, AM selected the following features, OH_BL_0.1, ChlA_1.2, DMS_4.5, and LSRR_1.2.

OH_BL_0.1 shows a linearly, positive effect on the model output of MSA_p from low levels to ~0.75×10^{-5} ppbv, thereafter a plateau is reached and only increases MSA_p above 2×10^{-5} ppbv (Fig. S8a), although there are few data points in this range. This shows an OH-limited regime where MSA_p increases proportionately with OH_BL_0.1 and a plateau regime where MSA_p is insensitive to increasing OH_BL levels. The plateau regime could be due to MSA_p production is likely limited by other environmental factors, such as DMS and/or its intermediates, warm temperatures affecting the oxidation pathways of DMS to SO_2 (gas-phase MSA formation is favored over SO_2 at lower temperatures), or steady-state conditions where MSA is formed at a similar rate as it is further oxidized (Barnes et al., 2006; Chen et al., 2018; Shen et al., 2022).

DMS_4.5 displays a mostly linearly, positive relationship with model output of MSA_p, although the slope of this relationship is steeper at lower and higher regimes of DMS emissions (Fig. S8b). This could indicate the presence of a DMS-limited regime (over the range of 0-200 kg), a DMS-insensitive regime where MSA_p is only slightly increasing with increased DMS emissions (likely limited by other factors such as oxidant levels) from ~200-600 kg, and a DMS-abundant regime where the model output of MSA_p again becomes sensitive to DMS emissions after 600 kg. Mansour et al. (2024) used a Gaussian process regressor to reconstruct MSA concentrations in the North Atlantic with good accuracy (R^2 values up to 0.86) and showed that the relationship between DMS ocean-air flux and model output of MSA is exponential and only significantly contributes after a certain level, similar to the results shown here.

Exposure of air masses to chlorophyll-a (ChlA) 1-2 days prior to measurement shows a linearly, positive relationship to model output of MSA_p from the lowest values up to ~4 mg m^{-3} (Fig. S8c). After 4 mg m^{-3}, the slope of the relationship between model output of MSA_p and ChlA becomes negative although the few data points in this range do not invoke confidence in this pattern. This shows that while DMS emissions 4-5 days back along the air mass are important for MSA_p model output, air masses recently spending time over biologically active open waters closer to the measurement site are also important.

For air masses recently experiencing precipitation (LSRR_1.2), a sharp reduction in model output of MSA_p is noticed with a slight increase in precipitation (Fig. S8d). After this initial reduction of MSA_p with increasing precipitation, a plateau is reached where increasing precipitation does not remove additional MSA_p up to ~2 mm. This relationship has been observed at Arctic and sub-Arctic sites (Isokääntä et al., 2022; Tunved et al., 2013). After this level, a gradual decrease in MSA_p is observed with increasing precipitation from ~2-3 mm. The initial sharp reduction in MSA_p with a slight increase in precipitation is likely due to the removal of the activated particles containing MSA which formed cloud droplets, these droplets

would be the first to precipitate during a rain event. With increasing precipitation after this initial reduction, the model output of MSA_p plateaus, which could be due to the remaining MSA-containing-particles do not possess the proper microphysical properties to be removed efficiently by precipitation. This could be due to particles either being too small or hydrophobic to nucleate into cloud droplets as well as being interstitially scavenged within a cloud or being scavenged below-cloud (or impaction scavenging) by falling raindrops (Andronache, 2003; Zanatta et al., 2023). After ~2 mm, the model output of MSA_p gradually decreases, which could be due to intense precipitation removing all particles in the atmosphere regardless of microphysical properties.

3.2. Pituffik/Thule

For Pituffik/Thule, AM selected the following features, SSRD_0.1, OPEN_WATER_4.5, and OH_FT_1.2.

SSRD_0.1 shows a minimum effect on the model output of MSA_p at ~200 W m^{-2} where after this value a linearly, positive relationship is observed which begins to plateau at high values of SSRD (Fig. S9a). Below 200 W m^{-2}, the model output of MSA_p increases with decreasing values of SSRD, although with few data points contributing to this pattern. This relationship resembles the one observed for Alert (Fig. 7a) but the model output of MSA_p for Pituffik/Thule does not decrease for very high values of SSRD. The beginning of plateau for the partial effects at high values of SSRD suggests that if Pituffik/Thule experienced higher values of SSRD then a more similar relationship compared to Alert might be observed.

The partial effects for the time air masses spent within the BL and over open water (OPEN_WATER, defined as sea ice concentration < 20 %, Table 2) for 4-5 days prior to measurement, used here as a source-related feature for MSA_p, shows an initial sharp increase in the lowest range of values and a mostly linearly, positive relationship afterward until ~25 sec km^{-2} (Fig. S9b). At higher values, the model output of MSA_p decreases which could be indicative of air masses arriving from marine areas that experienced moisture uptake and thus precipitation. Pituffik/Thule experiences almost exclusively air mass transport from Baffin Bay (Becagli et al., 2016, 2019) and lower Baffin Bay does experience frequent rainfall days (Boisvert et al., 2023), supporting this postulate. Another explanation could be these air masses originated from areas experiencing low biological activity at the time of air mass contact (thus minorly up-taking precursor material), although the low data density in this range shows a higher uncertainty.

The partial effects for OH_FT_1.2 show little variation, reaching a maximum between ~1 and ~2 × 10^{-5} ppbv, plateauing until reaching ~4.5 × 10^{-5} ppbv, and decreasing for higher values (Fig. S9c). The maximum partial effects for OH_FT_1.2 at ~1 × 10^{-5} ppbv could indicate an oxidant limited regime for DMS oxidation and maximum production occurs with OH_FT_1.2 levels between ~1 and ~2 × 10^{-5} ppbv. The plateau of the partial effects between ~2 and ~4 × 10^{-5} ppbv could indicate that MSA_p production is less sensitive to OH levels in this regime and possibly limited by other environmental factors (levels of DMS or its intermediates). The decrease of the partial effects after 4.5 × 10^{-5} ppbv could indicate the continued oxidation of MSA to sulfate (similar to the partial effects of OH_FT_1.2 at Alert, Fig. 7e). The selection of OH in the FT over the BL could indicate processes in the FT are relatively more important at Pituffik/Thule than BL processes but given that Pituffik/Thule is located ~200 m asl this is not unexpected (Table 1).

3.3. AllStations

The selected features for the AllStations subset (AS), with an equal number of test observations from each station, are OH_BL_1.2, WS_BL_0.1, DMS_3.4, LSRR_2.3, RT_BL_0.1, Q_FT_2.3, and ChlA_2.3 (Table 4). Each of these features was selected by AM for the individual stations with the exception of WS_BL_0.1 and RT_BL_0.1. Model output of MSA_p is sensitive to OH_BL_1.2 at low values and insensitive at higher values which could possibly be due to MSA production being oxidant limited at low values (Fig. S11a), similar to the relationship observed for Gruvebadet for OH_BL_0.1 (Fig. S8a). WS_BL_0.1 exhibits a linearly, negative relationship with model output of MSA_p (Fig. S11b), showing higher wind speeds act to decrease MSA_p concentrations either through dilution in a deeper boundary layer or increased dry deposition. This relationship agrees with other studies investigating the drivers of Arctic MSA showing a negative relationship (Song et al., 2022). Our recent study on Pan-Arctic MSA mainly displays a U-shaped relationship between wind speed and MSA_p, which was attributed to low and high wind speed affecting gas- and aqueous-phase oxidation, respectively, although the absolute magnitude of wind speed's effect on MSA was low (Pernov et al., 2024). This current study suggests that for models trained on the merged Pan-Arctic dataset, low wind speeds are generally conducive towards MSA_p production in the model, possibly due to low BLHs confining precursors and oxidants into a smaller volume of air. DMS_3.4 shows an overall positive relationship with the model output of MSA_p, although non-linear (Fig. S11c), which resembles the DMS partial effects observed at Alert, Gruvebadet, and Utqiaġvik/Barrow. The partial effects of LSRR_2.3 on model output of MSA_p resembles one of exponential decay, indicating that precipitation acts to remove MSA_p in the model but the efficiency decreases with increasing intensity (Fig. S11d) as observed for Alert, Gruvebadet, and Utqiaġvik/Barrow (after a certain threshold) as well as at other high-latitude sites (Isokääntä et al., 2022; Tunved et al., 2013). The retention time of air masses within the boundary layer 0-1 days prior to measurement (RT_BL_0.1) has two local maxima: one at $\sim 0.5 \times 10^6$ sec and another at $\sim 3 \times 10^6$ sec, with the latter having a greater impact on the model output of MSA_p (Fig. S11e). These two local maxima could possibly indicate the effect of both boundary layer and free troposphere processes contributing to the model output of MSA_p as discussed above. Q_FT_2.3 displays two local maxima: one at ~ 0.001 and another at ~ 0.00375 kg kg^{-1} (Fig. S11f), likely suggesting gas-phase oxidation at lower values and aqueous-phase oxidation at higher values. The partial effects for Q_FT_2.3 for AS show a similar pattern as does LWC_BL_2.3 and LWC_FT_0.1 (Fig. 10e and f) and slightly similar to Q_BL_0.1 (Fig. 10a) as those observed at Utqiaġvik/Barrow. ChlA_2.3 exhibits a local maximum at ~ 20 mg m^{-3}, which slightly decreases afterwards until ~ 60 mg m^{-3}, the partial effects increase after this level but there are few data points in this region (Fig. S8g).

