



Pan-Arctic methanesulfonic acid aerosol: source regions, atmospheric drivers, and future projections



Jakob Boyd Pernov ¹✉, Eliza Harris ², Michele Volpi ², Tamara Baumgartner¹, Benjamin Hohermuth³, Stephan Henne⁴, William H. Aeberhard², Silvia Becagli^{5,6}, Patricia K. Quinn⁷, Rita Traversi^{5,6}, Lucia M. Upchurch ^{7,8} & Julia Schmale¹✉

Natural aerosols are an important, yet understudied, part of the Arctic climate system. Natural marine biogenic aerosol components (e.g., methanesulfonic acid, MSA) are becoming increasingly important due to changing environmental conditions. In this study, we combine in situ aerosol observations with atmospheric transport modeling and meteorological reanalysis data in a data-driven framework with the aim to (1) identify the seasonal cycles and source regions of MSA, (2) elucidate the relationships between MSA and atmospheric variables, and (3) project the response of MSA based on trends extrapolated from reanalysis variables and determine which variables are contributing to these projected changes. We have identified the main source areas of MSA to be the Atlantic and Pacific sectors of the Arctic. Using gradient-boosted trees, we were able to explain 84% of the variance and find that the most important variables for MSA are indirectly related to either the gas- or aqueous-phase oxidation of dimethyl sulfide (DMS): shortwave and longwave downwelling radiation, temperature, and low cloud cover. We project MSA to undergo a seasonal shift, with non-monotonic decreases in April/May and increases in June–September, over the next 50 years. Different variables in different months are driving these changes, highlighting the complexity of influences on this natural aerosol component. Although the response of MSA due to changing oceanic variables (sea surface temperature, DMS emissions, and sea ice) and precipitation remains to be seen, here we are able to show that MSA will likely undergo a seasonal shift solely due to changes in atmospheric variables.

The Arctic is an especially climate-sensitive region of the Earth system, which is evident from the enhanced rate of warming compared to the global average¹. In addition to temperature changes, the Arctic has also experienced substantial declines in e.g., sea ice extent, thickness, and concentration² as well as thawing permafrost³, Greenlandic ice sheet melt⁴, increasing riverine fresh-water and sediment input into the ocean⁵, and increasing marine biological activity⁶. While anthropogenic CO₂ emissions are still increasing⁷, short-lived climate forcers (e.g., black carbon and sulfate aerosols) have been decreasing or plateauing in recent decades^{8,9}. These changes suggest that the anthropogenic influence on Arctic aerosols is decreasing in importance while natural influences on processes affecting

aerosols are relatively increasing⁹. These changes motivate the need to explore the natural components of the Arctic system, especially natural aerosols, which are an important factor in determining the Arctic radiative budget due to their direct and indirect effects^{10,11}. Aerosols are responsible for a large proportion of climate modeling uncertainty^{12–14} and including natural aerosol processes in climate models can help reduce their uncertainty¹⁵. Aerosols are expected to respond to the changes in the Arctic environment but determining the direction and magnitude of this response is complex^{16–18}. Therefore, investigations on the role and drivers of natural aerosols in the climate system are crucial for evaluating the effects of climate change in this delicate environment⁹.

¹Extreme Environments Research Laboratory, École Polytechnique Fédérale de Lausanne, Sion, Switzerland. ²Swiss Data Science Center, ETH Zurich and École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. ³Schroders Capital ILS, Zürich, Switzerland. ⁴Empa, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland. ⁵Department of Chemistry, University of Florence, Florence, Italy. ⁶Institute of Polar Sciences, National Research Council, Venice, Italy. ⁷Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration, Seattle, WA, USA. ⁸Cooperative Institute for Climate, Ocean, and Ecosystem Studies, University of Washington, Seattle, WA, USA. ✉e-mail: jakob.pernov@epfl.ch; julia.schmale@epfl.ch

Methanesulfonic acid (MSA) in the Arctic is derived mainly from the oxidation of naturally emitted, marine dimethyl sulfide (DMS)¹⁹. While the open ocean is the major source of DMS emissions, other sources can contribute as well such as melt ponds, biomass burning, coastal tundra, and lakes^{20–22}. DMS is produced from the enzymatic cleavage of dimethylsulfoniopropionate^{23,24} and represents the largest source of marine biogenic sulfur in the atmosphere^{25,26}. DMS mixing ratios follow a seasonal cycle with near-zero values during polar night (October to February) and elevated and highly variable values during polar day (March to September), typically ranging between 30 and 180 parts per trillion by volume (pptv) for monthly averages^{21,27–31}. Once ventilated into the atmosphere, DMS is oxidized by OH, O₃, NO₃, and halogens, a process that proceeds via either the addition or abstraction pathways, which are temperature dependent¹⁹. In the atmosphere, DMS has a lifetime of approx. 1–2 days²¹, which is dependent on latitude in the Arctic³². Gas-phase DMS oxidation ultimately yields MSA or SO₂ (which is further oxidized to sulfuric acid), with the addition pathway resulting in higher yields of MSA³³. DMS can also be oxidized in the aqueous phase (through processing in clouds or on deliquesced particles³⁴) mainly via O₃ and OH, which has been demonstrated through modeling and measurements to be an important formation mechanism^{35–40}. After dissolving into a droplet, DMS is oxidized into MSA through multi-phase chemistry and during droplet evaporation, MSA is released into the gas-phase and can further influence secondary aerosol production^{37,39,40}. Particulate MSA also follows a seasonal cycle with near-zero values during polar night and elevated values during polar day^{41–44}. Peak MSA concentrations are experienced during May (~10–40 ng m⁻³) and typically remain elevated until the end of summer (the seasonal cycle of MSA will be discussed in Section “Seasonal cycle and source regions of Pan-Arctic MSA”). Trends in MSA have shown high spatial and temporal variability in the Arctic, for instance, MSA has increased at Utqiāġvik/Barrow during July–September by 2.5% yr⁻¹ over 1998–2017⁴⁵, at Kevo, Finland, MSA has increased during June/July by 0.68% yr⁻¹ over 1964–2010⁴⁶, and at Alert, CA, MSA has shown decreasing trends during the cold and warm periods over 1980–2000 but increased during the warm period over 2000–2009⁸. The variable trends in MSA are likely related to interannual variability in source strength and transport patterns over the different considered periods^{8,45,47,48}. MSA has a lifetime of several days depending on the environmental conditions⁴⁹. MSA has been observed to reside mainly in the accumulation mode (aerosols with a diameter > 100 nm)^{50–52} although recent observations have indicated MSA to be in the Aitken mode (~25 < diameter < 100 nm) during a formation event⁵³. MSA is removed from the atmosphere via dry and wet deposition (with wet deposition being the most important³⁷) or further oxidized into sulfate in the aqueous phase³⁹.

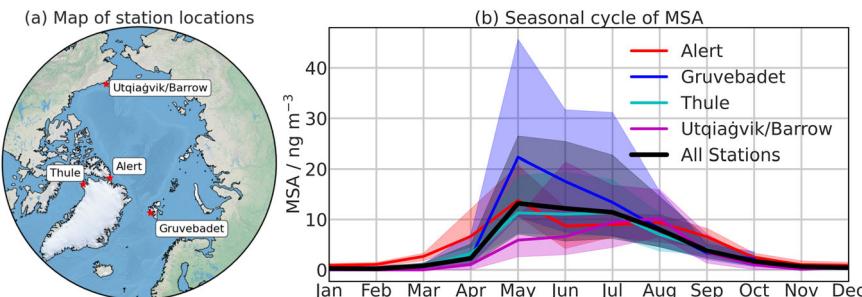
MSA is an important constituent in the Arctic atmospheric system, as it has mainly been observed to aid in the condensational growth of aerosols, but has also been modeled to participate in the nucleation of new aerosol particles^{54–56}. While its role in nucleation has yet to be directly observed in the field^{57,58}, it has been observed in chamber studies⁵⁹. More importantly, while only representing a fraction of the total aerosol mass, MSA is critical for the growth of particles to cloud condensation nuclei (CCN) sizes^{31,32,60} thereby affecting cloud microphysical properties such as cloud lifetime,

albedo, and precipitation efficiency^{61–64}. MSA has been shown to selectively condense onto alkaline particles and not acidic or hydrophobic particles in polar regions^{65,66}. The summertime Arctic is characterized by low particle concentrations, and thus an aerosol-sensitive CCN regime^{67–69}, therefore, variations in the number concentration of CCN-active aerosols can have a large effect on the radiative balance of clouds¹². Understanding the sources and atmospheric drivers of MSA is therefore vital for accurate modeling of the Arctic climate system, given that aerosol–cloud interactions are one of the largest sources of uncertainty in global climate modeling¹⁵.

Machine learning (ML) algorithms and models are powerful tools for understanding environmental phenomena since they can account for non-linear, complex relationships from large datasets that simpler analysis methods wouldn't be able to accommodate. ML has several advantages over numerical models (NMs), namely computational efficiency, flexibility, and the ability to handle complex non-linear tasks⁷⁰. ML models are data-driven, meaning they rely solely on the explanatory data and not on pre-defined physical equations, which allows them to efficiently capture complex non-linear relationships for many different applications. NMs rely on physical equations, emission inventories, and assumptions that require accurate parameterizations and estimates, thus limiting their flexibility. As NMs increase in their complexity, the computational costs and runtime increase disproportionately therefore capturing complex atmospheric phenomena can become prohibitively computationally expensive. ML is not without its disadvantages, including the need for large amounts of curated training data and the difficulty in interpreting complex models. Efforts are being made to create hybrid ML-NM models that maximize the advantages of both approaches while minimizing their disadvantages. ML has been demonstrated to be effective in the analysis of environmental processes including predicting sea surface DMS concentrations^{71,72}, aerosol–cloud interactions⁷³, CCN concentrations^{70,74}, and global particle number concentrations⁷⁵. In the Arctic, k-means clustering has been utilized at several sites to understand the dynamics of aerosol particle number size distributions^{76–80} and their links to environmental conditions (e.g., open water extent, transport patterns) as well as their sources and climate-relevant properties. Song et al.⁸¹ used a random forest regressor to understand the drivers of different aerosol types on Svalbard and found that solar radiation, surface pressure, and temperature were drivers of biogenic-type aerosols. Hu et al.⁸² used principal component analysis coupled with a generalized additive model to show the marine and terrestrial influences on isoprene and monoterpene-derived secondary organic aerosols in the Arctic. Overall, this highlights the applicability of ML models to understand atmospheric phenomena and their advantages over NMs.

In this study, we leverage in situ observations of particulate MSA from several stations geographically dispersed around the Arctic (Fig. 1a), along with FLEXible PARTicle dispersion model (FLEXPART⁸³) backward dispersion simulations and ERA5⁸⁴ reanalysis meteorological datasets. Our objectives are to (1) identify the Pan-Arctic seasonal cycle and source regions of particulate MSA, (2) elucidate the relationships between MSA and atmospheric variables, and (3) project the future response of MSA based on trends of environmental variables in ERA5. First, we explore the Pan-Arctic seasonal cycles and source regions of MSA. We then use the average trajectory-weighted source regions of MSA and the derived average

Fig. 1 | Locations and MSA seasonal cycle for each station. **a** Map of the four Arctic stations used in this study. Stations are indicated with a red star. The map background is from Natural Earth. **b** The seasonal cycle of MSA at Alert (red), Gruvebadet (blue), Thule (cyan), Utqiāġvik/Barrow (magenta), and all stations combined (Pan-Arctic) (black). The thick lines represent the median and the shading represents the interquartile range for each month over the period 2010–2017 for Alert, Gruvebadet, and Thule. For Utqiāġvik/Barrow, the period is 2008–2014.



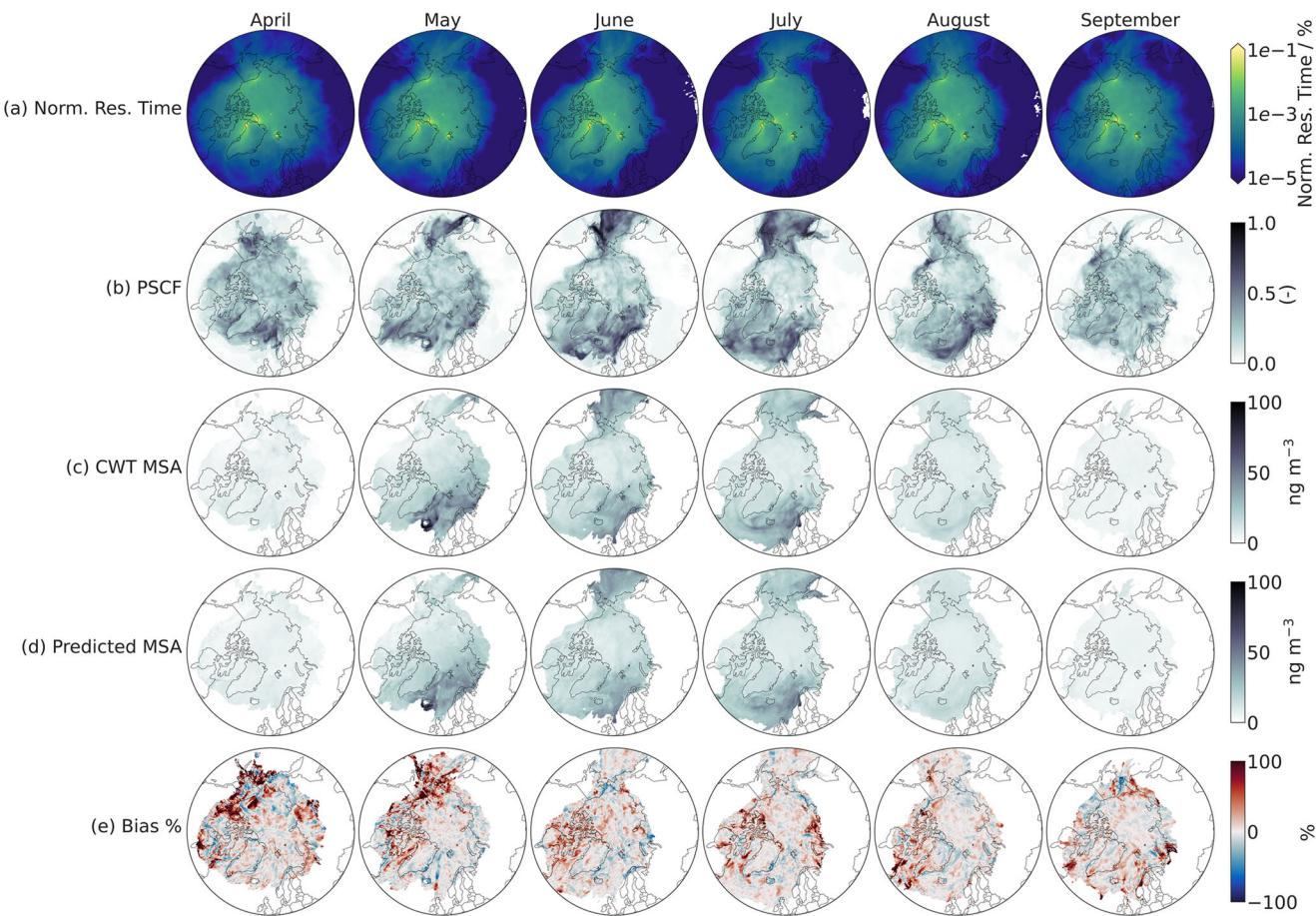


Fig. 2 | Monthly Pan-Arctic transport pathways, MSA source regions, and model prediction/bias. **a** Transport pathways for all stations expressed as the normalized residence time. **b** Potential source contribution function for MSA. **c** Concentration

weighted trajectory of MSA using in situ observations. **d** Model prediction of the CWT MSA. **e** Model bias (difference between modeled and observed values) expressed as a percentage. Each row uses the same color scale on the right.

trajectory-weighted meteorological conditions from ERA5 as input into an ML model and use the SHapley Additive exPlanations (SHAP⁸⁵) approach for model explainability (see Methods for further details about trajectory-weighted source regions and meteorological conditions as well as SHAP methodology). Following this, we project the future MSA concentrations using the same trained ML model and extrapolated trends in the ERA5 data. We utilize a leave-one-out approach to explore which trends in ERA5 variables are contributing to future changes in MSA. Finally, we discuss the limitations and implications of our results.

