Simulating the radiative forcing of oceanic dimethylsulfide (DMS) in Asia based on machine learning estimates

Junri Zhao[1,2], Weichun Ma[1,3], Kelsey R Bilsback[4,7], Jeffrey R Pierce[4], Shengqian Zhou[1,2], Ying Chen[1,2], Guipeng yang[5], Yan Zhang[1,2,3,6*]

[1]Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP3), Department of Environmental Science and Engineering, Fudan University, Shanghai 200438, China
[2]Shanghai Institute of Eco-Chongming (SIEC), Shanghai 200062, China
[3]Institute of Digitalized Sustainable Transformation,Big Data Institute, Fudan University, Shanghai 200433, China
[4]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, United States of America
[5]Key Laboratory of Marine Chemistry Theory and Technology, Ministry of Education, College of Chemistry and Chemical Engineering, Ocean University of China, Qingdao 266100, China
[6]National Observations and Research Station for Wetland Ecosystems of the Yangtze Estuary,Shanghai, China
[7]PSE Healthy Energy, Oakland, CA, United States of America

Correspondence to: Yan Zhang (yan_zhang@fudan.edu.cn)

Abstract. DMS emitted from sea water is a key precursor to new particle formation and acts as a regulator in Earth's warming climate system. However, DMS's effects are not well understood in various ocean regions. In this study, we estimated DMS emissions based on a machine learning method and used the GEOS-Chem global 3D chemical transport model coupled with the TwO Moment Aerosol Sectional (TOMAS) microphysics scheme to simulate the atmospheric chemistry and radiative effects of DMS. The contributions of DMS to atmospheric SO_4^{2-} aerosol and cloud condensation nuclei (CCN) concentrations along with the radiative effects over the Asian region were evaluated for the first time. First, we constructed novel monthly-resolved DMS emissions ($0.5° \times 0.5°$) for the year 2017 using a machine learning model. 4351 seawater DMS measurements (including the recent measurements made over the Chinese Seas) and 12 relevant environment parameters were selected for model training. We found the model could predict the observed DMS concentrations with a correlation coefficient of 0.75 and fill the values in regions lacking observations. Across the Asian Seas, the highest seasonal mean DMS concentration occurred in Mar-Apr-May (MAM), and we estimate annual DMS emission flux of 1.25 Tg (S), which equivalent to 15.4% of anthropogenic sulfur emissions over the entire simulation domain (which covered most of Asia) in 2017. The model estimates of DMS and methane sulfonic acid (MSA), using updated DMS emissions, were evaluated by comparing with cruise survey experiments and long-term online measurement site data. The improvement in model performance can be observed compared with simulation results derived from the global-database DMS emissions. The relative contributions of DMS to SO_4^{2-} and CCN were higher in remote oceanic areas, contributing 88% and 42% of all sources, respectively. Correspondingly, the sulfate direct radiative forcing (DRF) and indirect radiative forcing (IRF) contributed by DMS ranged from -200 to -20 mW m^{-2} and -900 to -100 mW m^{-2}, respectively, with levels varying by season. The strong negative IRF is

mainly over remote ocean regions (-900 to -600 mW m^{-2}). Generally, the magnitude of IRF derived by DMS was twice as large as its DRF. This work provides insights into the source strength of DMS, the impact of DMS on climate, and addresses knowledge gaps related to factors controlling aerosols in the marine boundary layer and their climate impacts.

1 Introduction

Ocean-emitted DMS is a precursor of non-sea-salt SO_4^{2-} and controls the composition, size distribution, and number concentration of aerosols over the remote oceanic areas. SO_4^{2-} directly influences the climate system directly by reflecting solar radiation back into the space and indirectly by acting as CCN and altering the albedo of clouds and changing cloud radiative properties (Andreae and Rosenfeld, 2008). The "CLAW" hypothesis proposed by Charlson et al. (1987) assumed that negative feedback interactions between ocean plankton and climate system, where the Earth system acted to buffer itself from warming, was linked through DMS production. Thereafter, several studies found that significant impacts of DMS induced aerosols on CCN and cloud albedos in remote oceans (Park et al., 2017;Quinn et al., 2017;Kulmala et al., 2014;Vallina and Simó, 2007), which lend credence to CLAW hypothesis. Nevertheless, due to low sensitivity of each step of the interactions to changes in force factors in the CLAW climate feedback loop (e.g., low sensitivity of DMS production to changes in incident solar radiation), Quinn and Bates (2011) disproved the hypothesis. Whether the CLAW climate feedback is positive or negative is still uncertain and further research is need to quantify the climate effects of DMS.