3.4. AllStationsFull

For the AllStationsFull (ASF) subset, utilizing all test data from all stations, the selected features are OH_BL_0.1, DMS_2.3, WS_BL_0.1, LSRR_2.3, DMS_4.5, OPEN_WATER_0.1 (Table 4). OH_BL_0.1 displays a sharp increase in the model output of MSA_p up to $\sim 0.5 \times 10^{-5}$ ppbv and plateaus afterward (Fig. S12a), similar to the AS OH_BL_1.2 relationship (Fig. S11a) and likely due to similar reasons. DMS_2.3 shows an almost linear, positive relationship with model output of

MSA$_p$ for ASF (Fig. S12b) as opposed to the non-linearity present in the AS subset for DMS_3.4 (Fig. S11c). WS_BL_0.1 shows a linearly, negative relationship with model output of MSA$_p$ that reaches a minimum at ~10 m s^{-1} and increases slightly thereafter (Fig. S12c). These partial effects suggest that low wind speeds act to enhance MSA$_p$ concentrations in the model output and that higher wind speeds slightly increase MSA$_p$, which could be due to enhanced transport of MSA$_p$-containing air masses, enhanced oceanic nutrient mixing, or elevated DMS emissions (Becagli et al., 2016, 2019). LSRR_2.3 exhibits a linearly, negative relationship with model output of MSA$_p$ with some complex structure present (Fig. S12d) as opposed to the exponential decay observed for AS for the same feature (Fig. S11d). DMS_4.5, similar to DMS_2.3, displays a linearly, positive relationship to model output of MSA$_p$ (Fig. S12d), showing the straightforward impact of MSA precursors on all Arctic stations. OPEN_WATER_0.1 shows a rather flat relationship with the model output of MSA$_p$ with an initial increase and local maximum (minimum) at ~250 (~1500) sec km^{-2} and increases thereafter (Fig. S12f). This timestamp of the OPEN_WATER feature (0-1) was only selected for the ASF subset (Table 4). Each station has different surrounding marine environments and the rather flat relationship could be a result of a smoothing effect induced by the model due to the merger of all test data, given that the effect of OPEN_WATER_0.1 would likely be vastly different at Alert (inland station usually surrounded by ice-covered oceans) versus Gruvebadet (coastal station with open water located to the south).