Results and discussion

Seasonal cycle and source regions of Pan-Arctic MSA

The seasonal cycle of particulate MSA at each of the four Arctic stations (locations are displayed in Fig. 1a) and for all stations combined, is displayed in Fig. 1b. For a description of each station see the Methods section. For all stations, MSA increases beginning in April and decreases in September, except for Alert which experiences increases beginning in March and decreases beginning in October. This period (April–September) corresponds to polar day, receding sea ice, increase in atmospheric oxidants as well as phytoplankton blooms^{2,28,86,87}. During October–March, MSA is at near-zero concentrations, mainly due to polar night in the Arctic when little biological activity and a minimum of photochemical activity occur^{26,88}. Combining these geographically dispersed stations with their varying seasonal cycles allows us to understand the Pan-Arctic seasonal cycle of MSA as represented by the median of data from all stations (Fig. 1b). Pan-Arctic MSA shows a sharp increase from April to May, stays elevated from May to July (11–13 ng m⁻³), and gradually decreases until September. This period of elevated MSA during polar day corresponds to enhanced solar radiation

and marine biological activity in the Arctic^{25,26,88}. The seasonal cycles presented in this study agree with previous studies which analyzed different periods^{8,42,47,48}. The Pan-Arctic seasonal cycle shows variations from April through September; therefore, we limit our analysis to these months.

The Pan-Arctic seasonal cycle can be attributed to the different geographical locations of each station, as well as varying transport patterns, in relation to biologically active regions of open water. To help contextualize the Pan-Arctic MSA seasonal cycle, the transport climatology was explored for all stations combined for April through September (Fig. 2a). A noticeable feature is the greater geographical extent of air masses during April and September compared to May through August, which is likely a reflection of the extent of the polar dome^{89,90}, which allows for efficient transport from lower latitudes. Whereas during May through August, air masses are mainly confined to marine areas in the Arctic⁹¹. It should be noted that the polar dome begins to recede during May, therefore May is often classified as a transition month. Overall, during the peak summer months, air masses mainly arrive from the Greenland/Barents Seas, Baffin Bay/Canadian Archipelago, the Bering Strait/Sea, and the central Arctic Ocean (Fig. 2a).

To elucidate the source regions of MSA between April and September, we applied a Potential Source Contribution Function (PSCF) and a Concentration Weighted Trajectory (CWT) analysis using in situ observations of MSA and FLEXPART (see Methods Section for details). Figure 2b and c displays the Pan-Arctic PSCF and CWT, respectively. The PSCF shows the *probability* for potential source regions, however, it doesn't show the *contribution* to observed concentrations that can be expected for a given air mass origin, therefore, we complemented the PSCF with a CWT analysis (Fig. 2c). The PSCF for April shows two regions with a high probability of being source regions, (i) the area extending from the Greenland Sea to lower

Fig. 3 | Evaluation metrics for the model performance. **a** Train and test sets metrics and **b** monthly metrics for the test sets. The color in each cell indicates the highest value for that row. The upper number in each panel represents the mean and the numbers in brackets represent the standard deviation.

	(a)		(b)						
	Train	Test	Slope	0.51 (0.01)	0.70 (0.03)	0.74 (0.06)	0.83 (0.04)	0.75 (0.02)	0.57 (0.07)
			R^2	0.33 (0.04)	0.74 (0.04)	0.76 (0.02)	0.83 (0.01)	0.53 (0.02)	0.39 (0.09)
			RMSE %	28.35 (1.89)	29.91 (2.63)	26.18 (1.76)	24.24 (1.48)	19.93 (0.56)	17.24 (1.87)
Slope	0.91 (0.00)	0.82 (0.02)							
R^2	0.94 (0.00)	0.84 (0.02)							
RMSE %	17.52 (0.33)	28.80 (1.85)							

Baffin Bay, and (ii) continental Alaska. The first area is likely related to the biological activity there⁶, while the second likely represents transport pathways from the northeastern subarctic Pacific Ocean^{45,47,71}, as observed from May to August (Fig. 2b). During September, the potential source regions of MSA are delocalized around the Arctic. Similar source regions of MSA are observed from May to August (Fig. 2b), which include the Greenland, Barents, Labrador, and Bering Seas. Overall, the marine areas of the Atlantic and Pacific sectors of the Arctic exhibit the highest probability of being an MSA source region during May through July. The CWT analysis supports the PSCF as it shows that the most probable source regions also exhibit the highest contributions to MSA. However, differences exist, for example, during May the PSCF shows that both the Atlantic and Pacific sectors appear to have similar probabilities of being source regions while the CWT shows the Atlantic sector has a higher relative contribution, highlighting the complementary nature of these two approaches. The most probable source regions of the Atlantic and Pacific sectors also make the largest contributions to MSA during the summer, which is likely due to a combination of relatively stronger biological activity and air mass transport from these regions.

Elucidating the relationships between MSA and atmospheric variables

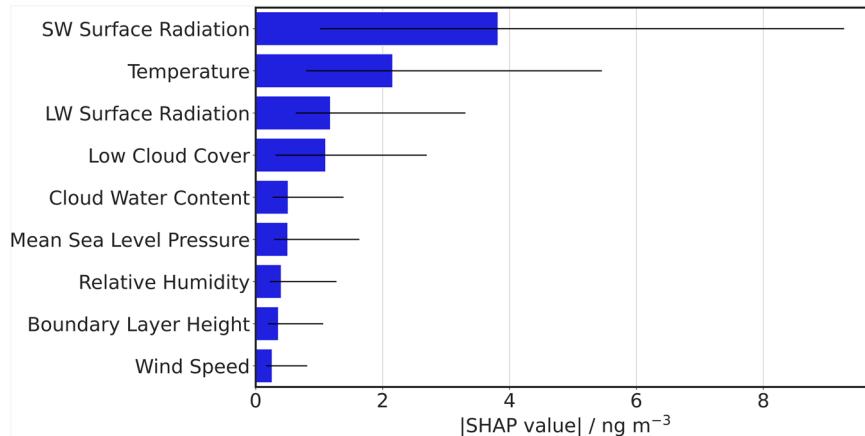
A transport climatology and source region analysis are useful, but they cannot elucidate the relationships between atmospheric variables and observed MSA concentrations. To explore these links, we developed an ML approach to model the trajectory-weighted concentration of MSA (or concentration-weighted trajectory, CWT), we derived the trajectory-weighted meteorological conditions (or meteorological-weighted trajectory, MWT) using FLEXPART and ERA5 to use as input for the machine learning model (See Supplementary Fig. 1 for maps of the MWTs). MWTs consider the average meteorological conditions for the actual location of air masses (using the gridded ERA5 data) and are analogous to our target variable (CWT), thus providing appropriate explanatory variables for our specific ML task and the target variable. Appropriate explanatory variables are important as the prediction efficacy of ML models largely depends on the appropriateness of the explanatory data provided (see Methods for a full description of MWTs and their calculation). The XGBoost library implementing gradient-boosted decision trees was selected among other competing models due to computational efficiency, performance, and high generalization accuracy on many data modeling tasks⁹². The relevant variables were selected using recursive feature elimination (See Supplementary Fig. 3). The model was optimized using the Optuna⁹³ library, which implements a Bayesian approach to hyperparameter tuning, and the SHAP approach was utilized to explain the model. See the Methods Section for further details about the steps in this methodology.

Before interpreting the model results, we evaluate its performance using spatially stratified groups for the training and test sets for all months combined (See Supplementary Fig. 4 for a schematic of the spatially stratified groups) as well as evaluation of the test set on a monthly basis using the slope, R^2 , and root mean squared error expressed as a percentage of the mean (RMSE %). Figure 3a and b display the results of this model evaluation. The

training and test set evaluation shows overall good performance of our model, with a test set mean (\pm SD) slope and R^2 of 0.82 ± 0.02 and 0.84 ± 0.02 , respectively, compared to the training set slope and R^2 of 0.91 ± 0.003 and 0.94 ± 0.002 , respectively (Fig. 3a). These metrics indicate that our model has good performance (high R^2 /slope and low RMSE % values) and is not strongly overfitting the data (which would be characterized by very high training set metrics and a large difference between training and test set metrics with the latter usually being very poor). When evaluating our model on a monthly basis, we can see that May through August are better reproduced by the model than April and September. Supplementary Fig. 5 shows further investigation of the model performance with boxplots of the percent bias for deciles of the CWT MSA for each month. In this study, bias is defined as the difference between modeled and observed values. A common theme for the distribution of the percent bias in each decile is apparent: at lower values of CWT MSA, the model overpredicts, while at higher values of CWT MSA, the model underpredicts, leading to an overall underestimation of MSA by our model (Fig. 3a). This pattern is also observed when analyzing the spatial bias of our model. Figure 2d and e shows the predicted MSA source regions and the percent bias from the CWT values, which displays this concentration regime bias. It should be noted that the observed CWT MSA is likely biased by the stations being mainly located in the European and North American Arctic (Fig. 1a) and three of the stations experience peak MSA during May and the other in August. There are no long-term observations of MSA in the Siberian Arctic⁸ and while air masses arriving at these stations cover the entire Arctic area, the Eastern Arctic is likely under-represented in the CWT MSA dataset. Our model overestimates MSA during April/May in the Canadian Archipelago and Alaska, which is likely a result of the low concentrations observed at Utqiagvik/Barrow during these months (Fig. 2d). These geographic and temporal variabilities should be kept in mind when interpreting the results. Overall, the spatial bias appears to be randomly distributed (with a slight influence from the geography and timing of MSA observations) and the model gives overall satisfactory results, indicating that the model is robust, accurate, and appropriate for the given modeling task.

To investigate the drivers of MSA using our model, we used the SHAP methodology which quantifies a variable's contribution to the model output for a given observation⁸⁵, whereby variables with positive (negative) SHAP values are actively increasing (decreasing) the model prediction (see Methods for further details). Figure 4 shows the overall variable importance while Supplementary Fig. 6 shows the variable importance on a monthly basis. The overall (or global) variable importance is represented by the median of the absolute SHAP values, with larger values having a greater impact on the model output. It should be noted that the SHAP values do not represent how well these explanatory variables explain the behavior of our target variable in the natural environment but how well these variables explain the behavior of our target variable in our model, therefore representing purely statistical relationships. Given these explanatory variables are environmental data, some correlations between them are unavoidable (Table S3), this can lead to difficulties when explaining the model using the SHAP methodology⁹⁴. Therefore, to avoid over-interpreting the model results due to correlated variables, we refrain from discussing the rankings of

Fig. 4 | The overall importance of each explanatory variable. The bars represent the median of the absolute SHAP value while the black lines represent the interquartile range.



individual explanatory variables but rather discuss which explanatory variables can represent certain processes in the model. From Fig. 4, we can see that the variables with the highest impact on the model are SW radiation, temperature, LW radiation, and low cloud cover, after which the variable importance is low with little variation for the remaining variables. These top four variables are all connected to either the gas- or aqueous-phase oxidation of DMS, which indicates that our model gives physically meaningful and interpretable results. It has been observed before that SW radiation is an important factor connected to MSA^{81,95,96}. SW radiation is necessary for the photochemical production of oxidants (OH, O₃, halogens)⁹⁷, both in the gas- and aqueous-phase, and for stimulating marine biological activity⁹⁸. SW radiation is therefore expected to have a positive impact on MSA concentrations, indeed MSA is mainly observed during the sunlit, summer months^{8,47}. Temperature affects DMS emission and oxidation to MSA through multiple different mechanisms. Importantly, temperature will affect the sea surface temperature and sea ice extent/concentration, with positive temperatures increasing the DMS emission and reducing sea ice cover^{26,99–101}. Temperature would also affect the oxidation of DMS after emission. The gas-phase DMS oxidation produces higher yields of MSA at colder temperatures^{19,102}. The thermodynamic phase of clouds will also be affected by temperature, with positive temperatures indicating the presence of mostly liquid-phase clouds and negative temperatures suggesting mixed- or ice-phase clouds, although this relationship is non-linear¹⁰³. The thermodynamic phase of clouds would affect the aqueous-phase oxidation of DMS^{19,36–38}, with liquid-containing clouds being more efficient than ice-containing clouds. Overall, it appears temperature will have a positive relationship with DMS emission and oxidation to MSA. LW radiation is used here as a proxy for the presence of liquid-containing clouds as liquid cloud water will enhance the downwelling LW radiation compared to clear sky conditions¹⁰⁴. While low-level liquid-containing clouds will affect the downwelling LW radiation, other factors including temperature and liquid water content will also have an effect, this should be kept in mind when interpreting the results. Low cloud cover indicates the fraction of low-altitude clouds present although it does not specify the phase (which has implications for the aqueous-phase oxidation of DMS and LW downwelling radiation). The presence of low-level clouds has been demonstrated to aid in the aqueous-phase processing of DMS^{39,50}, although low-level clouds can also reduce the amount of incoming SW radiation. LW radiation and low cloud cover therefore serve as proxies for aqueous-phase processes and are both expected to have an overall positive impact on MSA. These overall top four variables are also consistently amongst the most important variables on a monthly basis (Supplementary Fig. 6), therefore, we will focus our further discussion on these top four variables.

To investigate the relationship of SHAP values between different values of the explanatory variables (explanatory variables refer to variables input into the ML model. i.e., MWTs), we binned the explanatory variables into 10 equally spaced bins and calculated the median and interquartile range of

SHAP values for each bin for each month, as displayed in Fig. 5. Supplementary Fig. 7 shows the same figure but including all explanatory variables instead of only the top four. This allows us to discover how certain values of our explanatory variables affect the model output. We also show the geographic distribution of SHAP values for the top four variables and all variables in Fig. 5 and Supplementary Fig. 8, respectively.

For SW radiation, a complex pattern emerges for the relationship between the explanatory and the SHAP values. During April, July, and August, a mostly linear, monotonic relationship is displayed whilst during May and June a positive, non-monotonic relationship exists (Fig. 5a). During May and September, an entirely positive and negative relationship is observed, respectively. The overall relationship between explanatory and SHAP values of SW radiation supports the postulate that the photochemical oxidation of DMS or the production of key oxidants (OH, halogens, or O₃, within the aqueous phase or transferred to afterward) is a key atmospheric process leading to the formation of MSA, which is plausible based on theory. Geographically, SW radiation shows ubiquitously positive SHAP values over the Arctic region during May through July and mainly negative SHAP values during April, August, and September (Fig. 6a), corresponding to the seasonal cycle of SW radiation from ERA5 (Supplementary Fig. 2).