Building an accurate emission inventory is key to simulating the climate effects of DMS. As many previous studies have shown (Chen et al., 2018;Hodshire et al., 2019;Rap et al., 2013;Yang et al., 2017;Zhao et al., 2021), the marine DMS emissions used in numerical models are mainly estimated using an interpolation scheme (Kettle et al., 1999;Lana et al., 2011), which estimates DMS climatology by interpolating observed DMS data at limited sites to the global ocean. Previously, observations from Global Surface Seawater DMS Database have been grouped into 57 ecological geographic ocean provinces and weighted interpolations from nearby provinces have been used to fill the values without observations. Wang et al., (2020) has pointed out that there are uncertainties in using spatial and temporal averaged data to fill regions without observations. However, artificial neural networks can potentially be trained and used to fill measurement gaps (Wang et al., 2020). Galí et al. (2018) created a remote sensing algorithm to estimate DMS concentrations which is based on the relationship between a precursor of DMS and plankton light exposure. Their results (Galí et al., 2018) indicated that the remote sensing algorithms have better ability to reproduce the climatological features of DMS seasonality than interpolated DMS climatologies, which also outweigh the disadvantage of the interpolation scheme used in previous study (Lana et al., 2011). In a recent study (Bell et al., 2021), long-term in-situ DMS measurements conducted in North Atlantic Ocean from 2015 to 2018 were compared with the interpolated DMS climatologies (Lana et al., 2011), predicted DMS concentrations from the remote sensing algorithm (Galí et al., 2018), and a neural network approach (Wang et al., 2020). The analysis revealed that that both the remote sensing algorithm and the neural network model were better able to reproduce the sea water DMS trends better than the interpolated

65 climatologies. However, DMS predictions from two of the models (Galí et al., 2018;Wang et al., 2020) underpredicted DMS concentrations, likely because the primary biological processes of DMS production was not accounted for(Bell et al., 2021).

There are several modelling studies which have quantified the aerosol direct and indirect radiative forcing of DMS on a global scale. The global annual mean DMS aerosol indirect radiative forcing estimates have ranged from -6.55 to -0.23 W/m^2 in previous studies (Mahajan et al., 2015;Thomas et al., 2010;Rap et al., 2013;Yang et al., 2017;Jin et al., 2018). However,

70 there have been few studies that have reported the radiative effect of DMS on a regional scale. Choi et al. (2020) adopted an empirical algorithm to estimate DMS concentrations and calculated the direct radiative effect of DMS aerosol to be -1.3 W/m^2 for the year 2014-2016 over East Asian seas, which was higher than the global average results (Yang et al., 2017;Rap et al., 2013). There were no evaluations of the DMS predictions in the seawater or atmosphere in these studies, leading to an unknown reliability of the results. The annual-mean direct radiative forcing due to DMS produced aerosol were -0.2 to -0.1 W/m^2 over

75 East Asia reported by Li et al. (2019) who used a DMS climatology(Lana et al., 2011) with $1° \times 1°$ horizontal resolution for radiative forcing calculation. As mentioned before, some uncertainties in the DMS climatology estimated by an interpolation scheme and coarse grid ($1° \times 1°$) may not be appropriate for regional simulations. In the previous studies, (Li et al., 2020b;Li et al., 2020a) used long-term DMS measurements in 2011, 2013, 2015, 2016, and 2017 from a series of shipboard field experiments and performed interpolation to map DMS concentrations in Chinese Seas. The newest DMS measurements were

80 used to explore the impact of DMS on air quality over coastal areas of China, but the radiative effect of DMS was not reported.

To our knowledge, this is the first systematic study of the Asia region that quantifies the impacts of DMS on sulfate, particle number concentration, and radiative forcing by using state-of-the-art aerosol microphysics model coupled on global 3D chemical transport model. ~~In this study, we first trained 12 environmental parameters and newly updated DMS measurements by utilizing eXtreme Gradient Boosting (XGBoost) machine learning algorithms (Chen and Guestrin, 2016)~~

85 ~~used to construct a new regional DMS emissions for the year 2017 with a plausible underlying basis in ocean environmental parameters.~~ In this study, we developed the regional DMS emissions for the year 2017 by training eXtreme Gradient Boosting (XGBoost) machine learning algorithms (Chen and Guestrin, 2016) combined with a newly updated dataset. Then, the model estimates of DMS and MSA were evaluated by comparing the model simulations with shipboard field measurements and long-term online measurement site data. Finally, the annual-average and seasonal impacts of DMS on sulfate/CCN concentrations

90 and direct/indirect radiative forcing were quantified.