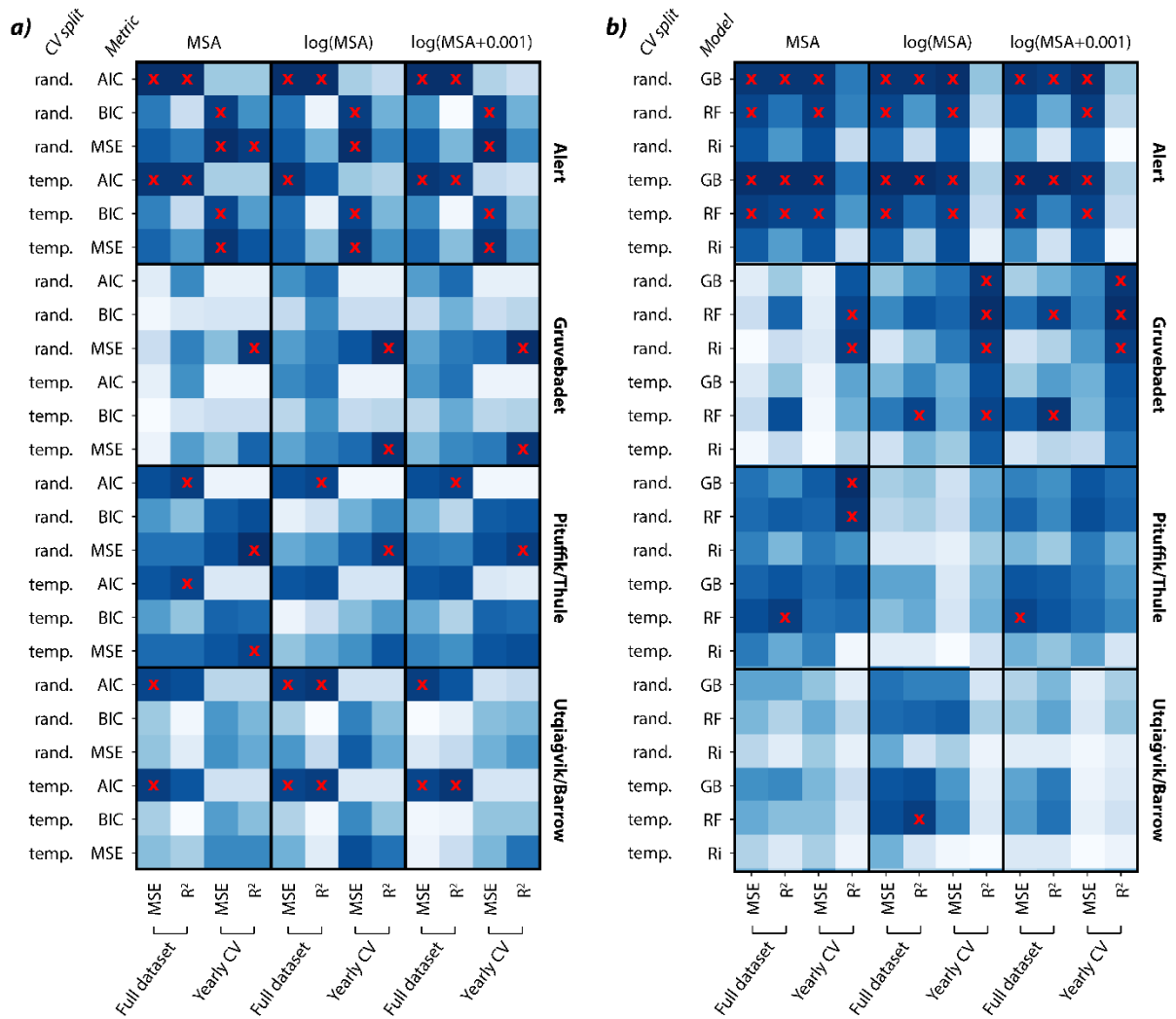



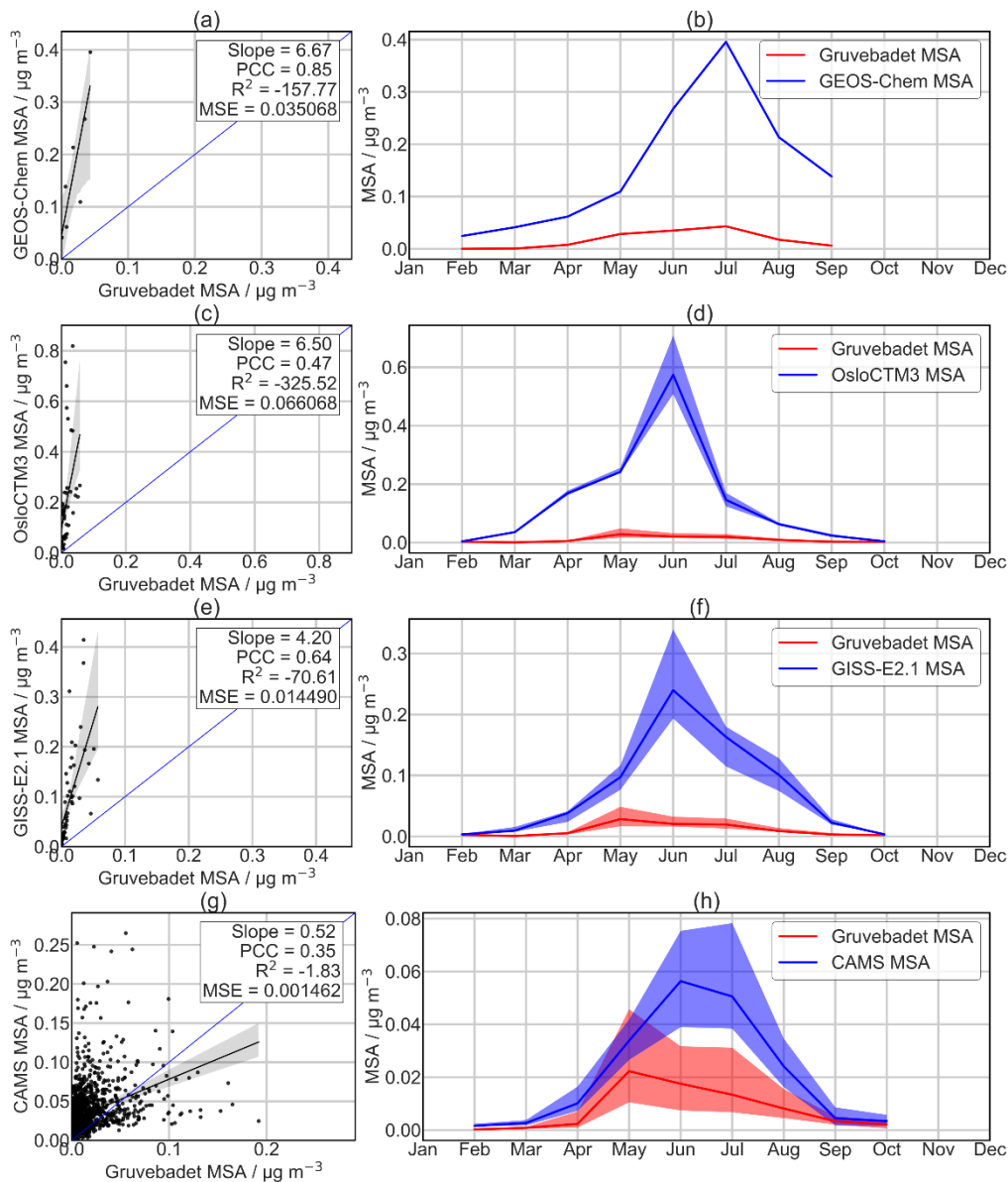
Figure S1: Overview of model performance for the AM (a) and RF (b) models using the Group A+B set of features.
'CV split' refers to the cross-validation split used during feature selection (random or temporal). 'Metric' in (**a**) refers to the metric used to evaluate the AM within the cross-validation during feature selection: AIC = Akaike information criterion, BIC = Bayesian information criterion, MSE = mean squared error. 'Model' in (**b**) refers to the different baseline models tested: RF = random forest regression tree, Ri = ridge regression. Model performance (MSE and R^2) is shown for the models run using the selected features and then trained and then tested on the full dataset ('Full dataset') and using a yearly cross-validation scheme ('Yearly CV') for the original data (MSA$_p$) and the log and log+constant transformed data (log(MSA) and log(MSA + 10^{-3}) respectively). The color scale shows the ranking of the performance for the column where darkest = best performance and the best-performing models in each column are highlighted with an overlaid red cross.

215

Figure S2. Comparison of modeled MSA against in situ MSA observations from Gruvebadet. Scatterplots on the left compare only April to September (over the available period for each station) with the 1:1 line in blue, linear fit in black, 95% confidence intervals estimated through bootstrapping in the shading and seasonal cycles on the right (the thick line is the median and shading is the interquartile range) for GEOS-Chem (**a** and **b**), OsloCTM3 (**c** and **d**), GISS-E2.1 (**e** and **f**), and

220 CAMS (**g** and **h**). The MSE, R^2, and PCC values are calculated according to Eqs. (1), (2), and (3), respectively.

Figure S3. Comparison of modeled MSA against in situ MSA observations from Pituffik/Thule. Scatterplots on the left compare only April to September (over the available period for each station) with the 1:1 line in blue, linear fit in black, 95% confidence intervals estimated through bootstrapping in the shading and seasonal cycles on the right (the thick line is the median and shading is the interquartile range) for GEOS-Chem (**a** and **b**), OsloCTM3 (**c** and **d**), GISS-E2.1 (**e** and **f**), and CAMS (**g** and **h**). The MSE, R^2, and PCC values are calculated according to Eqs. (1), (2), and (3), respectively.

230

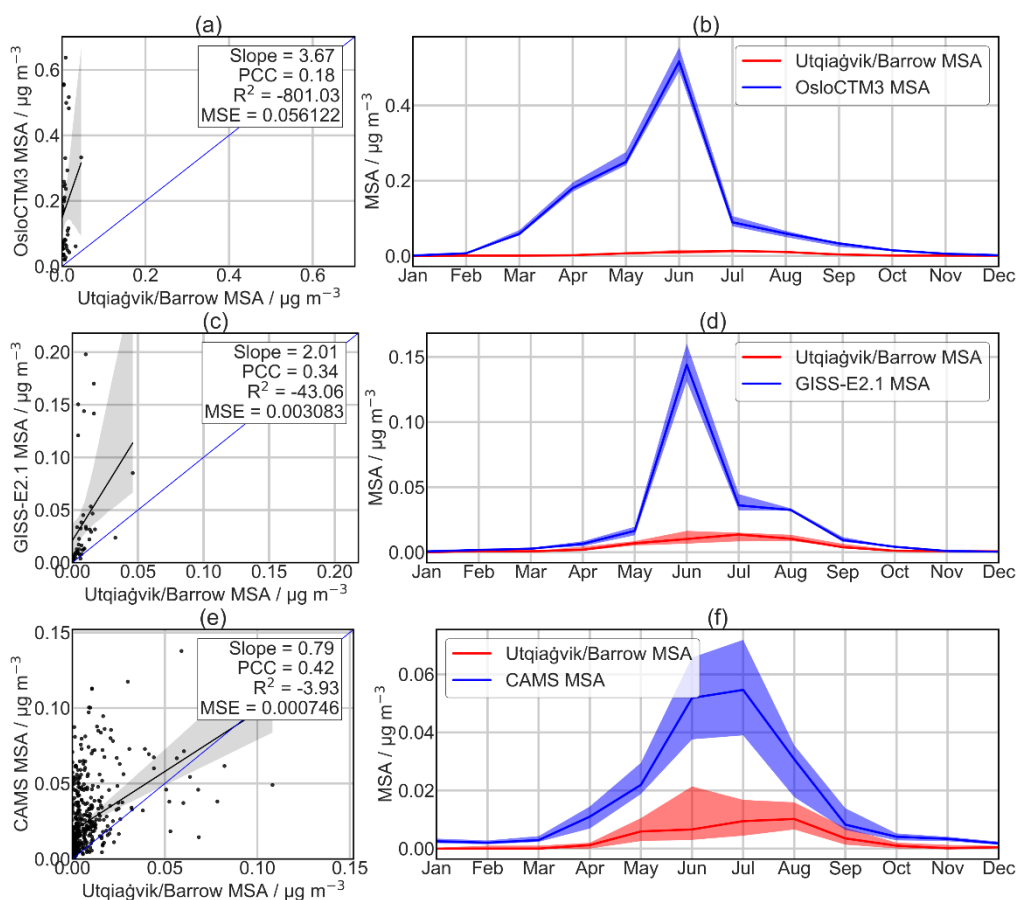



Figure S4. Comparison of modeled MSA against in situ MSA observations from Utqiaġvik/Barrow. Scatterplots on the
left compare only April to September (over the available period for each station) with the 1:1 line in blue, linear fit in black,
95% confidence intervals estimated through bootstrapping in the shading and seasonal cycles on the right (the thick line is the
median and shading is the interquartile range) for OsloCTM3 (**a** and **b**), GISS-E2.1 (**c** and **d**), and CAMS (**e** and **f**). The MSE,
R^2, and PCC values are calculated according to Eqs. (1), (2), and (3), respectively.