For temperature, during each month the overall relationship between the explanatory and SHAP temperature values is non-linear and there appears to be a threshold value of ~273 K, where afterwards SHAP values turn positive. Below this value, the average temperature SHAP values are all negative and mainly show a flat relationship with temperature values (Fig. 5b). This indicates that areas with above-freezing temperatures will make a positive contribution to the model output of MSA and areas with below-freezing temperatures will make a negative contribution. This is likely related to areas of open ocean with no sea ice, as melting sea ice regulates the overlying atmospheric temperature to around 273 K, thus areas with no sea ice can regularly experience above-freezing temperatures, whereas areas with melting sea ice in the summer experience temperatures around freezing. The absence of sea ice also allows for the emission of DMS into the atmosphere^{22,105}. Song et al.⁸¹ observed an opposite relationship between biogenic type aerosols (which were characterized by high fractions of MSA) and temperature, with a positive impact from temperature up until 273 K and a negative impact afterwards⁸¹, suggesting above-freezing temperatures would act to decrease MSA. This pattern was attributed to the role of disappearing sea ice with positive temperatures (which likely serves as a habitat for sympagic algae¹⁰⁶) rather than the temperature-dependent kinetics of DMS oxidation. It should be noted that Song et al.⁸¹ was only for a single year (2015) and used in situ measurements of temperature (representative of the conditions at the measurement location) as opposed to spatially-resolved temperature as in this study (representative of conditions in source regions and transport pathways), which would affect the interpretations. Studies from Utqiágik/Barrow show positive relationships

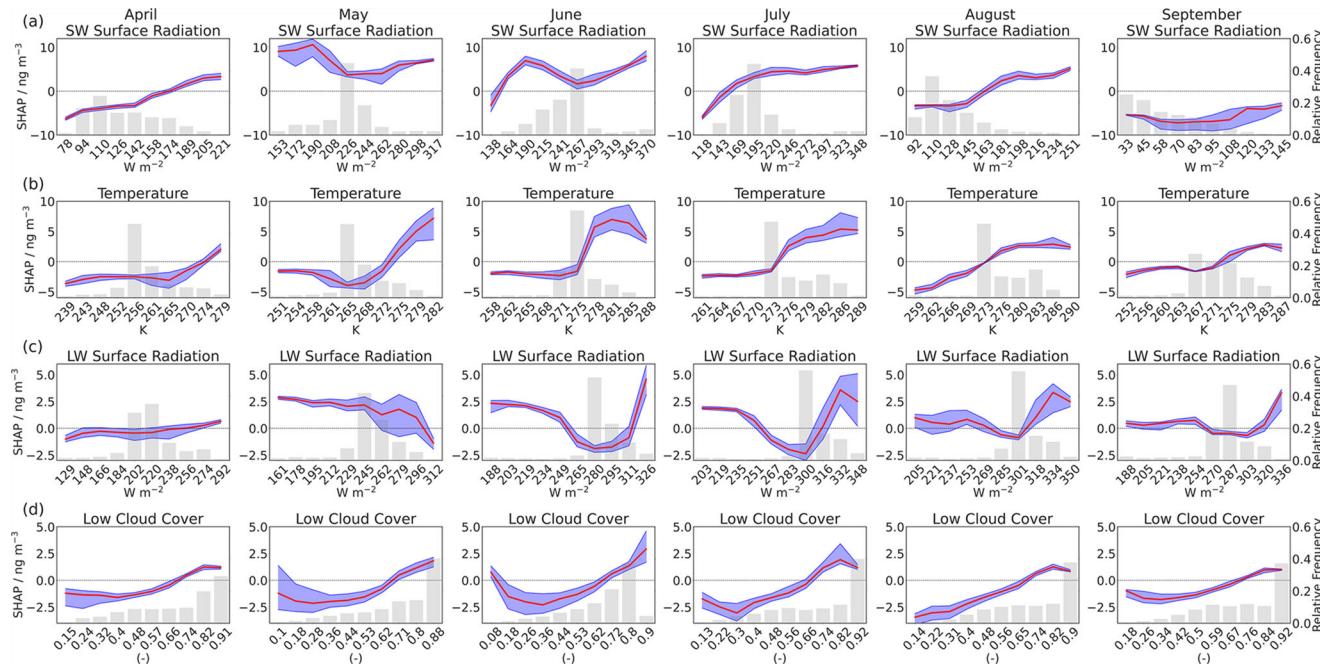


Fig. 5 | Relationships between SHAP and MWT values for the top four explanatory variables. The relationships between the explanatory and SHAP values for a SW surface radiation, b temperature, c LW surface radiation, and d low cloud cover for each month (columns). Ten equally spaced bins were calculated for each variable,

and the median (red line) and IQR (blue shading) of the SHAP values were computed for each bin. The value listed on the x-axis is the midpoint of each bin. The light gray bars represent the relative frequency of the MWT values for each bin (right y-axis).

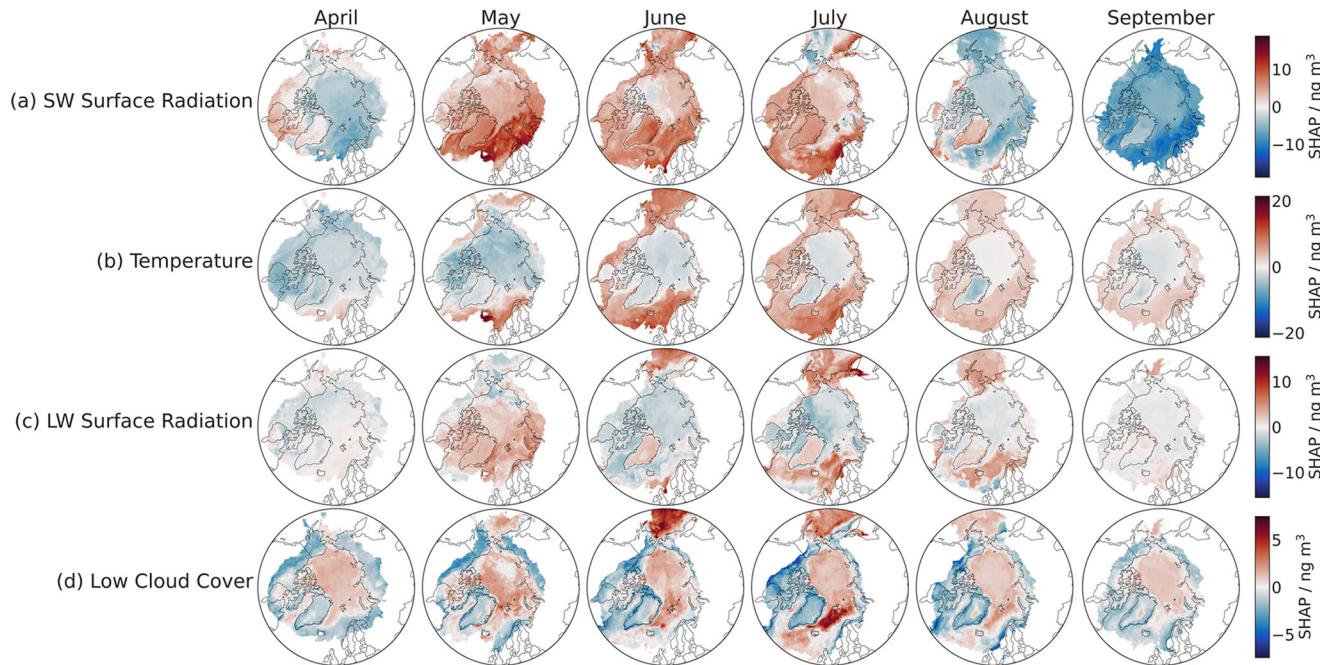


Fig. 6 | Spatial distribution of SHAP values for the top four explanatory variables. The spatial distribution of SHAP values for a SW surface radiation, b temperature, c LW surface radiation, and d low cloud cover for each month (columns). Each row uses the same color scale on the right.

between MSA and air temperature⁴⁵. Above-freezing temperatures will also affect the sea surface temperature, the stability of the boundary layer, and the presence of liquid-containing clouds. While the gas-phase oxidation of DMS into MSA has been shown to proceed at faster rates under colder temperatures^{19,102}, the lifetime of intermediates in the aqueous-phase oxidation of DMS into MSA is shorter at warmer temperatures¹⁰⁷ and the aqueous-phase oxidation requires the presence of liquid water. Conversely,

in areas with below-freezing temperatures, the presence of sea ice (physical barrier to the emission of DMS) as well as mixed-phase clouds (containing ice and supercooled water) is likely negatively affecting this relationship. The geographic distribution of temperature SHAP values closely follows the relationship between explanatory and SHAP values for temperature, with positive SHAP values in regions with positive temperatures (Fig. 6b), which are mainly located at more southerly latitudes. Explanatory and SHAP

values for temperature reach their maximum in the marine areas of the Atlantic and Pacific sectors during June and July (Fig. 6b and Supplementary Fig. 1), highlighting the importance of these two regions to MSA formation.

For LW radiation, a different pattern is observed. During April, the contribution of LW radiation to the model output is mostly flat and negative but becomes positive at the highest values. Conversely, during May, a mostly positive contribution from LW radiation is observed which becomes negative at the higher values (Fig. 5c). During June through September, a U-shaped relationship is observed, with a positive contribution from LW radiation at the highest and lowest values, although the most frequent values of LW radiation make a negative contribution. This pattern also suggests that gas- and aqueous-phase processes are both contributing to the model output, with higher values of LW radiation indicating aqueous-phase oxidation and lower values reflecting gas-phase oxidation. Spatially, LW radiation SHAP values show contributions of little magnitude during April and September (Fig. 6c), with slightly positive values in the Atlantic and Pacific sectors and slightly negative over the central Arctic Ocean and continental regions. During May, LW radiation SHAP values are mainly positive geographically except for the Bering Sea region while during June through August, LW radiation SHAP values are positive over the marine areas of the Atlantic and Pacific sectors as well as over the Greenlandic continent (Fig. 6c).

For low cloud cover, the overall relationship is mainly linearly positive and there also appears to be a threshold (~0.7) for positive SHAP values (Fig. 5d). The highest SHAP values are found for the highest values of low cloud cover (which are also the most frequent) and at lower values of low cloud cover, the SHAP values appear to show an increasing tendency (Fig. 5d). This suggests that for low cloud cover both high and low values are contributing to a positive model output of MSA, which also indicates the dual processes of gas- and aqueous phase oxidation of DMS¹⁹ during clear and cloudy conditions, respectively, as highlighted for LW radiation. However, it should be noted that the variability of the response for small values of low cloud cover is large (Fig. 5d). For low cloud cover, positive SHAP values are mainly confined to marine areas while continental areas mainly show negative SHAP values. Low cloud cover SHAP values are consistently positive over the central Arctic Ocean, where the MWT is at its highest values (Supplementary Fig. 1). Low cloud cover SHAP values reach a maximum in the Bering Sea during June and in the Greenland/Barents Sea during July (Fig. 6d). The geographic distribution of low cloud cover SHAP values is similar to those of LW radiation, although with noticeable differences which highlights the complementary nature of these two variables as proxies for aqueous-phase processes. LW radiation SHAP values are negative over the central Arctic Ocean while low cloud cover's SHAP values are mainly positive. This could indicate differences in the cloud phase for this region compared to the marine regions of the Atlantic and Pacific sectors. While low cloud cover indicates that clouds are present, the absence of elevated LW radiation values over the central Arctic Ocean suggests that these clouds are either ice- or mixed-phase clouds, which would emit less LW radiation downwards, as opposed to the more likely liquid-containing clouds in the more southerly regions (Supplementary Fig. 1).

The overall relationships suggest that on average there are two processes making positive contributions to the model output of MSA, gas- and aqueous-phase oxidation of DMS. The positive SHAP values at higher values of SW radiation and lower values of LW radiation suggest that clear sky conditions are also contributing to MSA production through the gas-phase oxidation of DMS. The temperature and low cloud cover threshold values for positive SHAP values as well as the positive SHAP values at higher values of LW radiation (and conversely the negative SHAP values at lower values of SW radiation) indicate that liquid-containing clouds in sufficient amounts are vital for MSA production (i.e., aqueous-phase oxidation of DMS). The model utilized in this study analyzes the geographic and temporal distributions of the explanatory variables based on monthly aggregates, therefore, any temporal information is seasonal and not instantaneous, hence the dual processes identified (gas-phase oxidation in

clear conditions and aqueous-phase oxidation in cloudy conditions). This is supported by the relationships observed for other variables. Higher values of total column cloud liquid water and RH both show positive SHAP values (Supplementary Fig. 7), indicating aqueous-phase processes. It should be noted that both these variables are associated with higher uncertainty in ERA5^{108–110}. Boundary layer height, wind speed, and mean sea level pressure all show indications of a U-shaped pattern for positive SHAP values (Supplementary Fig. 7). Low values of mean sea level pressure are associated with storms that bring cloudy conditions and high wind speeds (which also increases the boundary layer height and promotes vertical mixing in the ocean surface), these conditions are likely associated with the aqueous-phase processes of MSA production. High values of mean sea level pressure are associated with calm and sunny conditions with low wind speeds, which is likely related to the gas-phase processes contributing to MSA production. Peak MSA concentrations are observed in May (Fig. 1b), therefore, we examined the causes for this peak in more detail (Supplementary Note 1). The peak MSA concentrations in May can be linked to strong marine biological activity and meteorological processes which facilitate the efficient transport from southerly latitudes.

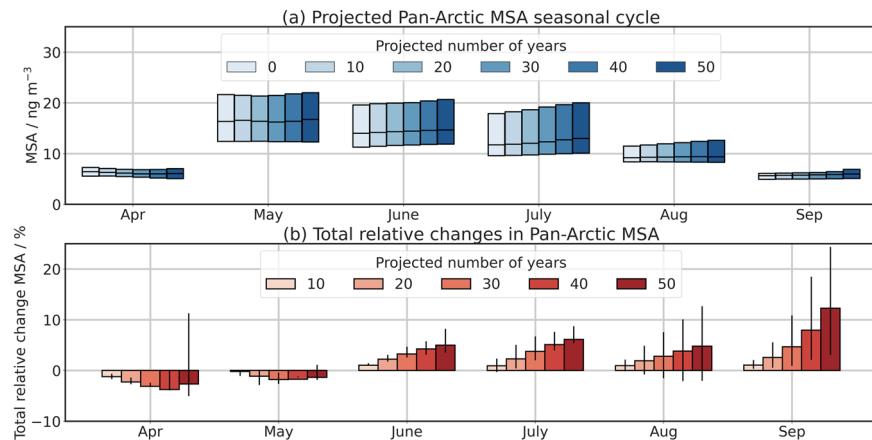
Overall, our model indicates that two main processes are affecting Pan-Arctic MSA concentrations: gas-phase oxidation in clear conditions (indicated through large and mainly positive SHAP values of SW radiation) and aqueous-phase oxidation in cloudy conditions (indicated through U-shaped pattern for LW radiation and low cloud cover values against their SHAP values) which is occurring mainly in the marine areas of the Atlantic and Pacific sectors with temperatures above freezing (threshold for positive temperature SHAP values above 273 K).

Future projections of Pan-Arctic MSA using trends in ERA5 data and machine learning

As discussed in the introduction, MSA concentrations have undergone significant trends, positive and negative, in various parts of the Arctic^{8,45–48} and it is plausible that MSA concentrations will change also in the future. Given the relevance of MSA for cloud formation in the Arctic, it is important to understand how the compound might evolve as environmental conditions change in the Arctic.