2 Methods and data

2.1 GEOS-Chem-TOMAS

In this study, the GEOS-Chem version 12.9.3 (https://doi.org/10.5281/zenodo.3974569, last access: 25 March 2021)
95 coupled with the online TOMAS aerosol microphysics model (Adams and Seinfeld, 2002) was adopted to calculate atmospheric aerosol size, number, and mass concentrations from marine DMS emissions. TOMAS was used to simulate

aerosol microphysics processes (i.e., nucleation, coagulation, condensation, cloud processing). The advantage of TOMAS is the full aerosol size resolution for all chemical species and the conservation of aerosol number, which allows modelers to construct aerosol and CCN number budgets that balance. GEOS-Chem-TOMAS (GC-TOMAS) has been used in a range of previous studies (Kodros and Pierce, 2017;Pierce and Adams, 2006;Kodros et al., 2016;D'Andrea et al., 2013;Westervelt et al., 2013;Lee et al., 2009;Trivitayanurak et al., 2008;Pierce et al., 2007;Adams and Seinfeld, 2002;Jathar et al., 2020). The model contains detailed hydrocarbon – nitrogen oxide (NO_x) – ozone(O_3) - volatile organic compounds(VOC) – bromine oxides(BrO_x) tropospheric chemistry (Bey et al., 2001) and aerosol species (including sulfate, nitrate, ammonium, black carbon, organic carbon, mineral dust, and sea salt) (Duncan Fairlie et al., 2007;Pye et al., 2009;Alexander et al., 2005;Park et al., 2004) that are fully coupled to gas-phase chemistry, with the ISORROPIA II algorithm to calculate the thermodynamic equilibrium between aerosols and their gas phase precursors (Fountoukis and Nenes, 2007). The model includes detailed wet and dry deposition scheme for aerosols and gas species which have been described in previous studies (Wesely, 2007;Liu et al., 2001;Wang et al., 1998;Amos et al., 2012). This version of GC-TOMAS tracks the total aerosol particle number and the mass of each aerosol species (sulfate, mineral dust, sea salt, hydrophilic and hydrophobic organic carbon, externally and internally mixed elemental carbon, and aerosol water) across 15 logarithmically size bins ranging from 3 nm to 10 μm (Lee and Adams, 2012;Lee et al., 2013). Since the ammonium nitrate size distribution is not explicitly tracked with GC-TOMAS, so we assume that it follows the aerosol water distribution(Bilsback et al., 2020a;Bilsback et al., 2020b).

The simulation domain covers most of Asia (11°S to 55°N, 60–150°E), was discretized with a horizontal grid resolution of 0.5° × 0.625° and 47 vertical layers, and uses Modern-Era Retrospective Analysis for Research and Applications Version (MERRA-2) assimilated meteorological field for meteorological inputs (Gelaro et al., 2017). To assess radiative impacts of DMS emissions at a regional scale, we performed three different annual simulations for the year 2017 (Table 1). The "XG" simulation represents DMS emissions that were calculated from our updated DMS emissions estimates (see Section 2.3) and the "LANA" simulation refer to DMS emissions from Lana DMS climatology (Lana et al., 2011), which is default setting in current version of Geos-Chem model. The "ND" simulation has DMS emissions tuned off. Each simulation was conducted with 1 month spin-up period (December 2016). The boundary conditions for the simulation domain were obtained from global simulations at 2° × 2.5° with 47 vertical layers.

For anthropogenic emissions in Asia, we used the recently updated Global anthropogenic emission inventories (0.5°x0.5°) or the year 2017 from the open-source Community Emissions Data System (CEDS) (McDuffie et al., 2020), which applied scale factors from Zheng et al., (2018) to update China's emissions for the year 2017. Since there is a significant reduction (62%) in SO_2 emissions in China from 2010 to 2017 (Zheng et al., 2018), updated emissions for China are crucial for quantifying contributions of biogenic sulfur source over Asia. Biomass burning emissions in the GC-TOMAS are obtained from Global Fire Emissions Database Version 4 (van der Werf et al., 2017). Dust, biogenic VOCs, sea salt, soil NO_x, and lighting NO_x emissions are calculated online based on MERRA-2 meteorological field. The Dust Entrainment and Deposition (DEAD) scheme from Zender et al., (2003) was implemented GEOS-Chem to simulate dust mobilization. The Model of Emissions of Gases and Aerosols from Nature from Guenther et al., (2012) was used to generate biogenic VOCs emissions.

Soil and lighting NO$_x$ emissions are calculated by parameterization scheme described in Hudman et al., (2012) and Price and Rind (1992), respectively.

The sea-air flux of DMS is estimated using the following the empirical formula as described in (Lana et al., 2011):

$$F = C_w \times k_w \times (1-\gamma) \tag{1}$$

Where, C_w is the seawater DMS concentrations and k_w is the water side gas transfer velocity and γ is the atmospheric gradient fraction. In this study, we selected the Nightingale et al. (2000) parameterization (hereafter N00) for k_w to represent the DMS emissions over the global ocean.