240

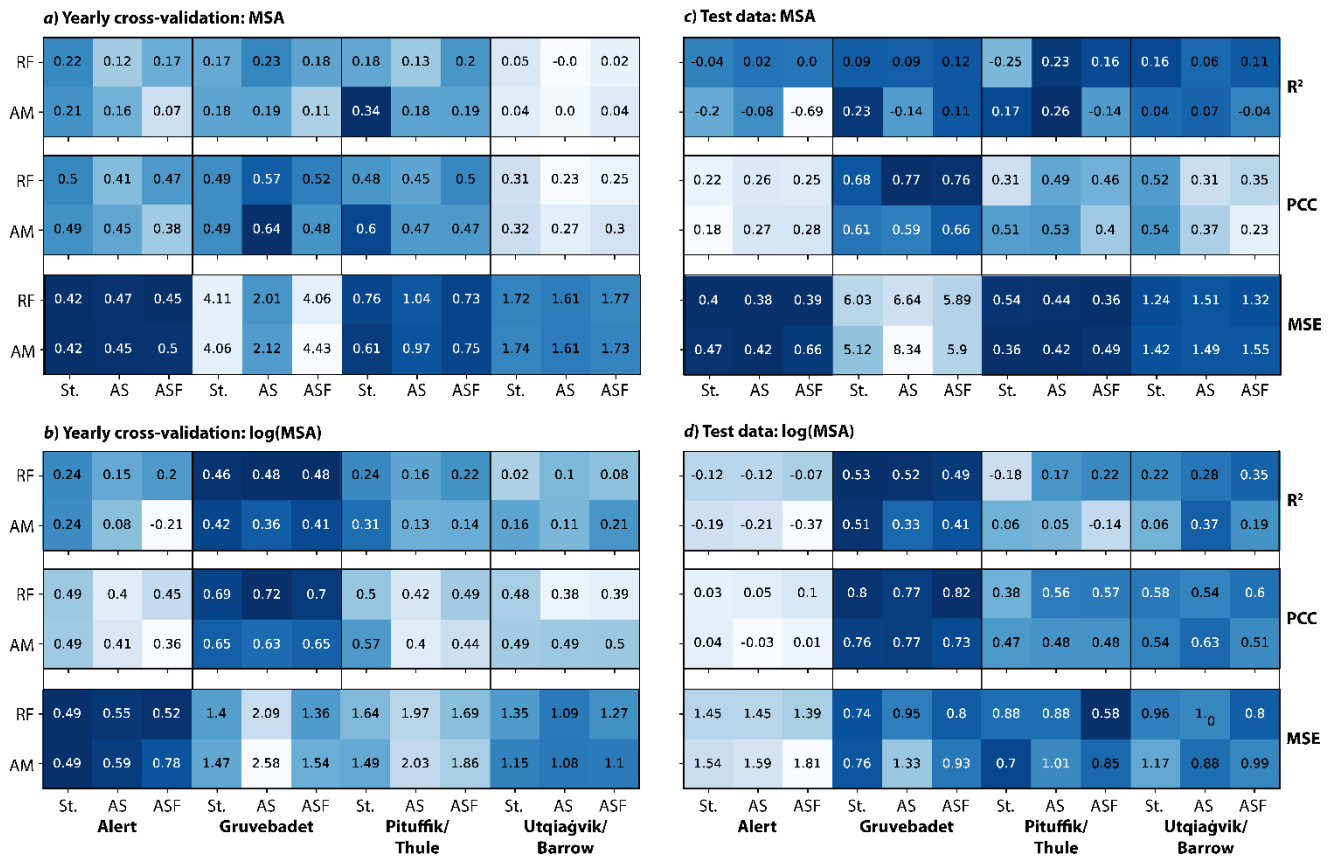



Figure S5: Prediction performance for the temporal cross-validation (CV) scheme and on the test set for the four stations, using the selected features from Group A for the RF and AM. (**a**) and (**b**) show CV performance on original and log scales, respectively. (**c**) and (**d**) show performance on the test set on original and log scales, respectively. In each panel, R^2 is shown in the top sub-panel, the PCC in the middle sub-panel, and the MSE at the bottom. St refers to a model trained and tested on the specified station, AS refers to a subset of the data with an equal number of observations from each station, and ASF refers to all data from all four stations and tested only on the specified station. MSE is multiplied by 10^4 to easily display three significant digits. The color scale indicates performance, where the darkest blue signifies the best performance (lowest MSE, highest R^2, and highest PCC within each row).

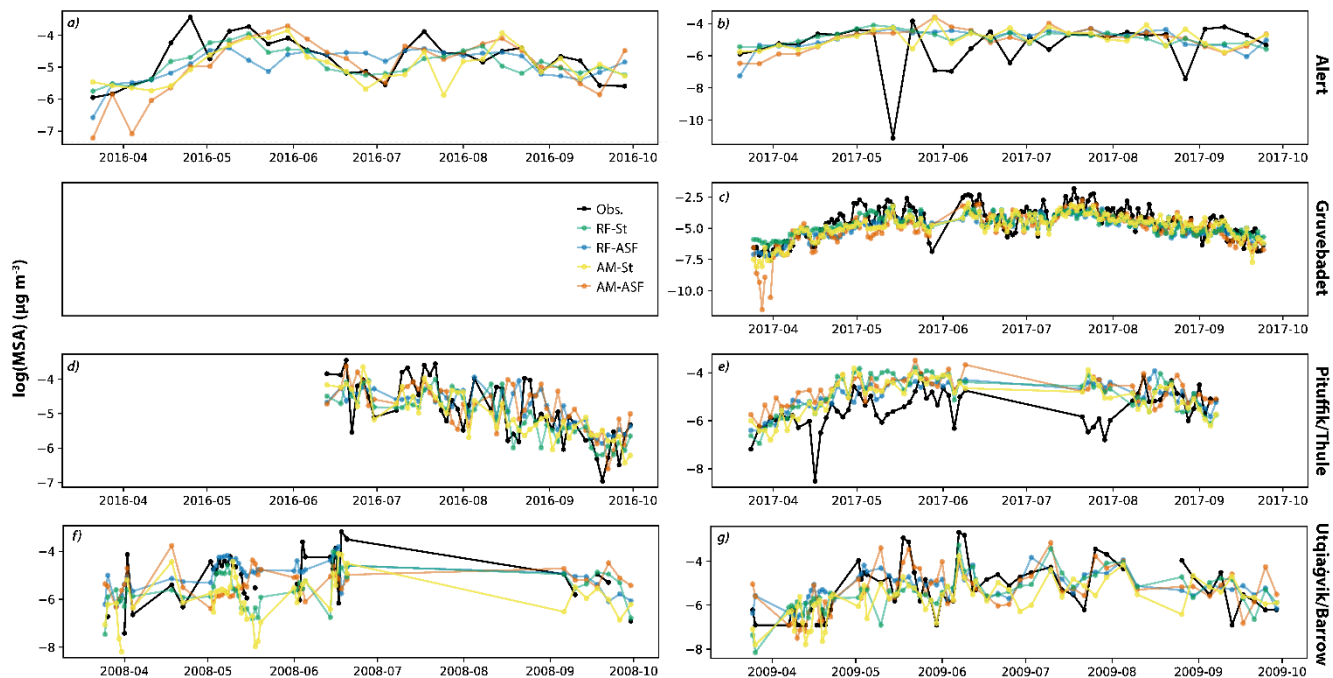



Figure S6: Observed and modeled time series of log-transformed MSA for the test dataset. (**a**) and (**b**) Alert, (**c**) Gruvebadet, (**d**) and (**e**) Pituffik/Thule, and (**f**) and (**g**) Utqiaġvik/Barrow. RF-St and AM-St refer to models trained and tested on the specified station; RF-ASF and AM-ASF refer to models trained on the ASF merged dataset and tested on the specified site dataset. The observations are shown in black.

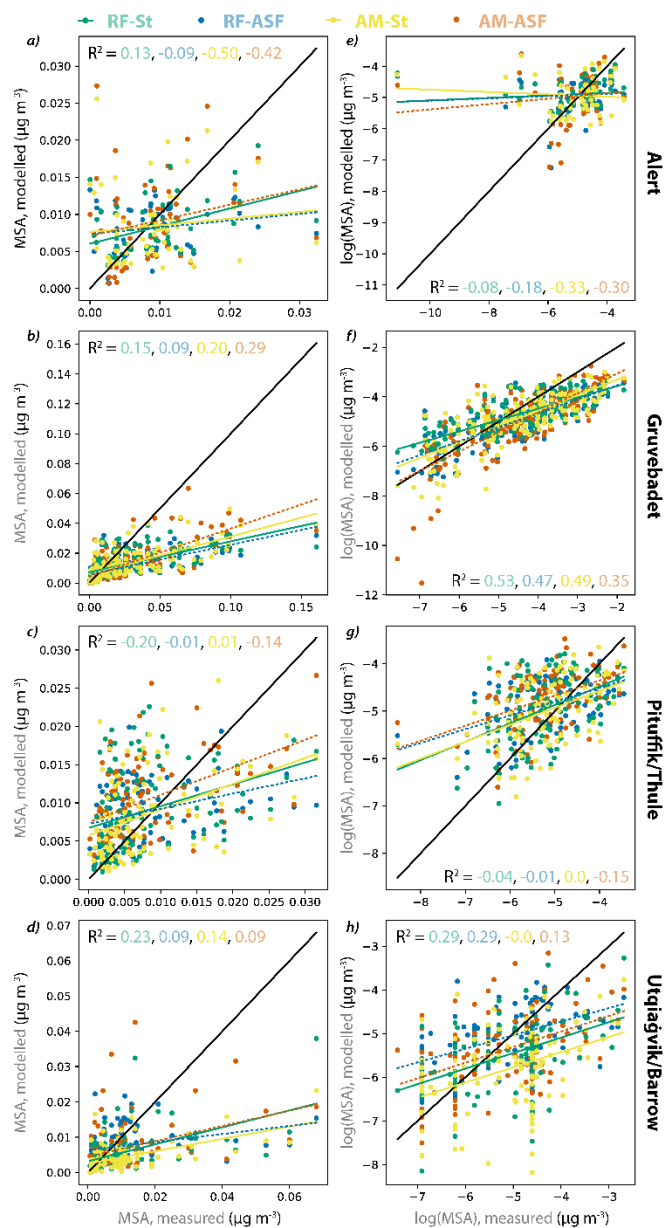



Figure S7: Scatterplots of modeled vs observed MSA$_p$. The left column (**a-d**) is for the original scale and the right column (**e-h**) is for the natural logarithm-transformed data for the RF (green and blue) and AM (yellow and orange) models for the four stations (row). The colored dots and lines indicate the model and data subset used with St referring to models trained and tested on the specified station, AS refers to a subset of the data with an equal number of observations from each station, and ASF refers to all data from all four stations and tested only on the specified station. The dashed lines are the ASF subsets. The black solid line shows a 1:1 line, and the linear regression lines for each model are shown, with the R^2 indicated in the legend.