The ability of our model to represent the seasonal cycle of MSA concentrations and source regions with confidence allows us to project the future response of Pan-Arctic MSA based on past trends of ERA5. We took the entire available record (1979–2022) of the surface level ERA5 variables used in the ML model, aggregated it to monthly means, and performed a trend analysis using the Mann-Kendall/Theil-Sen methodology at a confidence level (CL) of 95% for each month separately (see Methods for further details). Monthly trends for each variable are displayed in Supplementary Fig. 9. The percentage of variables with statistically significant (SS) trends in each grid cell is displayed in Supplementary Fig. 10 which shows that during April and September, most variables are changing in the central Arctic Ocean and during the other summer months the SS trends are more geographically dispersed. Temperature and LW radiation display mainly positive trends during all months and the highest fraction of grid cells with SS trends, wind speed and mean sea level pressure show the lowest fraction of grid cells with SS trends, while all other variables display a mix of positive and negative trends (Supplementary Fig. 10). We project the slope of the trends for SS grid cells forward for different decadal intervals in the future (10 through 50 years in steps of 10 years) and add the total change over a given period to the original value. We assume only SS grid cells will change and the linear trends will continue at the same rate (see Supplementary Note 2 for discussion of limitations and assumptions of this methodology). We use these projected atmospheric conditions to model the future Pan-Arctic MSA, over different decadal intervals up to year 50, using the previously trained ML model. We have constrained these projected atmospheric conditions to physically realistic values (see Table S3 for details on the constraints). Overall, this resulted in a small fraction of data being imputed, which did not significantly affect the results but ensures our projected atmospheric conditions are physically reasonable.

Fig. 7 | Projected changes in the seasonal cycle of Pan-Arctic MSA. Changes in the distribution of projected Pan-Arctic MSA values for each decadal interval are displayed in **a**. The middle line represents the median while the upper and lower limits of the box represent the 25th/75th percentiles. The total relative change in MSA expressed as a percent for each decadal interval compared to year 0 is displayed in **b**. The vertical lines for each bar represent the upper/lower confidence intervals (CIs) of the Theil-Sen slope used for extrapolating the trends in ERA5.



The median and interquartile range of the projected Pan-Arctic MSA concentrations for each month and different decadal intervals is displayed in Fig. 7a. The distributions show slight changes over the decadal intervals, therefore, we tested for SS differences between MSA concentrations for year 0 and each decadal interval. Depending on the properties of the data (normality or homogeneity of variance), either a two-tailed Welsch's or Student's *t*-test was applied on the 95% CL. All of the distributions for each month and decadal interval are significantly different from year 0 ($p < 0.05$) except for year 10 in May ($p = 0.59$). The median and interquartile range of the absolute and relative changes in MSA concentrations for each month and decadal interval is displayed in Supplementary Fig. 11. It is evident that there are positive and negative changes in MSA over the entire Arctic. Therefore, we calculated the total relative changes in Pan-Arctic MSA for each decadal interval compared to year 0 (See Methods), the results are displayed in Fig. 7b. The bars and vertical lines represent the total relative changes using the slope and the lower/upper confidence intervals (CIs) from the Theil-Sen estimator, respectively, to project changes in MSA. Two patterns are readily evident from Fig. 7b, first the slightly negative, non-monotonic changes during April and May, and second, the positive, monotonic changes during June through September.

The two patterns presented in Fig. 7b allude to a shift in the seasonal cycle of future projections of Pan-Arctic MSA, with small decreases during the late spring/early summer and larger increases during mid to late summer. Currently, the Pan-Arctic seasonal cycle of MSA peaks during May and stays elevated until July (Fig. 1b), whereas our results postulate that the peak duration of the future seasonal cycle is expected to broaden and the peak month will possibly shift from May to later in the summer. Interannual variation in MSA concentrations is not uncommon in the Arctic^{8,41–43,47}, here we project (over long time-scales on the order of decades) that peak Pan-Arctic MSA concentrations will likely undergo a systematic shift to later in the summer although interannual variability is still observed (non-monotonic changes in April/May). The geographic distributions of the changes in MSA reveal that each month experiences both positive and negative changes depending on the geographic location (Supplementary Fig. 12). We focus on the geographic distributions of the year 50 projections as other decadal intervals produce similar patterns but with smaller magnitudes. During April, decreases can be observed over the Arctic Ocean and parts of Siberia while sporadic increases are mainly confined to the periphery of the Arctic, overall the large area with decreases outweighs the sporadic increases. During May, strong decreases are observed in the Greenland, Barents, and Kara Seas as well as the Arctic Ocean, while increases are mainly observed around the southern Greenland, Bering Strait, and the Chukchi Sea. The strong decreases offset the increases to produce an overall decrease during May. During June and July, decreases are mainly confined to the Barents Sea and the northern Pacific Ocean while increases are evident in all other regions, except for the central Arctic Ocean which shows little to no change. During August, decreases are mostly distributed around the marine areas

while increases are observed mainly around the Canadian Archipelago and Siberian coast. The Canadian Archipelago is an area where our model overestimates MSA (Fig. 2e) and the increases in this region are likely an upper limit given the overestimation by our model in this region. The low CWT values and presence of model bias in these regions likely contribute to the large CIs for the projections of MSA in Fig. 7b during August. During September, increases are observed mainly in the Greenland Sea region. Interestingly, the regions around Svalbard are experiencing decreases in MSA during all months except for September and the regions around the Bering Strait are mainly experiencing increases during all months (Supplementary Fig. 12). The geographic distribution of the projected changes in MSA is a combination of the trends for each variable and their contribution to the model output in these regions. Overall, while the Pan-Arctic MSA concentrations are projected to change in the future, these changes are heterogeneously distributed geographically.

To elucidate which variables are responsible for the shift in the seasonal cycle, we devised a leave-one-out approach to calculate the contribution of each variable to the projected changes. To inspect the relationship between the trends in a variable and how said variable contributes to the projected changes, we classified each grid cell according to the direction of the trend and contribution. The leave-one-out and classification methodology are described in more detail in the Methods section. The results of the leave-one-out analysis and classification of the year 50 projections are displayed in Fig. 8 and Supplementary Fig. 15, respectively. The effects of the main variables and their trends are summarized in Table 1. During April, positive RH trends are contributing to the projected decrease in MSA. During May and June, increases in MSA due to positive temperature trends are counterbalanced by decreases in MSA due to positive LW radiation trends. During July and August, a more complex picture is presented. For instance, increases in MSA can be attributed to positive LW radiation and negative BLH trends, and decreases in MSA are due to negative SW radiation trends. During September, positive temperature trends are driving the projected increase in MSA, whereas negative low cloud cover trends are contributing to decreases in MSA. Overall, the contribution of a variable and the relationship to said variable's trends to the projected changes in MSA is highly complex with different variables contributing in different ways in different months. The variables making the highest contribution to projected increases in MSA are temperature, LW radiation, and BLH, while RH, LW, and SW radiation are mainly contributing to the projected decreases in MSA. Of these only BLH and SW radiation show negative trends, highlighting the intricate balance between the (linear) trends of a variable and the (non-linear) relationships with said variable and MSA.

Atmospheric implications

Changes to the MSA seasonal cycle, as projected by our model, will have an effect on the properties of secondary particles in the Arctic which, in turn, will have direct and indirect implications for the radiative balance. Changes

Fig. 8 | Explanatory variable contribution to the projected changes in the Pan-Arctic MSA seasonal cycle. The total relative change in Pan-Arctic MSA as a result of the leave-one-out analysis for each explanatory variable and decadal interval. The colors represent the different explanatory variables and the hatching represents different decadal intervals.

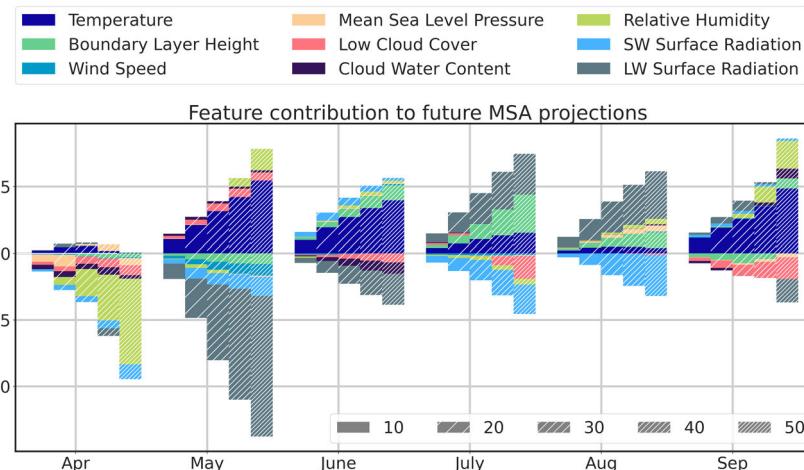


Table 1 | Summary of future projections of MSA, which main explanatory variables are contributing to these projections, and how

	April	May	June	July	August	September
MSA	↓	↓	↑	↑	↑	↑
Variable contribution to leave-one-out analysis	RH ↓	Temp ↑ LW ↓	Temp ↑ LW ↓	LW ↑ BLH ↑ SW ↓	LW ↑ BLH ↑ SW ↓	Temp ↑ LCC ↓
Trend	RH ↑	Temp ↑ LW ↑	Temp ↑ LW ↑	LW ↑ BLH ↓ SW ↓	LW ↑ BLH ↓ SW ↓	Temp ↑ LCC ↓

Changes in MSA, the effect of the main variables on the leave-one-out analysis, and the trends are displayed for each month.

in MSA will affect the physical properties of aerosols, namely mass and volume, which would be especially meaningful during summer and early autumn when total aerosol mass/volume concentrations are at a minimum^{8,91}. The increase in mass/volume will also increase the scattering efficiency, thus having direct implications on the radiative balance. This will also depend on future changes to the underlying surface properties (snow-covered vs bare land or sea ice vs open ocean in a future environment). The changes in MSA concentrations will also likely indirectly affect the aerosol number concentration and therefore CCN concentrations. The additional particulate MSA could affect the number concentration by partitioning into the gas-phase following aqueous-phase formation and droplet evaporation thus providing an additional precursor (or condensable) vapor source. The additional gas-phase MSA could aid in the growth of newly formed particles and possibly participate in nucleation^{54,55}, although this will also depend on changes in sulfuric acid¹¹¹, iodic acid^{56,112}, and other constituents (e.g., reduced nitrogen compounds and organic acids)¹¹³. The number of particles arising from the additional gas-phase MSA that could potentially reach CCN sizes will impact cloud microphysical properties and thus the radiative balance¹¹⁴. This also applies to particles whose hygroscopicity changes from changes in MSA¹¹⁵, as more hygroscopic particles (from additional or less MSA depending on the chemical composition) will more effectively form clouds. These effects would be most pronounced during summer when the sub-micron organic aerosol fraction (which is less hygroscopic) is high^{44,116–119}, sea salt aerosol is low¹²⁰, the condensation sink (CS) is lowest, and the cloud regime is CCN-sensitive⁶⁷. This would also depend on the meteorological conditions (updraught velocities and cloud cover)⁶⁹ and pre-existing aerosol load in a future environment. Clouds mainly reflect solar radiation during the summer (due to decreases in the underlying surface albedo) and our model projects the largest relative increase in MSA during June and July, therefore, the changes in MSA would likely result in additional negative radiative forcing thus cooling the surface. The changes in MSA could also affect the fate of aerosols. The increase in CCN active particles (either through growth to CCN sizes or changing hygroscopicity in existing particles) will affect the wet removal of the aerosol population. More

numerous or more active CCN particles might be preferentially removed via precipitation, therefore even though the source strength is increasing the overall effect on the atmospheric burden might be decreasing, highlighting the complex nature of projecting future changes in aerosol properties^{9,16,18}.

Two important aspects of the environmental system that are pertinent to MSA burdens but are missing from our model are precursor sources (DMS and oxidants) and MSA sinks (wet and dry removal). Studies have shown that in recent decades DMS emission, net primary production, and chlorophyll- α are all increasing in the Arctic^{6,101,121,122}, which could be attributed to a combination of increased open water and nutrient input. Changes in the Arctic ecosystem (shallowing of the oceanic mixed layer depth, increases in sea surface temperature, and acidification)¹⁰⁵ will also likely affect the emission of DMS. Geographically, ice break-up, chlorophyll- α , and polar spring blooms are all moving northward¹²³. Biogeochemical modeling studies suggest that under enhanced CO₂ levels, DMS emissions will decrease annually overall although with vernal decreases related to increased oceanic mixed layer depth and summer increases due to decreases in sea ice extent¹²⁴. These studies suggest that not only are sources of DMS strengthening but they are also moving increasingly northward which will positively affect the overall MSA production. Trends in atmospheric oxidants relevant for MSA production indicate that in the Arctic, ozone^{125,126} and bromine monoxide¹²⁷ are increasing also. A recent modeling study showed bromine levels are expected to increase in the future due to an increase in first-year sea ice and sea salt aerosol¹²⁸. Stevenson et al.¹²⁹ showed that OH levels are increasing globally and in the Arctic. Together these studies advocate for increasing emission of DMS and atmospheric oxidant levels in the future, both of which will positively affect MSA production. Changes in future precipitation patterns will also affect future MSA concentrations, by physically removing aerosols from the atmosphere but also by lowering the CS for vapors. Using accumulated precipitation along back-trajectories, studies have shown that accumulated precipitation is increasing at two High Arctic sites, mainly during autumn^{80,130}. Interestingly, modeling studies have shown precipitation to increase substantially, especially during summer^{131–134}. These studies suggest wet removal will become more

important in a future Arctic climate and will likely affect our projections of MSA.

Anthropogenic activities (i.e., shipping, tourism, and resource extraction) will also undergo future changes^{135,136}, with predicted increases due to sea ice retreat^{137,138} and will have an effect on future aerosol and cloud properties¹³⁹. Increases in anthropogenic activity will likely result in increased emissions of nitrogen oxides (NO_x), volatile organic compounds (VOCs), BC, CO, sulfur oxides (SO_x), greenhouse gases, and production of O₃¹⁴⁰. Increased atmospheric burdens of BC and organic aerosol (from VOC oxidation and condensation) from Arctic anthropogenic activities would increase the CS. This would suggest more condensation of MSA onto pre-existing aerosol surfaces rather than contributing to growth of newly formed particles. This would depend on the chemical composition of the coating material and aerosol physicochemical properties^{65,66}. Indeed increased SO_x emissions (and thus increased sulfate levels) would increase aerosol acidity thus affecting the partitioning of MSA to the gas-phase following aqueous-phase formation¹¹⁴. Increased levels of NO_x and O₃ could alter the branching pathways of DMS oxidation to produce more SO₂ instead of MSA^{19,37,96}. These changes in the aerosol burden due to anthropogenic activities will in turn affect the Arctic environment, such as increased sea ice/snow melt due to increased BC deposition as well as altered nutrient cycling due to inorganic aerosol deposition. The expected increase of Arctic anthropogenic emissions of CO₂ and BC (and reduced emissions of SO_x due to regulations) will lead to positive radiative forcing¹⁴⁰. These increased emissions will likely have the most impact on local air quality in regions experiencing expanded anthropogenic activities (shipping lanes and areas in proximity to industrial activities) and their overall impact on the aerosol population might be small and will likely have negligible impacts on the radiative budget¹³⁸. Overall, the increased anthropogenic activity in the Arctic will affect natural aerosol burdens and dynamics, however, how exactly remains elusive.