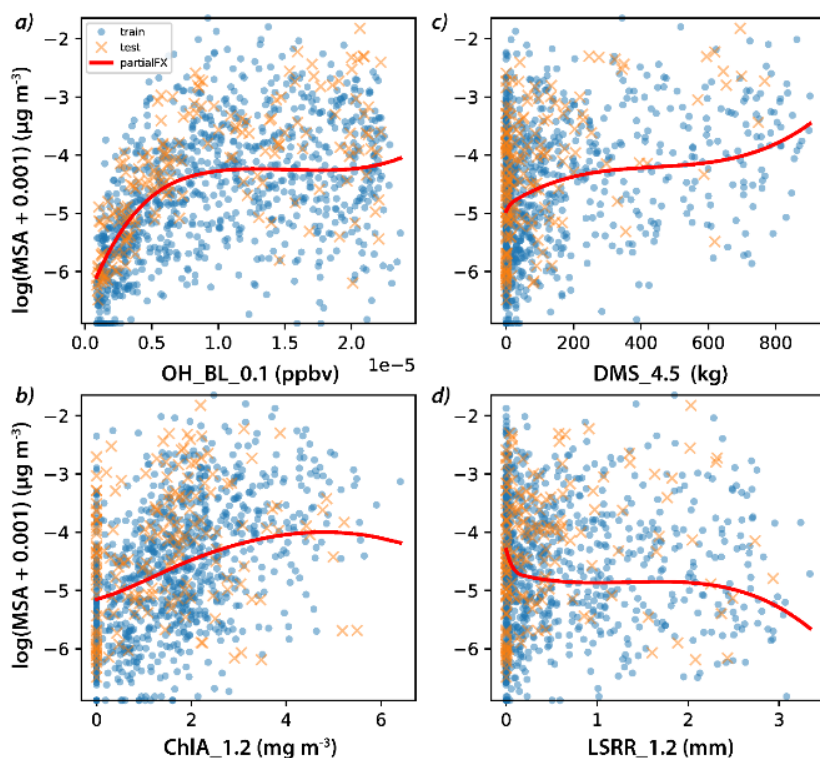



Figure S8: AM-St partial effects for the selected features at Gruvebadet. The red solid line is the partial effect for a different feature in each panel, the blue points are the training observations, and the orange crosses are the test data. Feature abbreviations are defined in Table 2. St refers to models trained and tested on the specified station. Features aggregated as sums over filter time windows (see Table 2) are rescaled here by the average number of 3-hourly time steps to help compare partial effects between stations.

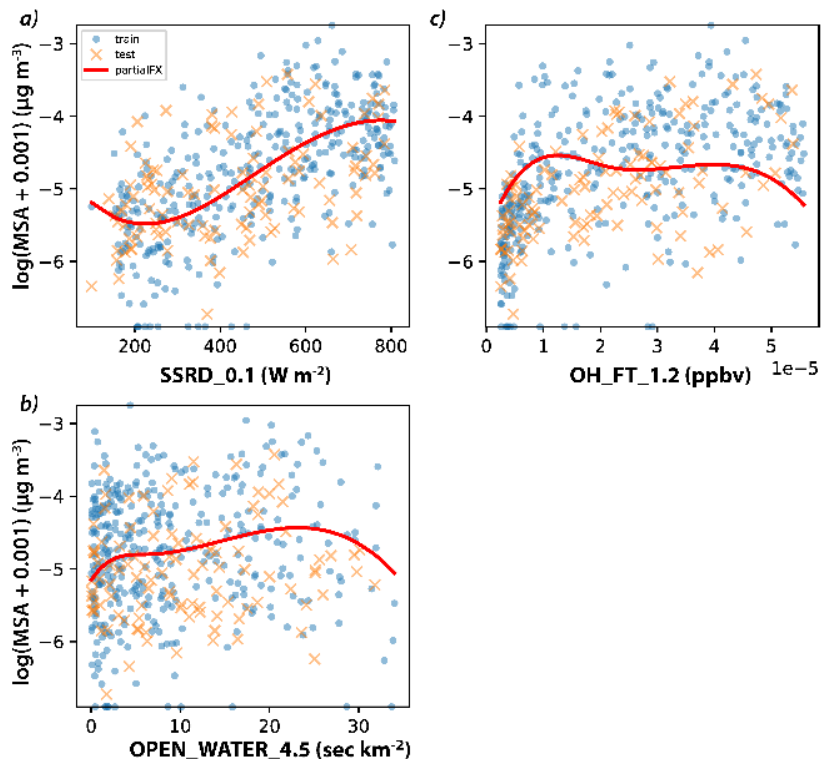



Figure S9: AM-St partial effects for the selected features at Pituffik/Thule. The red solid line is the partial effect for a
different feature in each panel, the blue points are the training observations, and the orange crosses are the test data. Feature
abbreviations are defined in Table 2. St refers to models trained and tested on the specified station. Features aggregated as
sums over filter time windows (see Table 2) are rescaled here by the average number of 3-hourly time steps to help compare
partial effects between stations.