This analysis shows that the seasonal cycle of the geographic source regions of MSA can be accurately modeled using the average atmospheric conditions and ML, highlighting the potential use of this methodology to be applied to other natural components of the Arctic atmospheric system, i.e., sea salt aerosol and gas-phase DMS. We also provide guidance as to which atmospheric variables are most important for explaining variations in MSA and how, thus emphasizing the need for an accurate representation of the conditions affecting gas- and aqueous-phase oxidation of DMS in chemical transport models. We also show, based only on changes in atmospheric conditions and not changes in oceanic and biological factors affecting DMS emission/oxidation, that MSA is projected to undergo a shift in the seasonal cycle. Other models and observations have shown that these oceanic and biological factors (sea ice extent/concentration, chlorophyll- α , and DMS emissions) are changing over these time ranges, however, precipitation patterns and intensity are also changing. This suggests that changes in oceanic variables will act in conjunction with changes in atmospheric variables to increase Pan-Arctic MSA concentrations although precipitation will likely affect these changes. The competition between source and sink strength will ultimately determine how aerosol components change in the future. How the Arctic environment responds to future climate change (including atmospheric and oceanic conditions) remains an open research question and provides an avenue for future research.

Methods

In situ aerosol observations

In situ filter samples of methane sulfonic acid (MSA) were collected at four Arctic stations (Alert, Gruvebadet, Thule, and Utqiagvik/Barrow). For Alert, Gruvebadet, and Thule, the sampling period covered 2010–2017, while for Utqiagvik/Barrow, it covered 2008–2014. The period 2010–2017 was selected since MSA measurements were available at most stations from April through September, while MSA observations were available at Utqiagvik/Barrow from 2015–2017 the sampling frequency changed relative to the previous years (2008–2014) and the data coverage is sparse⁴⁵. Therefore, 2008–2014 was selected as the period for Utqiagvik/Barrow due to a consistent sampling frequency and adequate data coverage. Figure 1

displays the location of each station and details about each station are given below. While there are differences in sampling (different inlet and temporal resolution) and analysis (different ion chromatographs) at each station, these are not expected to drastically affect the concentration levels reported from each station. Ion chromatography analysis by two different laboratories for samples from Alert in 2018 showed good agreement¹¹⁸ and in general ion chromatography displays reproducible results¹⁴¹. Therefore, we can be confident that the reported MSA concentrations from different stations are comparable.

Alert, located on Ellesmere Island in Nunavut, Canada (82.5° N, 62.4° W, 210 m above sea level (asl)) is the most northern land-based research station in the High Arctic and has been operational since 1980¹⁴². Aerosols, with no upper size limits, were sampled onto filters (20 × 25 cm Whatman 41) using a high-volume sampler with a duration of seven days. Filters were shipped from Alert to Toronto for analysis. MSA was extracted from one-eighth of the filters using 12 mL of deionized water and sonication for 1 h (>97% extraction efficiency)¹⁴³. Samples were left at room temperature for 48 h before analysis by ion chromatography, with a detection limit ranging from 0.03 to 0.4 ng m⁻³^{47,48}. 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 Na₂B₄O₇ at 2.0 mL min⁻¹. In-between analyses, the column was flushed with 28 mM Na₂B₄O₇ and then re-equilibrated with 5 mM Na₂B₄O₇. A micromembrane suppressor (H₂SO₄) 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 anions¹⁴³. The analytical precision and accuracy are listed at 5 and 2%, respectively, while the uncertainty from random errors during field sampling and laboratory handling/analysis is estimated to be <13%^{143,144}.

Gruvebadet Observatory, located in Ny-Ålesund, Norway on the Svalbard Archipelago (78.9° N, 11.9° E, 50 m asl), has been sampling aerosol chemical composition since 2010 from March to September. The time resolution of each sample was 24 h. Main wind directions indicate that the station is not influenced by local anthropogenic emission sources which are a few kilometers away¹⁴⁵.

The Thule High Arctic Atmospheric Observatory (THAAO) is located in Thule, Greenland (76.5° N, 68.8° W, 220 m asl) off the coast of Baffin Bay. The time resolution of each sample was 48 h. The dominant wind directions indicate that local anthropogenic emission sources, which are located a few kilometers away, are not influencing the station¹⁴⁶.

At Gruvebadet and Thule, aerosols, with a diameter of less than 10 micrometers (PM₁₀), were collected on 47 mm diameter Teflon filters (2 μm nominal porosity and 99% capture efficiency for 0.3 μm diameter particles) using a TECORA Skypost sequential sampler at a flow rate of 2.3 m³ h⁻¹ according to UNI-EN1234⁴². The filters were placed in plastic Petri dishes, frozen, and shipped to Italy for extraction and analysis. Filter quarters were extracted using 10 mL of Milli-Q ultrapure (>18 MΩ) in an ultrasonic bath for 20 min. The complete ionic content of the extracts was determined by using 3 chromatographic systems: 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, using a Gilson 222 autosampler, into a Dionex Thermo-Fischer DX600 ion chromatograph utilizing a Thermo-Fischer Dionex TAC-2 pre-concentration column (50 mL dead-volume) and a Thermo-Fischer Dionex AS11 separation column using a gradient elution of Na₂B₄O₇ solution from 0.075 mM to 2.5 mM as well as electrochemical suppression. This chromatographic method allows the complete separation of MSA from the nearest peak of pyruvate (peak resolution = 0.9). Prior to each analysis, a cleaning step using a 45 mM Na₂B₄O₇ solution was implemented to ensure reproducible results¹⁴⁷. The detection limit was 0.1 μg L⁻¹¹⁴⁷, which field blanks were always below. The reproducibility of the MSA analysis should not differ by more than 5% (<5% relative standard deviation)^{41,42}.

The Barrow Atmospheric Baseline Observatory site is located near the town of Utqiagvik, Alaska (referred to as Utqiagvik/Barrow) (71° N,

156.6° W, 10 m asl) and is part of the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL) Global Monitoring Division (GMD). The wind direction in real-time, only from the clean air sector (0°–129°), was used to avoid influence from local pollution¹⁴⁸. Aerosols were collected using a Berner-type multi-jet cascade impactor with aerodynamic D₅₀ cutoff diameters of 1 and 10 µm (only submicron chemical composition data was used in this study). Sample volumes of 1 (43 m³) to 5 (172 m³) days were collected depending on the time of year. After collection, samples were sealed in tubes and shipped to NOAA's Pacific Marine Environmental Laboratory (PMEL) for chemical analysis by ion chromatography (Metrohm Compact IC 761). Filters were extracted by wetting with 1 mL of spectral grade methanol followed by an additional 5 mL of distilled deionized water and sonicating for 30 min. MSA was analyzed using a Phenomenex Star Ion™ A300 anion column in front of a Metrosep ASUPP5 250/4.0 column with a 1.0 mM NaHCO₃ and 5.0 mM Na₂CO₃ eluent and a 70 mM H₂SO₄ regenerant. The column and eluent combination ensure that MSA is completely resolved from other organic acids (e.g., pyruvic acid). The detection limit was 0.001 mg L⁻¹⁴⁵. Concentrations of MSA are reported at 0 °C and 1013 mbar. The relative uncertainty including errors arising from the ion chromatography analysis, blanks, liquid extract volume, and volume of sampled air is ±11% (95% confidence level)^{149–151}.

ERA5

ERA5 is the fifth-generation atmospheric reanalysis product from ECMWF⁸⁴. The Integrated Forecast System (IFS) cycle 41r2 is the numerical model forming the basis of ERA5. ERA5 is an improvement on the previous reanalysis from ECMWF (ERA-Interim), with improved spatial and temporal resolution, and is considered the best guess of the state of the atmosphere. ERA5 is available for the entire globe at hourly resolution, although in this study ERA5 output for north of 45° N and every third hour was used to match the geographical extent and temporal resolution of the FLEXPART output. The surface-level (or single-level) ERA5 dataset on a 0.5° × 0.5° regular latitude-longitude grid with a spatial resolution of ~31 km was utilized. Relative humidity was calculated using 2 m air temperature and dew point temperature following the method of Pernov et al.¹¹⁰. Here we use ERA5 data from 1 April to 30 September for 2008–2017 for the MWTs and data from 1979–2022 for the trend analysis. Data was obtained from the ECMWF Climate Data Store (CDS) (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview>, last accessed November 24, 2022). For a complete list of the variables used in this study see Table S1. While ERA5 is one of the best-performing reanalysis datasets available¹⁵², it is not without its limitations including non-physical variability and trends due to changes in the observing system (although mainly affects high altitudes), shifts in tropospheric oceanic humidity due to the incorporation of microwave imager data (this is only present in the 1990s), and a warm bias of 5 K in the surface temperature over Arctic sea ice (mainly present during winter)^{84,153,154}. Recently, ERA5 surface level variables were compared against continental ground-based stations spanning at least one decade for most sites. Overall ERA5 performed well for temperature, solar radiation, and pressure although less so for relative humidity and wind speed/direction¹¹⁰. These limitations will not affect our results or our interpretations.

FLEXPART

Simulations of gridded air mass residence times were performed with the Lagrangian particle dispersion model FLEXPART v9.1⁸³ to determine the source origins and transport pathways. The model has been driven with meteorological data from the ERA5 reanalysis with 0.5° × 0.5° resolution and 137 vertical levels available every third hour. A release of 50,000 tracer particles was initialized every three hours at the location of the atmospheric observatories and traced backward in time for up to 10 days. For Alert, Thule, and Utqiagvik/Barrow, particles were released at 10 m above ground level (agl). For Gruvebadet, to account for the complex topography, particles were released within the altitude range of 10–100 m agl. The simulations

were computed for a passive air tracer without removal processes. The main output from FLEXPART consists of 3-dimensional fields of residence time in units of seconds. Lagrangian dispersion models have been shown to be more representative than Eulerian or semi-Lagrangian models but they are not without their limitations including simplification of atmospheric processes (deposition, advection, and diffusion)⁸³, limitations in the vertical resolution¹⁵⁵, and choice of input meteorological data¹⁵⁶. We calculated a Pan-Arctic transport climatology which is expressed as a normalized residence time for each month separately. For every FLEXPART release during a given month, the residence time in each grid cell was summed below an altitude of 1000 m and for a transport length of seven days backward in time. This total residence time was then divided by the grid cell area and normalized by the sum of all grid cells over the geographical domain (>50° N) and expressed as a percentage to give the normalized residence time. The output of the FLEXPART model is seconds and divided by the grid cell area gives units of s km⁻², therefore, the normalized residence time allows for a meaningful comparison across all latitudes. Similar approaches have been utilized in the Arctic before to display transport patterns^{80,91,157}. Considering these different transport parameters, we selected a transport length of 7 days to capture the timeframe of DMS emission and oxidation to MSA and a maximum altitude of 1000 m to capture the transport pathways not only in the boundary layer but also in the lower free troposphere. We performed a sensitivity analysis for these selections by varying the maximum altitude (200, 500, and 1000 m) and transport length (5, 7, and 10 days). Overall, extending the transport length and maximum altitude expanded the trajectory footprint, although, the general interpretations of the transport climatologies, potential source contribution function, and concentration-weighted trajectory remain unchanged.

Potential source contribution function

The Potential Source Contribution Function (PSCF) was used to identify the probability of MSA source regions at the different stations^{158–160}. It should be noted that the PSCF represents the probability for potential source regions and preferred transport pathways for an atmospheric species at a receptor site and is a non-quantitative source analysis and as such is limited in distinguishing strong from moderate sources¹⁶⁰. The 75th percentile of MSA in a given month was used as a threshold value as is common in the literature^{79,161–163}. Due to the long sampling times of the filter packs, all trajectories occurring during a sample above the threshold were used in the PSCF calculation. The sum of the residence time for samples above the 75th percentile was normalized by the sum of the residence for all periods when MSA observations were available according to Eq. (1).

$$PSCF_{ij} = \frac{m_{ij}}{n_{ij}} \quad (1)$$

Where n_{ij} is the total residence time in grid cell ij and m_{ij} is the total residence time when the concentration was above the 75th percentile. The weighting scheme (W) adapted from Bressi et al.¹⁶⁴ was used to down-weight grid cells with low residence times.

$$W_{ij} = \begin{cases} 0.75 & 0.75 * \max(\log(n_{ij} + 1)) < n_{ij} < \max(\log(n_{ij} + 1)) \\ 0.5 & 0.5 * \max(\log(n_{ij} + 1)) < n_{ij} < 0.75 * \max(\log(n_{ij} + 1)) \\ 0.15 & 0.25 * \max(\log(n_{ij} + 1)) < n_{ij} < 0.5 * \max(\log(n_{ij} + 1)) \\ 0.05 & n_{ij} < 0.25 * \max(\log(n_{ij} + 1)) \end{cases}$$

Concentration weighted trajectory

The concentration-weighted trajectory (CWT) was used to estimate the source contribution of geographical regions to the observed MSA at the receptor sites and expresses the average contribution to the concentration expected if a trajectory passes over a given grid cell^{165,166}. CWT is capable of determining the relative strength of potential sources and transport

pathways by incorporating MSA concentrations at the receptor site with FLEXPART trajectories according to Eq. (2):

$$C_{ij} = \frac{\sum_{l=1}^S C_l \tau_{i,j,l}}{\sum_{l=1}^S \tau_{i,j,l}} \quad (2)$$

where C_{ij} is the trajectory weighted average concentration in the grid cell ij , l is the index of the sample, S is the total number of samples, C_l is the concentration observed at the sampling location (receptor site) during sample l , and $\tau_{i,j,l}$ is the residence time in grid cell ij during sample l . The CWT was downweighted in the same manner as the meteorological weighted trajectory (MWT) as described below.

Applying the weighting scheme used for the PSCF would result in meteorological variables being substantially reduced beyond realistic values if grid cells experienced the lower threshold of the weighting scheme. This approach is valid for variables that have a lower limit of zero, although for variables with a continuous distribution (e.g., temperature), applying this weighting scheme would result in unrealistic values. To overcome this, we applied a cutoff value of $0.25 * \max(\log(n_{ij} + 1))$, with grid cells below this value being replaced with NaN and grid cells above this value being kept as is. Different levels for this threshold were tested and were found not to majorly influence the results. We ultimately chose this value (i.e., the second level of the weighting scheme) as a compromise between including grid cells with a large enough residence time to be statistically robust and obtaining a large enough geographic footprint that represents the actual transport pathways.

Meteorological weighted trajectory

The MWT was used to estimate the environmental conditions of geographical regions of air masses that influenced the receptor sites. The MWT can be thought of as trajectory-weighted average meteorological conditions of air masses ending at the receptor site. The MWT was calculated using ERA5 reanalysis data and FLEXPART back trajectories. The ERA5 data was averaged using the arithmetic mean during the filter collection periods as well as during the preceding days corresponding to the trajectory length (7 days). This allows us to weigh the average meteorological conditions by the air mass residence time thus producing the trajectory-weighted meteorological conditions during this filter sample. Averaging over the filter collection period and preceding days was employed to resemble the situation for MSA collection, as conditions during filter collection and the preceding days of transport to the measurement site affect the measured concentration for a particular filter. The MWT closely reflects the monthly mean ERA5 values (Supplementary Figs. 1 and 2) but accounts for spatial variability in meteorological conditions and air mass transport pathways. The MWT also requires a cutoff value to be applied to reduce the influence of grid cells with low residence times (therefore we have less statistical confidence in them), which is the same used in the CWT calculation. The MWT was calculated according to Eq. (3):

$$M_{ij} = \frac{\sum_{l=1}^S M_l \tau_{i,j,l}}{\sum_{l=1}^S \tau_{i,j,l}} \quad (3)$$

where M_{ij} is the trajectory weighted average meteorological conditions in the grid cell ij , l is the index of the sample, S is the total number of samples, M_l is the mean meteorological conditions during sample l and preceding days, and $\tau_{i,j,l}$ is the residence time in grid cell ij during sample l .

Machine learning methodology

In this study, we utilize a supervised, regression form of ML. Here we describe the explanatory variables and target variables, variable selection, hyperparameter tuning, the machine learning model, and the explainability of the model.

The target variable used was the CWT values of MSA for individual grid cells over all summer months. The explanatory variables (or features) used were the MWT values of the ERA5 data for individual grid cells over all summer months. No additional feature engineering (e.g., standardization or normalization) was applied prior to modeling. All summer months were modeled together to better capture the seasonal cycle of the target and explanatory variables as well as capture the dynamic range of MSA values. Since the target we are modeling is geographically distributed and not temporally (i.e., time series), the variables and target were split into training and test sets based on their geographic distribution. The area north of 50° N was divided into 20 equally spaced groups of 18° each, these groups were further divided into 4 equally spaced sub-groups of 4.5° each to obtain a 75/25 train/test split. Three sub-groups were used as a training set while the fourth was used as a test set, this process was repeated using each sub-group as a test set. A schematic of the geographic distribution of train/test sets is depicted in Supplementary Fig. 4.

A set of 15 explanatory variables from ERA5, which plausibly could contribute to atmospheric MSA production, was initially selected as explanatory variables (Supplementary Table 1). From this initial set, the final explanatory variables were selected using recursive feature elimination. Starting with 15 variables, one variable per iteration was recursively eliminated. Elimination was based on the importance of each variable as assigned by the model, whereby the variable with the lowest importance was eliminated. This process was repeated until one variable remained. The number of variables was evaluated using the R^2 score from the training sets using a different number of variables. The R^2 scores from the test set were also calculated for comparison. The results of this analysis are displayed in Supplementary Fig. 3. Ultimately, nine variables were selected, based on the decrease of R^2 scores with a fewer number of variables. The recursive feature elimination was implemented via the scikit-learn package (v1.0.2). All modeling was performed in the Python programming language (v3.10.2).

We tested ensemble and boosted (ensemble) methods, such as Random Forests, Adaptive Boosting, (extreme) gradient-boosted regression trees, and gradient-boosted generalized additive models, as implemented in several python libraries (respectively, scikit-learn, AdaBoost, XGBoost, CatBoost, LightGBM, and InterpretML). All tested models produced similar validation metrics, although we found the Gradient-boosted regression trees as implemented in XGBoost to be the most computationally efficient approach. The XGBoost library implements gradient boosting of shallow decision trees. It works by sequentially training decision trees to minimize the residuals of the trees already in the model⁹². This allows for the residuals from the previous trees to be learned, reducing the regression loss function and obtaining accurate predictions by passing data through the trained (shallow) trees sequentially. XGBoost implements a regularized model formalization to prevent overfitting and to improve computational efficiency. Finally, XGBoost is also able to handle collinearity amongst the explanatory variables, which is important for this study as we include several variables related to the aqueous phase processes. We use the XGBoost package (v1.6.2)⁸⁵.

The XGBoost implementation of gradient-boosted trees requires hyperparameter tuning to ensure optimal performance. We used a Bayesian search approach for testing the optimum configuration of hyperparameters, which was implemented using the Optuna⁹³ library (v3.0.3). The tested hyperparameters the range of values explored, and the optimum values are listed in Supplementary Table 2. The objective of the hyperparameter tuning procedure is to maximize the mean R^2 score. Tuning was performed for 250 trials; preliminary tuning revealed a plateauing of the mean R^2 score after ~150 trials indicating that 250 trials are sufficient to ensure the optimal setting was selected. Hyperparameter values were sampled using the Tree-structured Parzen Estimator (TPE) algorithm¹⁶⁷ and trials were pruned using the Hyperband pruner¹⁶⁸. The final set of hyperparameters was selected based on the compromise between overall performance (high R^2 scores) and agreement between the training and test set evaluation metrics (prevention of overfitting).

We utilized SHapley Additive exPlanations (SHAP) values⁸⁵ which is based on Shapley values¹⁶⁹, to evaluate the impact of the explanatory variables on the model output. SHAP is a model-agnostic methodology designed to evaluate variable importance based on coalitional game theory¹⁷⁰, where each explanatory variable is treated as a “player” in a “game” (model framework) and SHAP aims to evaluate each player’s contribution to the “payout” (model output). For each grid cell, the SHAP values represent a variable’s marginal contribution over the mean model output considering all possible combinations of explanatory variables. SHAP can provide both local and global explanations as opposed to other commonly used variable importance methods (such as those provided by XGBoost itself: gain, split count, permutation importance) that only provide a measure of global importance¹⁷¹. The global importance for each variable is calculated as the median of the absolute SHAP values for said variable which gives an overview of the most important variables. However, it does not consider the relationships between the SHAP and explanatory variable value (positive or negative relationship, linear or non-linear). Therefore, we assessed the relationship between the explanatory and SHAP values by discretizing the explanatory values into 10 equally spaced bins and calculated the median and 25th/75th percentiles for each bin. These two approaches allow us to evaluate the overall global importance and the relationship between explanatory and SHAP values for each variable. The local importance was assessed by creating maps of SHAP values for each variable, which allows us to illuminate the geographic distribution of each variable’s contribution to the model output. The SHAP approach was applied via the shap package (v0.41.0).

Trend analysis

Trends in ERA5 environmental variables for each grid cell were determined using the non-parametric Mann-Kendall test and Theil Sen’s slope methodology¹⁷²⁻¹⁷⁵ and calculated using monthly means. These tests require the data to be serially independent and homogeneously distributed, however, the environmental variables often exhibit statistically significant (SS) lag-1 autocorrelation and high amounts of seasonality. We have calculated trends for each month separately to account for seasonality and used monthly means to minimize auto-correlation. SS trends were determined at the 95th % confidence level (CL). Uncertainty in trends was assessed using the 95th % confidence interval (CI) of the slope.

Future projections methodology

To project the response of MSA concentration in the future, we extrapolated the MWT values for grid cells that exhibited SS trends. We assumed the SS trends would continue at the same rate in the future and calculated the absolute change over a given decadal interval (10, 20, 30, 40, and 50 years in the future) and added these changes to the original MWT values (referred to year 0). Extrapolating the trends of meteorological variables into the future can result in physically unrealistic values (e.g., fractional cloud cover greater than 1), therefore, we applied constraints to the extrapolated values (Table S4). We imputed values of the lower/upper limit if extrapolated values exceeded these limits for explanatory variables which are physically bounded (low cloud cover, relative humidity, cloud liquid water). For boundary layer height, we imposed a lower limit of 20 m to avoid physically unrealistic values, using lower limits of 10 and 30 did not produce any meaningful differences. Overall, only a small percentage of the data was imputed to avoid physically unrealistic values (Table S4). We then used the previously trained XGBoost model to project the MSA concentrations for each of the five decades. To assess the total changes in Pan-Arctic MSA, we use Eq. (4) to calculate the total relative change of MSA expressed as a percentage for each year T .

$$\text{Total Relative Delta}_T = \frac{\sum_{i=1}^I \sum_{j=1}^J (MSA_{T,i,j} - MSA_{0,i,j})}{\sum_{i=1}^I \sum_{j=1}^J MSA_{0,i,j}} * 100\% \quad (4)$$

Where I and J are the total number of meridional and zonal grid cells, respectively. MSA_T and MSA_0 are the Pan-Arctic MSA concentrations for year T in the future and year 0, respectively.

To evaluate how each variable contributed to the future projections of MSA in certain decades, we applied a leave-one-out approach to the trends for each variable. We projected the future changes in MSA by not applying the extrapolated trends for one variable at a time for each decade interval. We compared these results to the results obtained by including all variables for the same decade interval to give the total relative delta for a decade interval T and without variable V , as displayed in Eq. (5).

$$\text{Total Relative Delta}_{T,V} = \frac{\sum_{i=1}^I \sum_{j=1}^J (MSA_{T,A,i,j} - MSA_{X,V,i,j})}{\sum_{i=1}^I \sum_{j=1}^J MSA_{T,A,i,j}} * 100\% \quad (5)$$

Where I and J are the total number of meridional and zonal grid cells, respectively. $MSA_{T,V}$ and $MSA_{T,A}$ are the Pan-Arctic MSA concentrations for year T without variable V and for year T with all variables included, respectively. To investigate the relationship between the contribution of each variable to the leave-one-out analysis and the trends in said variable, we classified each grid cell according to the direction of the trend (positive or negative) and the contribution to the leave-one-out analysis (positive or negative). This was performed only for grid cells that exhibited SS trends. The classification included four categories: positive trend and positive variable contribution, positive trend and negative variable contribution, negative trend and positive variable contribution, and negative trend and negative variable contribution. The results of this analysis are displayed in Supplementary Fig. 15.

Data availability

The datasets used and/or analyzed during the current study are available on reasonable request from the corresponding author Jakob Boyd Pernov (jakob.pernov@epfl.ch). ERA5 data is available from the CDS (<https://cds.climate.copernicus.eu/#/home>, last accessed 08/11/2022). In situ MSA is available online for Utqiagvik/Barrow (<https://data.pmel.noaa.gov/pmel/erddap/tabledap/submicron.html>) and Alert (<https://ebas.nilu.no/>) while Thule and Gruvebadet are available upon request. FLEXPART is available upon reasonable request.

Code availability

The underlying code for this study is available as a Renkulab project (<https://renkulab.io/projects/arcticnap/arcticnap/files/blob/src>) and by contacting the corresponding author Jakob Boyd Pernov (jakob.pernov@epfl.ch). Code for FLEXPART, Optuna, XGBoost, and SHAP packages are available online

Received: 21 February 2024; Accepted: 3 July 2024;

Published online: 13 July 2024

References

1. Rantanen, M. et al. The Arctic has warmed nearly four times faster than the globe since 1979. *Commun. Earth Environ.* **3**, 1–10 (2022).
2. Stroeve, J. & Notz, D. Changing state of Arctic sea ice across all seasons. *Environ. Res. Lett.* **13**, 103001 (2018).
3. O’Neill, H. B., Smith, S. L., Burn, C. R., Duchesne, C. & Zhang, Y. Widespread Permafrost Degradation and Thaw Subsidence in Northwest Canada. *J. Geophys. Res. Earth Surf.* **128**, e2023JF007262 (2023).
4. King, M. D. et al. Dynamic ice loss from the Greenland Ice Sheet driven by sustained glacier retreat. *Commun. Earth Environ.* **1**, 1 (2020).
5. Morison, J. et al. Changing Arctic Ocean freshwater pathways. *Nature* **481**, 66–70 (2012).
6. Oziel, L., Massicotte, P., Babin, M. & Devred, E. Decadal changes in Arctic Ocean Chlorophyll a: Bridging ocean color observations from

the 1980s to present time. *Remote Sens. Environ.* **275**, 113020 (2022).

7. Intergovernmental Panel on Climate Change (IPCC). In IPCC Special Report on the Ocean and Cryosphere in a Changing Climate (eds. Pörtner, H.-O. et al.) (IPCC, 2019).
8. Schmale, J. et al. Pan-Arctic seasonal cycles and long-term trends of aerosol properties from 10 observatories. *Atmos. Chem. Phys.* **22**, 3067–3096 (2022).
9. Schmale, J., Zieger, P. & Ekman, A. M. L. Aerosols in current and future Arctic climate. *Nat. Clim. Chang.* **11**, 95–105 (2021).
10. Twomey, S. The Influence of Pollution on the Shortwave Albedo of Clouds. *J. Atmos. Sci.* **34**, 1149–1152 (1977).
11. Albrecht, B. A. Aerosols, Cloud Microphysics, and Fractional Cloudiness. *Science* **245**, 1227–1230 (1989).
12. Carslaw, K. S. et al. Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature* **503**, 67–71 (2013).
13. Regayre, L. A. et al. Uncertainty in the magnitude of aerosol-cloud radiative forcing over recent decades. *Geophys. Res. Lett.* **41**, 9040–9049 (2014).
14. Watson-Parris, D. et al. Constraining Uncertainty in Aerosol Direct Forcing. *Geophys. Res. Lett.* **47**, e2020GL087141 (2020).
15. Regayre, L. A. et al. The value of remote marine aerosol measurements for constraining radiative forcing uncertainty. *Atmos. Chem. Phys.* **20**, 10063–10072 (2020).
16. Browse, J. et al. The complex response of Arctic aerosol to sea-ice retreat. *Atmos. Chem. Phys.* **14**, 7543–7557 (2014).
17. Ridley, J. K., Ringer, M. A. & Sheward, R. M. The transformation of Arctic clouds with warming. *Clim. Change* **139**, 325–337 (2016).
18. Mahmood, R., von Salzen, K., Norman, A.-L., Galí, M. & Levasseur, M. Sensitivity of Arctic sulfate aerosol and clouds to changes in future surface seawater dimethylsulfide concentrations. *Atmos. Chem. Phys.* **19**, 6419–6435 (2019).
19. Barnes, I., Hjorth, J. & Mihalopoulos, N. Dimethyl Sulfide and Dimethyl Sulfoxide and Their Oxidation in the Atmosphere. *Chem. Rev.* **106**, 940–975 (2006).
20. Levasseur, M. Impact of Arctic meltdown on the microbial cycling of sulphur. *Nat. Geosci.* **6**, 691–700 (2013).
21. Mungall, E. L. et al. Dimethyl sulfide in the summertime Arctic atmosphere: measurements and source sensitivity simulations. *Atmos. Chem. Phys.* **16**, 6665–6680 (2016).
22. Park, K. et al. Unexpectedly high dimethyl sulfide concentration in high-latitude Arctic sea ice melt ponds. *Environ. Sci.: Process. Impacts* **21**, 1642–1649 (2019).
23. Andreae, M. O. Ocean-atmosphere interactions in the global biogeochemical sulfur cycle. *Mar. Chem.* **30**, 1–29 (1990).
24. Kettle, A. J. et al. A global database of sea surface dimethylsulfide (DMS) measurements and a procedure to predict sea surface DMS as a function of latitude, longitude, and month. *Glob. Biogeochem. Cycles* **13**, 399–444 (1999).
25. Lana, A. et al. An updated climatology of surface dimethylsulfide concentrations and emission fluxes in the global ocean. *Global Biogeochem. Cycles* **25**, GB1004 (2011).
26. Hulswar, S. et al. Third revision of the global surface seawater dimethyl sulfide climatology (DMS-Rev3). *Earth Syst. Sci. Data* **14**, 2963–2987 (2022).
27. Sjostedt, S. J. et al. Evidence for the uptake of atmospheric acetone and methanol by the Arctic Ocean during late summer DMS-Emission plumes. *J. Geophys. Res. Atmos.* **117**, n/a–n/a (2012).
28. Park, K.-T. et al. Linking atmospheric dimethyl sulfide and the Arctic Ocean spring bloom. *Geophys. Res. Lett.* **40**, 155–160 (2013).
29. Park, K.-T. et al. Atmospheric DMS in the Arctic Ocean and Its Relation to Phytoplankton Biomass. *Glob. Biogeochem. Cycles* **32**, 351–359 (2018).
30. Pernov, J. B. et al. Atmospheric VOC measurements at a High Arctic site: characteristics and source apportionment. *Atmos. Chem. Phys.* **21**, 2895–2916 (2021).
31. Park, K.-T. et al. Dimethyl Sulfide-Induced Increase in Cloud Condensation Nuclei in the Arctic Atmosphere. *Glob. Biogeochem. Cycles* **35**, e2021GB006969 (2021).
32. Ghahreman, R. et al. Dimethyl sulfide and its role in aerosol formation and growth in the Arctic summer – a modelling study. *Atmos. Chem. Phys.* **19**, 14455–14476 (2019).
33. Sørensen, S., Falbe-Hansen, H., Mangoni, M., Hjorth, J. & Jensen, N. R. Observation of DMSO and CH₃S(O)OH from the gas phase reaction between DMS and OH. *J. Atmos. Chem.* **24**, 299–315 (1996).
34. Wollesen de Jonge, R. et al. Secondary aerosol formation from dimethyl sulfide – improved mechanistic understanding based on smog chamber experiments and modelling. *Atmos. Chem. Phys.* **21**, 9955–9976 (2021).
35. von Glasow, R. & Crutzen, P. J. Model study of multiphase DMS oxidation with a focus on halogens. *Atmos. Chem. Phys.* **4**, 589–608 (2004).
36. Hoffmann, E. H. et al. An advanced modeling study on the impacts and atmospheric implications of multiphase dimethyl sulfide chemistry. *Proc. Natl. Acad. Sci. USA* **113**, 11776–11781 (2016).
37. Chen, Q., Sherwen, T., Evans, M. & Alexander, B. DMS oxidation and sulfur aerosol formation in the marine troposphere: a focus on reactive halogen and multiphase chemistry. *Atmos. Chem. Phys.* **18**, 13617–13637 (2018).
38. Baccarini, A. et al. Low-Volatility Vapors and New Particle Formation Over the Southern Ocean During the Antarctic Circumnavigation Expedition. *J. Geophys. Res. Atmos.* **126**, e2021JD035126 (2021).
39. Fung, K. M. et al. Exploring dimethyl sulfide (DMS) oxidation and implications for global aerosol radiative forcing. *Atmos. Chem. Phys.* **22**, 1549–1573 (2022).
40. Kecorius, S. et al. Rapid growth of Aitken-mode particles during Arctic summer by fog chemical processing and its implication. *PNAS Nexus* **2**, pgad124 (2023).
41. Becagli, S. et al. Relationships linking primary production, sea ice melting, and biogenic aerosol in the Arctic. *Atmos. Environ.* **136**, 1–15 (2016).
42. Becagli, S. et al. Biogenic Aerosol in the Arctic from Eight Years of MSA Data from Ny Ålesund (Svalbard Islands) and Thule (Greenland). *Atmosphere* **10**, 349 (2019).
43. Jang, S. et al. Large seasonal and interannual variations of biogenic sulfur compounds in the Arctic atmosphere (Svalbard; 78.9° N, 11.9° E). *Atmos. Chem. Phys.* **21**, 9761–9777 (2021).
44. Nøgaard, J. K. et al. A local marine source of atmospheric particles in the High Arctic. *Atmos. Environ.* **285**, 119241 (2022).
45. Moffett, C. E. et al. Long-Term Trends for Marine Sulfur Aerosol in the Alaskan Arctic and Relationships With Temperature. *J. Geophys. Res. Atmos.* **125**, e2020JD033225 (2020).
46. Laing, J. R. et al. Long-term trends of biogenic sulfur aerosol and its relationship with sea surface temperature in Arctic Finland. *J. Geophys. Res. Atmos.* **118**, 11,770–11,776 (2013).
47. Sharma, S. et al. Influence of transport and ocean ice extent on biogenic aerosol sulfur in the Arctic atmosphere. *J. Geophys. Res. Atmos.* **117**, (2012).
48. Sharma, S. et al. A Factor and Trends Analysis of Multidecadal Lower Tropospheric Observations of Arctic Aerosol Composition, Black Carbon, Ozone, and Mercury at Alert. *Can. J. Geophys. Res. Atmos.* **124**, 14133–14161 (2019).
49. Mungall, E. L., Wong, J. P. S. & Abbott, J. P. D. Heterogeneous Oxidation of Particulate Methanesulfonic Acid by the Hydroxyl Radical: Kinetics and Atmospheric Implications. *ACS Earth Space Chem.* **2**, 48–55 (2018).