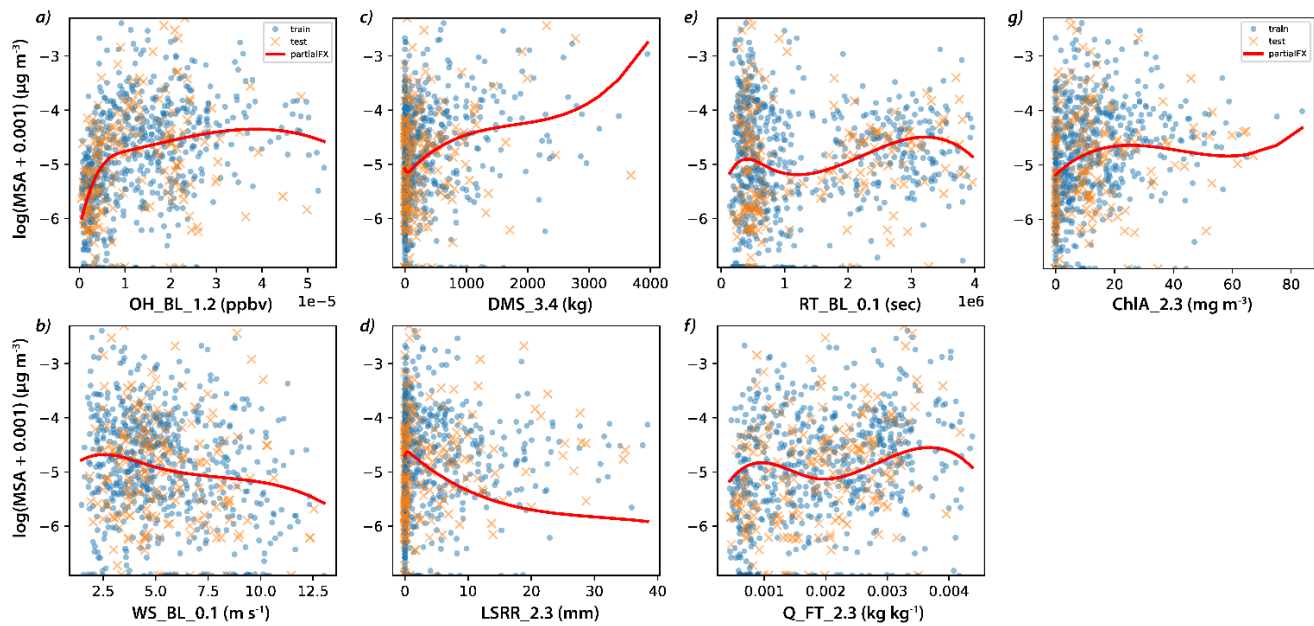



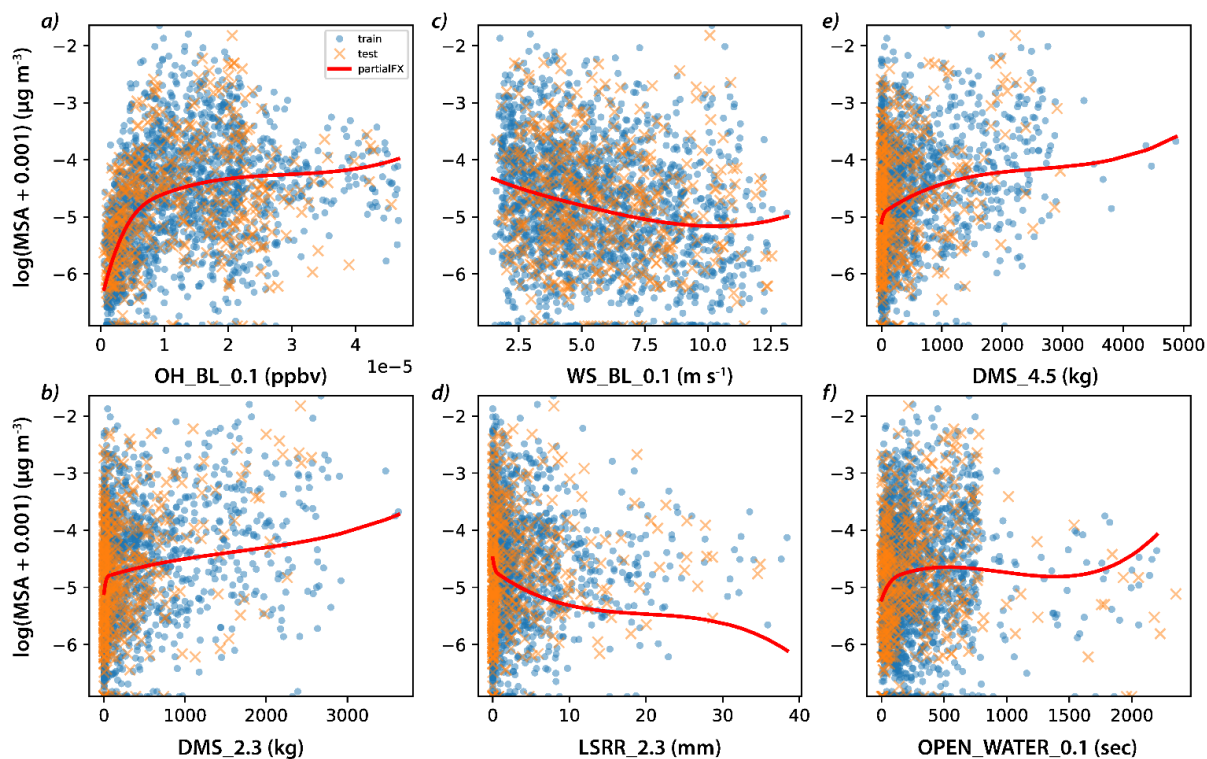
Figure S10: AM partial effects for the selected features for the AllStations dataset. The red solid line is the partial effect for a different feature in each panel, the blue points are the training observations, and the orange crosses are the test data. Feature abbreviations are defined in Table 2 of the main article. Features aggregated as sums over filter time windows (see Table 2) are rescaled here by the average number of 3-hourly time steps to help compare partial effects between stations.

305

Figure S11: AM partial effects for the selected features for the AllStationsFull dataset. The red solid line is the partial effect for a different feature in each panel, the blue points are the training observations, and the orange crosses are the test data. Feature abbreviations are defined in Table 2 of the main article. Features aggregated as sums over filter time windows (see Table 2) are rescaled here by the average number of 3-hourly time steps to help compare partial effects between stations.

310

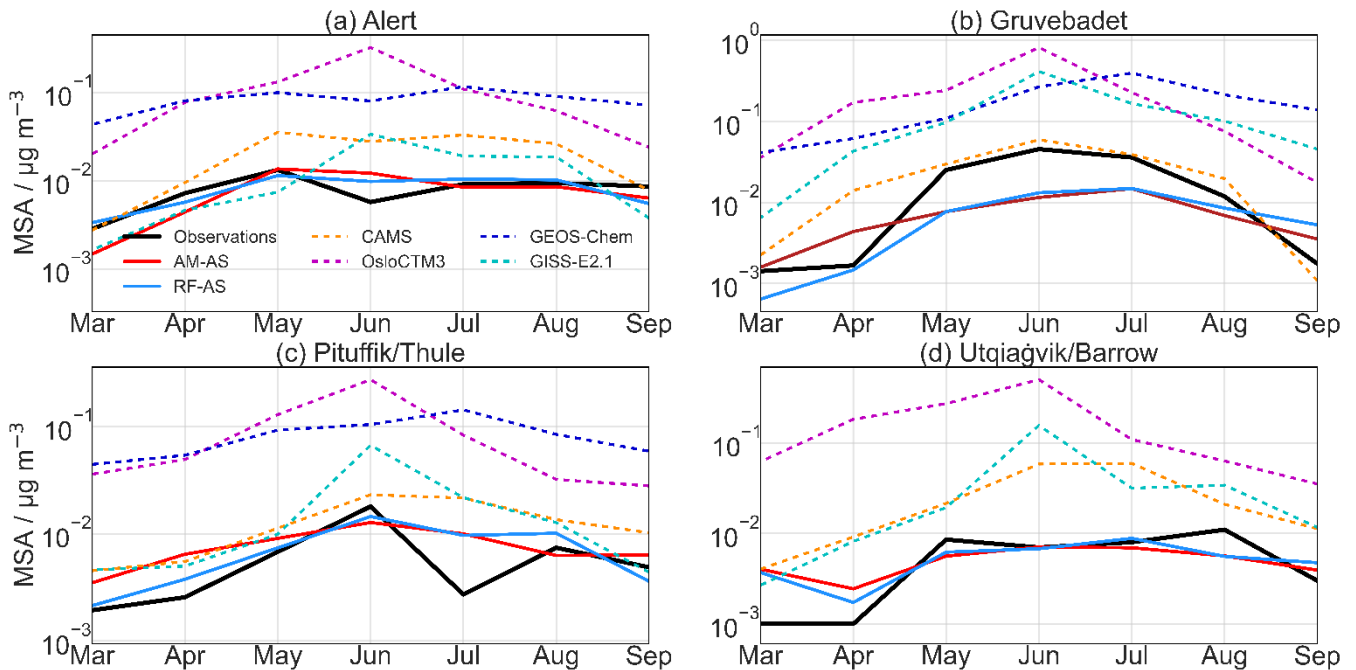



Figure S12: Comparison of seasonal cycles for observations, AS data-driven models, and numerical models. Monthly medians for observations (solid black), data-driven models (AM-AS in solid red and RF-AS in solid light blue), CAMS (dashed orange), GEOS-Chem (dashed dark blue), GISS-E2.1 (dashed cyan), and OsloCTM3 (dashed magenta) for (**a**) Alert, (**b**) Gruvebadet, (**c**) Pituffik/Thule, and (**d**) Utqiaġvik/Barrow. Only data for the tests were included in this analysis for a fair comparison, see Table 3 for dates. AS refers to a subset of the data with an equal number of observations from each station. The evaluation metrics for each data-driven and numerical model against in situ observations are given in Fig. 10.