50. Kerminen, V.-M., Aurela, M., Hillamo, R. E. & Virkkula, A. Formation of particulate MSA: deductions from size distribution measurements in the Finnish Arctic. *Tellus B: Chem. Phys. Meteorol.* **49**, 159–171 (1997).

51. Phinney, L. et al. Characterization of the aerosol over the sub-arctic north east Pacific Ocean. *Deep-Sea Res. II* **53**, 2410–2433 (2006).

52. Xavier, C. et al. Secondary aerosol formation in marine Arctic environments: a model measurement comparison at Ny-Ålesund. *Atmos. Chem. Phys.* **22**, 10023–10043 (2022).

53. Lawler, M. J. et al. New Insights Into the Composition and Origins of Ultrafine Aerosol in the Summertime High Arctic. *Geophys. Res. Lett.* **48**, e2021GL094395 (2021).

54. Chang, R. Y.-W. Relating atmospheric and oceanic DMS levels to particle nucleation events in the Canadian Arctic. *J. Geophys. Res. Atmos.* **116**, D00S03 (2011).

55. Ning, A. & Zhang, X. The synergistic effects of methanesulfonic acid (MSA) and methanesulfonic acid (MSIA) on marine new particle formation. *Atmos. Environ.* **269**, 118826 (2022).

56. Li, J., Wu, N., Chu, B., Ning, A. & Zhang, X. Molecular-level study on the role of methanesulfonic acid in iodine oxoacid nucleation. *Atmos. Chem. Phys.* **24**, 3989–4000 (2024).

57. Dall’Osto, M. et al. Abiotic and biotic sources influencing spring new particle formation in North East Greenland. *Atmos. Environ.* **190**, 126–134 (2018).

58. Beck, L. J. et al. Differing Mechanisms of New Particle Formation at Two Arctic Sites. *Geophys. Res. Lett.* **48**, e2020GL091334 (2021).

59. Rosati, B. et al. New Particle Formation and Growth from Dimethyl Sulfide Oxidation by Hydroxyl Radicals. *ACS Earth Space Chem.* **5**, 801–811 (2021).

60. Ghahreman, R. et al. Modeling Aerosol Effects on Liquid Clouds in the Summertime Arctic. *J. Geophys. Res. Atmos.* **126**, e2021JD034962 (2021).

61. Twomey, S. A., Piepgrass, M. & Wolfe, T. L. An assessment of the impact of pollution on global cloud albedo. *Tellus B: Chem. Phys. Meteorol.* **36**, 356–366 (1984).

62. Hansen, J., Sato, M. & Ruedy, R. Radiative forcing and climate response. *J. Geophys. Res. Atmos.* **102**, 6831–6864 (1997).

63. Rosenfeld, D. TRMM observed first direct evidence of smoke from forest fires inhibiting rainfall. *Geophys. Res. Lett.* **26**, 3105–3108 (1999).

64. Ramanathan, V., Crutzen, P. J., Kiehl, J. T. & Rosenfeld, D. Atmosphere - Aerosols, climate, and the hydrological cycle. *Science* **294**, 2119–2124 (2001).

65. Yan, J. et al. Uptake selectivity of methanesulfonic acid (MSA) on fine particles over polynya regions of the Ross Sea, Antarctica. *Atmos. Chem. Phys.* **20**, 3259–3271 (2020).

66. Dada, L. et al. A central arctic extreme aerosol event triggered by a warm air-mass intrusion. *Nat. Commun.* **13**, 5290 (2022).

67. Mauritsen, T. et al. An Arctic CCN-limited cloud-aerosol regime. *Atmos. Chem. Phys.* **11**, 165–173 (2011).

68. Birch, C. E. et al. Modelling atmospheric structure, cloud and their response to CCN in the central Arctic: ASCOS case studies. *Atmos. Chem. Phys.* **12**, 3419–3435 (2012).

69. Motos, G. et al. Aerosol and dynamical contributions to cloud droplet formation in Arctic low-level clouds. *Atmos. Chem. Phys.* **23**, 13941–13956 (2023).

70. Nair, A. A. et al. Machine Learning Uncovers Aerosol Size Information From Chemistry and Meteorology to Quantify Potential Cloud-Forming Particles. *Geophys. Res. Lett.* **48**, e2021GL094133 (2021).

71. McNabb, B. J. & Tortell, P. D. Improved prediction of dimethyl sulfide (DMS) distributions in the northeast subarctic Pacific using machine-learning algorithms. *Biogeosciences* **19**, 1705–1721 (2022).

72. Mansour, K., Decesari, S., Ceburnis, D., Ovadnevaite, J. & Rinaldi, M. Machine learning for prediction of daily sea surface dimethylsulfide concentration and emission flux over the North Atlantic Ocean (1998–2021). *Sci. Total Environ.* **871**, 162123 (2023).

73. Chen, Y. et al. Machine learning reveals climate forcing from aerosols is dominated by increased cloud cover. *Nat. Geosci.* 1–6 (2022) <https://doi.org/10.1038/s41561-022-00991-6>.

74. Nair, A. A. & Yu, F. Using machine learning to derive cloud condensation nuclei number concentrations from commonly available measurements. *Atmos. Chem. Phys.* **20**, 12853–12869 (2020).

75. Yu, F., Luo, G., Nair, A. A., Tsigaridis, K. & Bauer, S. E. Use of Machine Learning to Reduce Uncertainties in Particle Number Concentration and Aerosol Indirect Radiative Forcing Predicted by Climate Models. *Geophys. Res. Lett.* **49**, e2022GL098551 (2022).

76. Dall’Osto, M. et al. Arctic sea ice melt leads to atmospheric new particle formation. *Sci. Rep.* **7**, 3318 (2017).

77. Lange, R. et al. Characterization of distinct Arctic aerosol accumulation modes and their sources. *Atmos. Environ.* **183**, 1–10 (2018).

78. Dall’Osto, M. et al. Regions of open water and melting sea ice drive new particle formation in North East Greenland. *Sci. Rep.* **8**, 6109 (2018).

79. Song, C. et al. Differentiation of coarse-mode anthropogenic, marine and dust particles in the High Arctic islands of Svalbard. *Atmos. Chem. Phys.* **21**, 11317–11335 (2021).

80. Pernov, J. B. et al. Increased aerosol concentrations in the High Arctic attributable to changing atmospheric transport patterns. *npj Clim. Atmos. Sci.* **5**, 1–13 (2022).

81. Song, C. et al. 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 (2022).

82. Hu, C. et al. Molecular characteristics, sources and influencing factors of isoprene and monoterpenes secondary organic aerosol tracers in the marine atmosphere over the Arctic Ocean. *Sci. Total Environ.* **853**, 158645 (2022).

83. Pisso, I. et al. The Lagrangian particle dispersion model FLEXPART version 10.4. *Geosci. Model Dev.* **12**, 4955–4997 (2019).

84. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).

85. Lundberg, S. M. & Lee, S.-I. A Unified Approach to Interpreting Model Predictions. in *Adv Neural Inf Process Syst* (eds. Guyon, I. et al.) 30 (Curran Associates, Inc, 2017).

86. Grannas, A. M., Shepson, P. B. & Filley, T. R. Photochemistry and nature of organic matter in Arctic and Antarctic snow. *Glob. Biogeochem. Cycles* **18**, n/a–n/a (2004).

87. Grannas, A. M., Martin, C. B., Chin, Y.-P. & Platz, M. Hydroxyl Radical Production from Irradiated Arctic Dissolved Organic Matter. *Biogeochemistry* **78**, 51–66 (2006).

88. Maturilli, M., Herber, A. & König-Langlo, G. Surface radiation climatology for Ny-Ålesund, Svalbard (78.9°N), basic observations for trend detection. *Theor. Appl. Climatol.* **120**, 331–339 (2015).

89. Stohl, A. Characteristics of atmospheric transport into the Arctic troposphere. *J. Geophys. Res.* **111**, D11306 (2006).

90. Bozem, H. et al. Characterization of transport regimes and the polar dome during Arctic spring and summer using in situ aircraft measurements. *Atmos. Chem. Phys.* **19**, 15049–15071 (2019).

91. Freud, E. et al. Pan-Arctic aerosol number size distributions: seasonality and transport patterns. *Atmos. Chem. Phys.* **17**, 8101–8128 (2017).

92. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (Association for Computing Machinery, New York, NY, USA, 2016). <https://doi.org/10.1145/2939672.2939785>.

93. Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. Optuna: A Next-Generation Hyperparameter Optimization Framework. in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* 2623–2631 (Association for Computing Machinery, New York, NY, USA, 2019). <https://doi.org/10.1145/3292500.3330701>.

94. Flora, M. L., Potvin, C. K., McGovern, A. & Handler, S. A Machine Learning Explainability Tutorial for Atmospheric Sciences. *Artif. I. earth syst.* **3**, e230018 (2024).

95. Nielsen, I. E. et al. Biogenic and anthropogenic sources of aerosols at the High Arctic site Villum Research Station. *Atmos. Chem. Phys.* **19**, 10239–10256 (2019).

96. Jiang, Bss et al. Importance of Atmospheric Transport on Methanesulfonic Acid (MSA) Concentrations in the Arctic Ocean During Summer Under Global Warming. *J. Geophys. Res. Atmos.* **128**, e2022JD037271 (2023).

97. Seinfeld, J. H. & Pandis, S. N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. (John Wiley & Sons, 2016).

98. Sunda, W., Kieber, D. J., Kiene, R. P. & Huntsman, S. An antioxidant function for DMSP and DMS in marine algae. *Nature* **418**, 317–320 (2002).

99. Woolf, D. K. Bubbles and their role in gas exchange. in *The Sea Surface and Global Change* (eds. Liss, P. S. & Duce, R. A.) 173–206 (Cambridge University Press, Cambridge, 1997). <https://doi.org/10.1017/CBO9780511525025.007>.

100. Nightingale, P. D. et al. In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers. *Glob. Biogeochem. Cycles* **14**, 373–387 (2000).

101. Galí, M., Devred, E., Babin, M. & Levasseur, M. Decadal increase in Arctic dimethylsulfide emission. *Proc. Natl Acad. Sci. USA* **116**, 19311–19317 (2019).

102. Shen, J. et al. High Gas-Phase Methanesulfonic Acid Production in the OH-Initiated Oxidation of Dimethyl Sulfide at Low Temperatures. *Environ. Sci. Technol.* **56**, 13931–13944 (2022).

103. Devasthale, A. & Thomas, M. A. Sensitivity of Cloud Liquid Water Content Estimates to the Temperature-Dependent Thermodynamic Phase: A Global Study Using CloudSat Data. *J. Clim.* **25**, 7297–7307 (2012).

104. Shupe, M. D. & Intrieri, J. M. Cloud Radiative Forcing of the Arctic Surface: The Influence of Cloud Properties, Surface Albedo, and Solar Zenith Angle. *J. Clim.* **17**, 616–628 (2004).

105. Lannuzel, D. et al. The future of Arctic sea-ice biogeochemistry and ice-associated ecosystems. *Nat. Clim. Chang.* **10**, 983–992 (2020).

106. Gradinger, R. Integrated abundance and biomass of sympagic meiobenthos in Arctic and Antarctic pack ice. *Polar Biol.* **22**, 169–177 (1999).

107. Zhu, L., Nicovich, J. M. & Wine, P. H. Temperature-dependent kinetics studies of aqueous phase reactions of hydroxyl radicals with dimethylsulfoxide, dimethylsulfone, and methanesulfonate. *Aquat. Sci.* **65**, 425–435 (2003).

108. Gu, M., Wang, Z., Wei, J. & Yu, X. An assessment of Arctic cloud water paths in atmospheric reanalyses. *Acta Oceanol. Sin.* **40**, 46–57 (2021).

109. Zhang, J. et al. Evaluation of Surface Relative Humidity in China from the CRA-40 and Current Reanalyses. *Adv. Atmos. Sci.* **38**, 1958–1976 (2021).

110. Pernov, J. B., Gros-Daillon, J. & Schmale, J. Comparison of selected surface level ERA5 variables against in situ observations in the continental Arctic. *Q. J. R. Meteorol. Soc.* 1–24 (2024) <https://doi.org/10.1002/qj.4700>.

111. Johnson, J. S. & Jen, C. N. Role of Methanesulfonic Acid in Sulfuric Acid–Amine and Ammonia New Particle Formation. *ACS Earth Space Chem.* **7**, 653–660 (2023).

112. He, X.-C. et al. Iodine oxoacids enhance nucleation of sulfuric acid particles in the atmosphere. *Science* **382**, 1308–1314 (2023).

113. Zhang, R., Shen, J., Xie, H.-B., Chen, J. & Elm, J. The role of organic acids in new particle formation from methanesulfonic acid and methylamine. *Atmos. Chem. Phys.* **22**, 2639–2650 (2022).

114. Hodshire, A. L. et al. The potential role of methanesulfonic acid (MSA) in aerosol formation and growth and the associated radiative forcings. *Atmos. Chem. Phys.* **19**, 3137–3160 (2019).

115. Rosati, B. et al. Hygroscopicity and CCN potential of DMS-derived aerosol particles. *Atmos. Chem. Phys.* **22**, 13449–13466 (2022).

116. Gunsch, M. J. et al. Contributions of transported Prudhoe Bay oil field emissions to the aerosol population in Utqiagvik, Alaska. *Atmos. Chem. Phys.* **17**, 10879–10892 (2017).

117. Kirpes, R. M. et al. Emerging investigator series: influence of marine emissions and atmospheric processing on individual particle composition of summertime Arctic aerosol over the Bering Strait and Chukchi Sea. *Environ. Sci.: Process. Impacts* **22**, 1201–1213 (2020).

118. Moschos, V. et al. Elucidating the present-day chemical composition, seasonality and source regions of climate-relevant aerosols across the Arctic land surface. *Environ. Res. Lett.* **17**, 034032 (2022).

119. Moschos, V. et al. Equal abundance of summertime natural and wintertime anthropogenic Arctic organic aerosols. *Nat. Geosci.* **15**, 196–202 (2022).

120. Huang, J. & Jaeglé, L. Wintertime enhancements of sea salt aerosol in polar regions consistent with a sea ice source from blowing snow. *Atmos. Chem. Phys.* **17**, 3699–3712 (2017).

121. Arrigo, K. R. & Dijken, G. Lvan Continued increases in Arctic Ocean primary production. *Prog. Oceanogr.* **136**, 60–70 (2015).

122. Lewis, K. M., Dijken, G. Lvan & Arrigo, K. R. Changes in phytoplankton concentration now drive increased Arctic Ocean primary production. *Science* **369**, 198–202 (2020).

123. Renaut, S., Devred, E. & Babin, M. Northward Expansion and Intensification of Phytoplankton Growth During the Early Ice-Free Season in Arctic. *Geophys. Res. Lett.* **45**, 10,590–10,598 (2018).

124. Qu, B., Gabric, A. J. & Jackson, R. Contemporary variability in dimethylsulfide flux in the Barents Sea and simulated change under 4 \times CO₂ climate conditions. *J. Mar. Syst.* **220**, 103573 (2021).

125. Cooper, O. R. et al. Multi-decadal surface ozone trends at globally distributed remote locations. *Elem. Sci. Anth.* **8**, 23 (2020).

126. Law, K. S. et al. Arctic Tropospheric Ozone Trends. *Geophys. Res. Lett.* **50**, e2023GL103096 (2023).

127. Bougoudis, I. et al. Long-term time series of Arctic tropospheric BrO derived from UV-VIS satellite remote sensing and its relation to first-year sea ice. *Atmos. Chem. Phys.* **20**, 11869–11892 (2020).

128. Confer, K. L. et al. Impact of Changing Arctic Sea Ice Extent, Sea Ice Age, and Snow Depth on Sea Salt Aerosol From Blowing Snow and the Open Ocean for 1980–2017. *J. Geophys. Res. Atmos.* **128**, e2022JD037667 (2023).

129. Stevenson, D. S. et al. Trends in global tropospheric hydroxyl radical and methane lifetime since 1850 from AerChemMIP. *Atmos. Chem. Phys.* **20**, 12905–12920 (2020).

130. Heslin-Rees, D. et al. From a polar to a marine environment: has the changing Arctic led to a shift in aerosol light scattering properties? *Atmos. Chem. Phys.* **20**, 13671–13686 (2020).

131. Bintanja, R. & Andry, O. Towards a rain-dominated Arctic. *Nat. Clim. Chang.* **7**, 263–267 (2017).

132. Bintanja, R. The impact of Arctic warming on increased rainfall. *Sci. Rep.* **8**, 16001 (2018).

133. Pithan, F. & Jung, T. Arctic Amplification of Precipitation Changes—The Energy Hypothesis. *Geophys. Res. Lett.* **48**, e2021GL094977 (2021).

134. Bintanja, R. et al. Strong future increases in Arctic precipitation variability linked to poleward moisture transport. *Sci. Adv.* **6**, eaax6869 (2020).

135. Law, K. S. et al. Local Arctic air pollution: Sources and impacts. *Ambio* **46**, 453–463 (2017).

136. Schmale, J. et al. Local Arctic Air Pollution: A Neglected but Serious Problem. *Earth's Future* **6**, 1385–1412 (2018).

137. Corbett, J. J. et al. Arctic shipping emissions inventories and future scenarios. *Atmos. Chem. Phys.* **10**, 9689–9704 (2010).

138. Gilgen, A., Huang, W. T. K., Ickes, L., Neubauer, D. & Lohmann, U. How important are future marine and shipping aerosol emissions in a warming Arctic summer and autumn? *Atmos. Chem. Phys.* **18**, 10521–10555 (2018).

139. Peters, G. P. et al. Future emissions from shipping and petroleum activities in the Arctic. *Atmos. Chem. Phys.* **11**, 5305–5320 (2011).

140. Dalsøren, S. B. et al. Environmental impacts of shipping in 2030 with a particular focus on the Arctic region. *Atmos. Chem. Phys.* **13**, 1941–1955 (2013).

141. Xu, J. et al. An interlaboratory comparison of aerosol inorganic ion measurements by ion chromatography: implications for aerosol pH estimate. *Atmos. Meas. Tech.* **13**, 6325–6341 (2020).

142. Barrie, L. A., Hoff, R. M. & Daggupaty, S. M. The influence of mid-latitudinal pollution sources on haze in the Canadian arctic. *Atmos. Environ.* **15**, 1407–1419 (1981).

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

144. Barrie, L. A., den Hartog, G., Bottenheim, J. W. & Landsberger, S. Anthropogenic aerosols and gases in the lower troposphere at Alert Canada in April 1986. *J. Atmos. Chem.* **9**, 101–127 (1989).

145. Maturilli, M., Herber, A. & König-Langlo, G. Climatology and time series of surface meteorology in Ny-Ålesund, Svalbard. *Earth Syst. Sci. Data* **5**, 155–163 (2013).

146. Muscari, G. et al. Observations of surface radiation and stratospheric processes at Thule Air Base, Greenland, during the IPY. *Ann. Geophys.* **57**, SS0323 (2014).

147. Becagli, S. et al. MBAS (Methylene Blue Active Substances) and LAS (Linear Alkylbenzene Sulphonates) in Mediterranean coastal aerosols: Sources and transport processes. *Atmos. Environ.* **45**, 6788–6801 (2011).

148. Kolesar, K. R. et al. Effect of Prudhoe Bay emissions on atmospheric aerosol growth events observed in Utqiagvik (Barrow), Alaska. *Atmos. Environ.* **152**, 146–155 (2017).

149. Quinn, P. K. et al. Surface submicron aerosol chemical composition: What fraction is not sulfate? *J. Geophys. Res. Atmos.* **105**, 6785–6805 (2000).

150. Quinn, P. K. et al. A 3-year record of simultaneously measured aerosol chemical and optical properties at Barrow, Alaska. *J. Geophys. Res. Atmos.* **107**, AAC 8–1 (2002).

151. Quinn, P. K., Bates, T. S., Schulz, K. & Shaw, G. E. Decadal trends in aerosol chemical composition at Barrow, Alaska: 1976–2008. *Atmos. Chem. Phys.* **9**, 8883–8888 (2009).

152. Graham, R. M., Hudson, S. R. & Maturilli, M. Improved Performance of ERA5 in Arctic Gateway Relative to Four Global Atmospheric Reanalyses. *Geophys. Res. Lett.* **46**, 6138–6147 (2019).

153. Adrian S. et al. Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5.1. *ECMWF Technical Memoranda* (2020) <https://doi.org/10.21957/rcxqfm0g>.

154. Bell, B. et al. The ERA5 global reanalysis: Preliminary extension to 1950. *Q. J. R. Meteorol. Soc.* **147**, 4186–4227 (2021).

155. Stohl, A., Hittenberger, M. & Wotawa, G. Validation of the lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data. *Atmos. Environ.* **32**, 4245–4264 (1998).

156. Brioude, J. et al. The Lagrangian particle dispersion model FLEXPART-WRF version 3.1. *Geosci. Model Dev.* **6**, 1889–1904 (2013).

157. Tunved, P., Ström, J. & 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 (2013).

158. Cheng, M. D. et al. Qualitative determination of source regions of aerosol in Canadian high Arctic. *Environ. Sci. Technol.* **27**, 2063–2071 (1993).

159. Hopke, P. K. et al. Possible sources and preferred pathways for biogenic and non-sea-salt sulfur for the high Arctic. *J. Geophys. Res. Atmos.* **100**, 16595–16603 (1995).

160. Polissar, A. V., Hopke, P. K. & Harris, J. M. Source Regions for Atmospheric Aerosol Measured at Barrow, Alaska. *Environ. Sci. Technol.* **35**, 4214–4226 (2001).

161. Cheng, M.-D. Geolocating Russian sources for Arctic black carbon. *Atmos. Environ.* **92**, 398–410 (2014).

162. Lee, H. et al. Atmospheric new particle formation characteristics in the Arctic as measured at Mount Zeppelin, Svalbard, from 2016 to 2018. *Atmos. Chem. Phys.* **20**, 13425–13441 (2020).

163. Kuttippurath, J., Patel, V. K., Roy, R. & Kumar, P. Sources, variability, long-term trends, and radiative forcing of aerosols in the Arctic: implications for Arctic amplification. *Environ. Sci. Pollut. Res* **31**, 1621–1636 (2024).

164. Bressi, M. et al. Sources and geographical origins of fine aerosols in Paris (France). *Atmos. Chem. Phys.* **14**, 8813–8839 (2014).

165. Seibert, P. et al. Trajectory analysis of aerosol measurements at high alpine sites. *Air Pollut. Modeling Its Appl. X. Nato. Chall. Mod. Soc.* **18**, 689–693 (1994).

166. Hsu, Y.-K., Holsen, T. M. & Hopke, P. K. Comparison of hybrid receptor models to locate PCB sources in Chicago. *Atmos. Environ.* **37**, 545–562 (2003).

167. Bergstra, J., Bardenet, R., Bengio, Y. & Kégl, B. Algorithms for Hyper-Parameter Optimization. In *Advances in Neural Information Processing Systems* (eds. Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F. & Weinberger, K. Q.) 24 (Curran Associates, Inc, 2011).

168. Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A. & Talwalkar, A. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *J. Mach. Learn. Res.* **18**, 1–52 (2018).

169. Shapley, L. S. 17. A Value for n-Person Games. In *Contributions to the Theory of Games (AM-28), Volume II* (eds. Kuhn, H. W. & Tucker, A. W.) 307–318 (Princeton University Press, Princeton, 1953). <https://doi.org/10.1515/9781400881970-018>.

170. Molnar, C. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable (2nd ed.). christophm.github.io/interpretable-ml-book/ (2022).

171. Lundberg, S. M., Erion, G. G. & Lee, S.-I. Consistent individualized feature attribution for tree ensembles. arXiv preprint arXiv:1802.03888 (2018).

172. Mann, H. B. Nonparametric Tests Against Trend. *Econometrica* **13**, 245–259 (1945).

173. Kendall, M. G. *Rank Correlation Methods*. (Griffin, Oxford, England, 1948).

174. Theil, H. A rank-invariant method of linear and polynomial regression analysis. *Nederl. Akad. Wetensch., Proc.* **53**, 386–392 (1950).

175. Sen, P. K. Estimates of the Regression Coefficient Based on Kendall's Tau. *J. Am. Stat. Assoc.* **63**, 1379–1389 (1968).

Acknowledgements

This research was funded by the Swiss Data Science Center project C20-01 Arctic climate change: exploring the Natural Aerosol baseline for improved model Predictions (ArcticNAP). J.S. holds the Ingvar Kamprad Chair for Extreme Environment Research sponsored by Ferring Pharmaceuticals. This project received funding from the Ingvar Kamprad Chair funded by Ferring Pharmaceuticals. This is a PMEL contribution 5608 and CICOES publication number 2024-1341. We would like to thank Environment and Climate Change Canada, CCMR, CRD, Sangeeta Sharma, technicians (Joe Kovalik, Armand Gaudenzi, Dave Halpin, Dan Veber, D. Toom, and Alina Chivulescu), students, and operators at the Alert station who assisted in the filter collection, processing, and laboratory analyses of methanesulfonic acid. The Department of National Defense and Canadian Armed Forces staff

is acknowledged for providing support at the Canadian Forces Station Alert. We would also like to thank the technicians, support staff, students, and scientists who helped obtain, analyze, and process the MSA measurements at Thule, Gruvebadet, and Utqiagvik/Barrow. This research was partially funded by the Italian Ministry of University and Research (MIUR) within the framework of the projects: Dirigibile Italia: A platform for a multidisciplinary study on climatic changes in the Arctic region and their influence on temperate latitudes (PRIN-2007); ARCTICA-ARCTic Research on the Inter-connections between Climate and Atmosphere (PRIN 2009); Observations of changes in chemical composition and physical properties of Polar Atmospheres from NDACC Stations (PNRA 2010–2012); ARCA-Arctic: present climatic change and past extreme events (MIUR 2014–2016); SVAAP—The study of the water vapor in the polar atmosphere (PNRA 2015–2016); OASIS-YOPP—Observations of the Arctic Stratosphere In Support of YOPP (PNRA 2016–2018).

Author contributions

Jakob Boyd Pernov—conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft, writing—review and editing, visualization, supervision, project administration. Eliza Harris—conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing—review and editing. Michele Volpi—conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing—review and editing. Tamara Baumgartner—methodology, software, validation, formal analysis, investigation, writing—review and editing. Benjamin Hohermuth—conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing—review and editing. Stephan Henne—methodology, software, writing—review and editing. William H. Aeberhard—conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing—review and editing. Julia Schmale—conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing—review and editing, funding acquisition. Silvia Becagli, Patricia K. Quinn, Rita Traversi, Lucia M. Upchurch—methodology, validation, resources, data curation, writing—review and

editing, funding acquisition. Julia Schmale—conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing—review and editing, funding acquisition.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41612-024-00712-3>.

Correspondence and requests for materials should be addressed to Jakob Boyd Pernov or Julia Schmale.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024