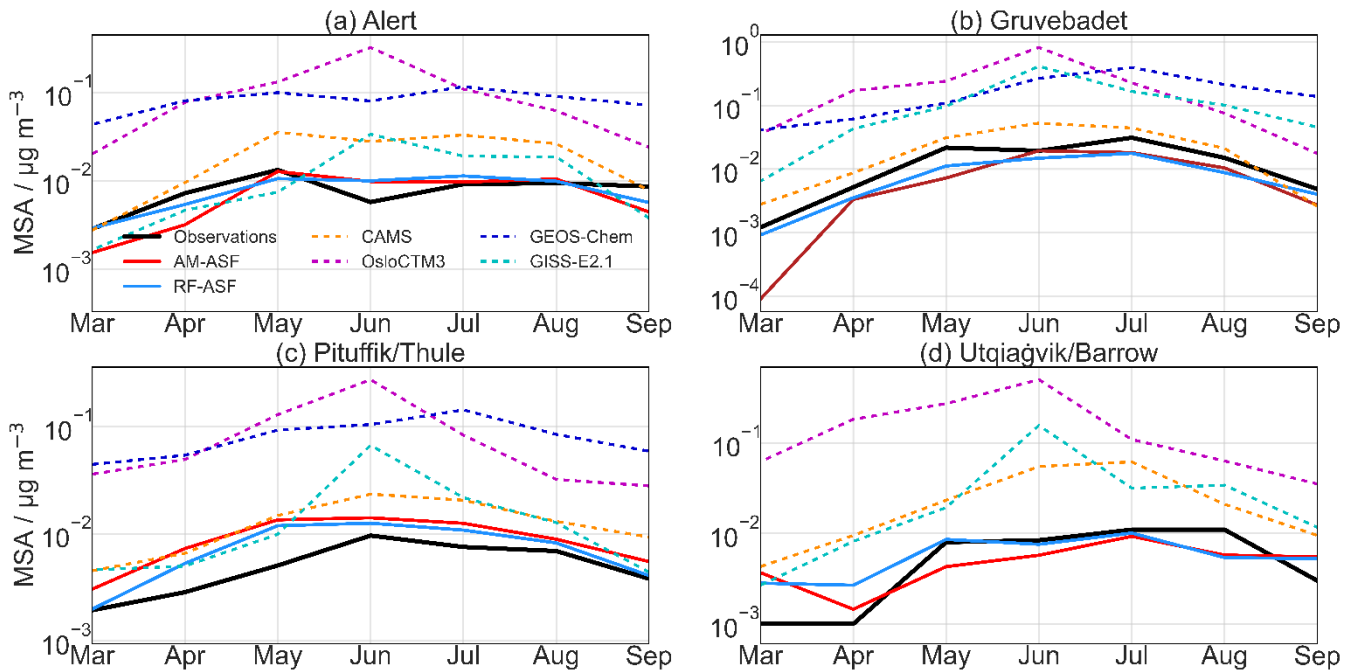



Figure S13: Comparison of seasonal cycles for observations, ASF data-driven model, and numerical models. Monthly medians for observations (solid black), data-driven model (AM-ASF in solid red and RF-AS in solid light blue), CAMS (dashed orange), GEOS-Chem (dashed blue), GISS-E2.1 (dashed cyan), and OsloCTM3 (dashed magenta) for (**a**) Alert, (**b**) Gruvebadet, (**c**) Pituffik/Thule, and (**d**) Utqiaġvik/Barrow. Only data for the tests were included in this analysis for a fair comparison, see Table 3 for dates. ASF refers to all data from all four stations and tested only on the specified station. The evaluation metrics for each data-driven and numerical model against in situ observations are given in Fig. 10.

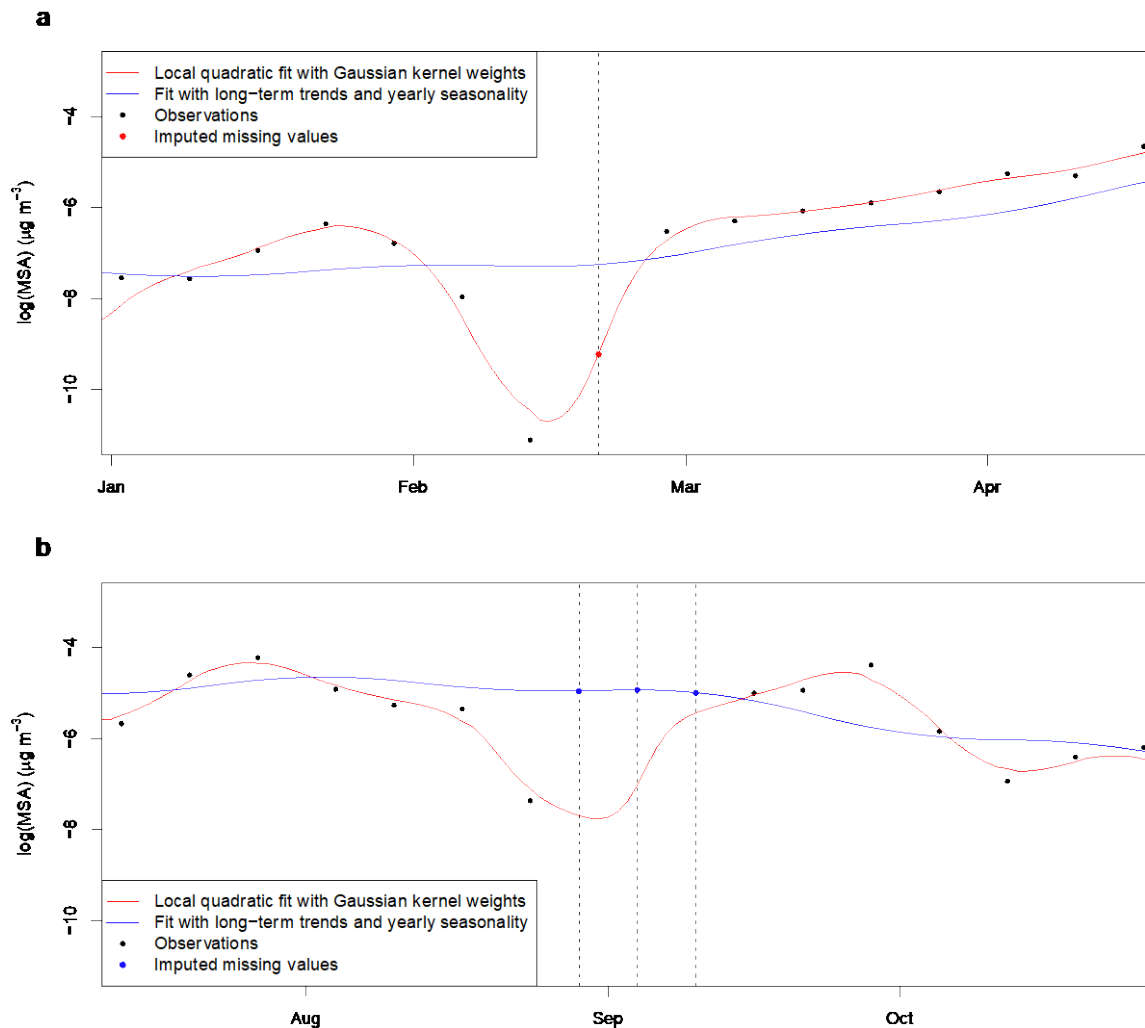



Figure S14: Illustration of missing value imputation schemes for Alert MSA$_p$ in situ measurements. Local quadratic fit with Gaussian kernel weighting (red curve), the fitted model with long-term time trends and yearly seasonality (blue curve), overlaid on top of the log-transformed MSA$_p$ observations (black dots), and the imputations (dots colored according to which method was applied). The vertical dashed lines over the shaded background indicate where missing values are to be imputed at Alert's weekly nominal temporal resolution. (**a**) Example of a short gap:14 days between available observations, starting on 2017-02-13, where a single value is imputed with the local quadratic fit. (**b**) Example of a long gap: 23 days between available observations, starting on 2014-08-24, where three values are imputed with long-term time trends and yearly seasonality.

4. References

345 Andronache, C.: Estimated variability of below-cloud aerosol removal by rainfall for observed aerosol size distributions, Atmos. Chem. Phys., 3, 131–143, https://doi.org/10.5194/acp-3-131-2003, 2003.

Barnes, I., Hjorth, J., and Mihalopoulos, N.: Dimethyl Sulfide and Dimethyl Sulfoxide and Their Oxidation in the Atmosphere, Chem. Rev., 106, 940–975, https://doi.org/10.1021/cr020529+, 2006.

Barrie, L. A., den Hartog, G., Bottenheim, J. W., and Landsberger, S.: Anthropogenic aerosols and gases in the lower 350 troposphere at Alert Canada in April 1986, J. Atmos. Chem., 9, 101–127, https://doi.org/10.1007/BF00052827, 1989.

Becagli, S., Ghedini, C., Peeters, S., Rottiers, A., Traversi, R., Udisti, R., Chiari, M., Jalba, A., Despiau, S., Dayan, U., and Temara, A.: MBAS (Methylene Blue Active Substances) and LAS (Linear Alkylbenzene Sulphonates) in Mediterranean coastal aerosols: Sources and transport processes, Atmos. Environ., 45, 6788–6801, https://doi.org/10.1016/j.atmosenv.2011.04.041, 2011.

355 Becagli, S., Lazzara, L., Marchese, C., Dayan, U., Ascanius, S. E., Cacciani, M., Caiazzo, L., Di Biagio, C., Di Iorio, T., di Sarra, A., Eriksen, P., Fani, F., Giardi, F., Meloni, D., Muscari, G., Pace, G., Severi, M., Traversi, R., and Udisti, R.: Relationships linking primary production, sea ice melting, and biogenic aerosol in the Arctic, Atmos. Environ., 136, 1–15, https://doi.org/10.1016/j.atmosenv.2016.04.002, 2016.

Becagli, S., Amore, A., Caiazzo, L., Iorio, T. D., Sarra, A. di, Lazzara, L., Marchese, C., Meloni, D., Mori, G., Muscari, G., 360 Nuccio, C., Pace, G., Severi, M., and Traversi, R.: Biogenic Aerosol in the Arctic from Eight Years of MSA Data from Ny Ålesund (Svalbard Islands) and Thule (Greenland), Atmosphere, 10, 349, 2019.

Boisvert, L. N., Webster, M. A., Parker, C. L., and Forbes, R. M.: Rainy Days in the Arctic, Journal of Climate, 36, 6855–6878, https://doi.org/10.1175/JCLI-D-22-0428.1, 2023.

Chen, Q., Sherwen, T., Evans, M., and Alexander, B.: DMS oxidation and sulfur aerosol formation in the marine troposphere: 365 a focus on reactive halogen and multiphase chemistry, Atmos. Chem. Phys., 18, 13617–13637, https://doi.org/10.5194/acp-18-13617-2018, 2018.

Delene, D. J. and Ogren, J. A.: Variability of Aerosol Optical Properties at Four North American Surface Monitoring Sites, Journal of the Atmospheric Sciences, 59, 1135–1150, https://doi.org/10.1175/1520-0469(2002)059<1135:VOAOPA>2.0.CO;2, 2002.

370 Isokääntä, S., Kim, P., Mikkonen, S., Kühn, T., Kokkola, H., Yli-Juuti, T., Heikkinen, L., Luoma, K., Petäjä, T., Kipling, Z., Partridge, D., and Virtanen, A.: The effect of clouds and precipitation on the aerosol concentrations and composition in a boreal forest environment, Atmospheric Chemistry and Physics, 22, 11823–11843, https://doi.org/10.5194/acp-22-11823-2022, 2022.

Kolesar, K. R., Cellini, J., Peterson, P. K., Jefferson, A., Tuch, T., Birmili, W., Wiedensohler, A., and Pratt, K. A.: Effect of 375 Prudhoe Bay emissions on atmospheric aerosol growth events observed in Utqiaġvik (Barrow), Alaska, Atmos. Environ., 152, 146–155, https://doi.org/10.1016/j.atmosenv.2016.12.019, 2017.

Li, S.-M. and Barrie, L. A.: Biogenic sulfur aerosol in the Arctic troposphere: 1. Contributions to total sulfate, J. Geophys. Res. Atmos., 98, 20613–20622, 1993.

Mansour, K., Decesari, S., Ceburnis, D., Ovadnevaite, J., Russell, L. M., Paglione, M., Poulain, L., Huang, S., O'Dowd, C., and Rinaldi, M.: IPB-MSA&SO₄: a daily 0.25\degree resolution dataset of in situ-produced biogenic methanesulfonic acid and sulfate over the North Atlantic during 1998–2022 based on machine learning, Earth System Science Data, 16, 2717–2740, https://doi.org/10.5194/essd-16-2717-2024, 2024.

Maturilli, M., Herber, A., and König-Langlo, G.: Climatology and time series of surface meteorology in Ny-Ålesund, Svalbard, Earth Syst. Sci. Data, 5, 155–163, https://doi.org/10.5194/essd-5-155-2013, 2013.

Moffett, C. E., Barrett, T. E., Liu, J., Gunsch, M. J., Upchurch, L. M., Quinn, P. K., Pratt, K. A., and Sheesley, R. J.: Long-Term Trends for Marine Sulfur Aerosol in the Alaskan Arctic and Relationships With Temperature, J. Geophys. Res. Atmos., 125, e2020JD033225, https://doi.org/10.1029/2020JD033225, 2020.

Muscari, G., Biagio, C. D., Sarra, A. di, Cacciani, M., Ascanius, S. E., Bertagnolio, P. P., Cesaroni, C., Zafra, R. L. de, Eriksen, P., Fiocco, G., Fiorucci, I., and Fuà, D.: Observations of surface radiation and stratospheric processes at Thule Air Base, Greenland, during the IPY, Ann Geophys, 57, SS0323, https://doi.org/10.4401/ag-6382, 2014.

Pernov, J. B., Harris, E., Volpi, M., Baumgartner, T., Hohermuth, B., Henne, S., Aeberhard, W. H., Becagli, S., Quinn, P. K., Traversi, R., Upchurch, L. M., and Schmale, J.: Pan-Arctic methanesulfonic acid aerosol: source regions, atmospheric drivers, and future projections, npj Clim Atmos Sci, 7, 1–18, https://doi.org/10.1038/s41612-024-00712-3, 2024.

Quinn, P. K., Bates, T. S., Miller, T. L., Coffman, D. J., Johnson, J. E., Harris, J. M., Ogren, J. A., Forbes, G., Anderson, T. L., Covert, D. S., and Rood, M. J.: Surface submicron aerosol chemical composition: What fraction is not sulfate?, J. Geophys. Res. Atmos., 105, 6785–6805, https://doi.org/10.1029/1999JD901034, 2000.

Quinn, P. K., Miller, T. L., Bates, T. S., Ogren, J. A., Andrews, E., and Shaw, G. E.: A 3-year record of simultaneously measured aerosol chemical and optical properties at Barrow, Alaska, J. Geophys. Res. Atmos., 107, AAC 8-1, https://doi.org/10.1029/2001JD001248, 2002.

Quinn, P. K., Bates, T. S., Schulz, K., and Shaw, G. E.: Decadal trends in aerosol chemical composition at Barrow, Alaska: 1976–2008, Atmos. Chem. Phys., 9, 8883–8888, https://doi.org/10.5194/acp-9-8883-2009, 2009.

Sharma, S., Chan, E., Ishizawa, M., Toom-Sauntry, D., Gong, S. L., Li, S. M., Tarasick, D. W., Leaitch, W. R., Norman, A., Quinn, P. K., Bates, T. S., Levasseur, M., Barrie, L. A., and Maenhaut, W.: Influence of transport and ocean ice extent on biogenic aerosol sulfur in the Arctic atmosphere, J. Geophys. Res. Atmos., 117, https://doi.org/10.1029/2011JD017074, 2012.

Sharma, S., Barrie, L. A., Magnusson, E., Brattström, G., Leaitch, W. R., Steffen, A., and Landsberger, S.: A Factor and Trends Analysis of Multidecadal Lower Tropospheric Observations of Arctic Aerosol Composition, Black Carbon, Ozone, and Mercury at Alert, Canada, J. Geophys. Res. Atmos., 124, 14133–14161, https://doi.org/10.1029/2019JD030844, 2019.

Shen, J., Scholz, W., He, X.-C., Zhou, P., Marie, G., Wang, M., Marten, R., Surdu, M., Rörup, B., Baalbaki, R., Amorim, A., Ataei, F., Bell, D. M., Bertozzi, B., Brasseur, Z., Caudillo, L., Chen, D., Chu, B., Dada, L., Duplissy, J., Finkenzeller, H., Granzin, M., Guida, R., Heinritzi, M., Hofbauer, V., Iyer, S., Kemppainen, D., Kong, W., Krechmer, J. E., Kürten, A., Lamkaddam, H., Lee, C. P., Lopez, B., Mahfouz, N. G. A., Manninen, H. E., Massabò, D., Mauldin, R. L., Mentler, B., Müller, T., Pfeifer, J., Philippov, M., Piedehierro, A. A., Roldin, P., Schobesberger, S., Simon, M., Stolzenburg, D., Tham, Y. J., Tomé, A., Umo, N. S., Wang, D., Wang, Y., Weber, S. K., Welti, A., Wollesen de Jonge, R., Wu, Y., Zauner-Wieczorek, M., Zust, F., Baltensperger, U., Curtius, J., Flagan, R. C., Hansel, A., Möhler, O., Petäjä, T., Volkamer, R., Kulmala, M., Lehtipalo, K., Rissanen, M., Kirkby, J., El-Haddad, I., Bianchi, F., Sipilä, M., Donahue, N. M., and Worsnop, D. R.: High Gas-Phase Methanesulfonic Acid Production in the OH-Initiated Oxidation of Dimethyl Sulfide at Low Temperatures, Environ. Sci. Technol., 56, 13931–13944, https://doi.org/10.1021/acs.est.2c05154, 2022.

Song, C., Becagli, S., Beddows, D. C. S., Brean, J., Browse, J., Dai, Q., Dall'Osto, M., Ferracci, V., Harrison, R. M., Harris, N., Li, W., Jones, A. E., Kirchgäßner, A., Kramawijaya, A. G., Kurganskiy, A., Lupi, A., Mazzola, M., Severi, M., Traversi, R., and Shi, Z.: Understanding Sources and Drivers of Size-Resolved Aerosol in the High Arctic Islands of Svalbard Using a Receptor Model Coupled with Machine Learning, Environ. Sci. Technol., 56, 11189–11198, https://doi.org/10.1021/acs.est.1c07796, 2022.

Tunved, P., Ström, J., and Krejci, R.: Arctic aerosol life cycle: linking aerosol size distributions observed between 2000 and 2010 with air mass transport and precipitation at Zeppelin station, Ny-Ålesund, Svalbard, Atmos. Chem. Phys., 13, 3643–3660, https://doi.org/10.5194/acp-13-3643-2013, 2013.

Zanatta, M., Mertes, S., Jourdan, O., Dupuy, R., Järvinen, E., Schnaiter, M., Eppers, O., Schneider, J., Jurányi, Z., and Herber, A.: Airborne investigation of black carbon interaction with low-level, persistent, mixed-phase clouds in the Arctic summer, Atmospheric Chemistry and Physics, 23, 7955–7973, https://doi.org/10.5194/acp-23-7955-2023, 2023.