2.2 Radiative forcing calculation scheme

To calculate the top-of-atmosphere (TOA) all sky DRF and cloud-albedo IRF, we used Rapid Radiative Transfer Model for Global Climate Models (RRTMG) (Iacono et al., 2008) with monthly averaged aerosol number and mass concentrations from GC-TOMAS output and meteorological variables from MERRA2. For the DRE, we calculated aerosol optical depth (AOD) single scattering albedo, and the asymmetry parameter based on Mie theory (Bohren and Huffman, 1983) and refractive indices from the Global Aerosol Database (Koepke et al., 1997). In all cases, the DRE was calculated for core-shell optical assumption, where, for each aerosol size bin, black carbon was represented as a spherical core within a homogenous shell of all other hydrophilic species. For the cloud-albedo IRF, we calculate cloud droplet number concentration (CDNC) using the activation parameterization from Abdul-Razzak and Ghan (2002). Cloud-liquid water content is prescribed from MERRA-2 and held fixed, and hence we only calculated the cloud-albedo (Twomey) indirect effect. The changes in effective cloud drop radii were estimated following the cloud-droplet-radius perturbation method used in previous studies (Rap et al., 2013;Kodros et al., 2016;Scott et al., 2014). Then, RRTMG was used to calculate the changes of TOA radiative flux from the changes effect cloud drop radii. We limited this calculation to liquid clouds, which is a limitation in this method. More detailed information about implementation of RRTMG in GC-TOMAS can be found in Kodros et al., (2016).

2.3 Machine learning estimates of sea-surface DMS concentration for calculating DMS emission flux

To better capture the nonlinear relationship between DMS and its influencing parameters, we trained an XGBoost model (machine learning algorithm under the Gradient Boosting framework) with the entire dataset to predict sea surface DMS concentrations where without the observations in the place of missing observations. XGBoost (machine learning algorithm under the Gradient Boosting framework) was used due to its many advantages. For example, XGBoost is computationally efficiency, has prediction accuracy, requires less tunning, and is scalable, has been widely used in area of geoscience (Sun et al., 2021;Ivatt and Evans, 2020;Pan, 2018;Qian et al., 2020;Silva et al., 2022;Cao et al., 2021), and generally outperformed other models. Moreover, Xgboost is good for tabular data and does not require large training datasets (Shwartz-Ziv and Armon, 2022). Thus, to better capture the nonlinear relationship between DMS and the parameters that influence it, we trained an

XGBoost model with the entire dataset to predict sea surface DMS concentrations ~~where without the observations~~ in the place of missing observations.

Figure S1 shows the spatial distribution of DMS measurements. The red points (1022 valid measurements) represent local DMS observations dataset (2011, 2013, 2015, 2016, and 2017) across several Chinese Seas from China Ocean University. Details can be found in our previous studies (Yang et al., 2015a;Yang et al., 2014;Yang et al., 2015b;Xu et al., 2021;Zhai et al., 2020;Wu et al., 2020;Jian et al., 2019;Yu et al., 2019;Mao et al., 2021). The blue points (3329 valid measurements) represent the observations from Global Surface Seawater DMS Database (http://saga.pmel.noaa.gov/dms/; last access: 1 May 2021). In total, 12 environmental parameters (Table S1) which strongly affect the growth of phytoplankton and the production of DMS (Wang et al., 2015) were included as predictors in machine learning estimates. Satellite remotely sensed chlorophyll (Chl), photosynthetically available radiation (PAR), particulate inorganic/organic carbon (PIC/POC), and diffuse attenuation coefficient at 490m (kd490) were from MODIS-Aqua products (daily,8-day, and monthly Level 3-binned 4km resolution data). Nutrient data (Silicate, Phosphate, and Nitrate), sea surface temperature (SST), salinity, and dissolved oxygen (DO) were obtained from World Ocean Atlas 2018 (monthly 0.25° and 1° climatology data). Monthly mixed layer depth (MLD) climatology (0.5° × 0.5°) was obtained from Monthly Isopycnal & Mixed-layer Ocean Climatology (MIMOC). Before the implementation of the algorithm to Asia's oceans, we performed a model validation. First, the environmental parameters were matched with DMS measurements according to sampling geographical coordinates and date. Take remotely sensed Chl data, for example, if the daily binned data failed to match the DMS observed data, we used the 8-day binned data to take the place of daily binned data. After the data matching, we then conducted filtering and quality control which followed methods from Wang et al., (2020), the number of data points in the simulation domain was reduced from 4351 to 3748 observation-based datasets for in-situ DMS and matched with environmental parameters. Table S1 has a description of the environmental parameters, sources, and their filtering thresholds. Avoiding the possible large latitudinal and seasonal variation in DMS, the sampling times and geographic coordinates were also included in machine learning estimates. To solve issues in data discontinuity, these datasets were converted to periodic functions as suggested in previous studies (Gade, 2010;Gregor et al., 2017;Wang et al., 2020). To verify the prediction performance of XGBoost model, we divided the datasets into two parts: a validation dataset and a training dataset. Considering that most of the northern part of the simulation domain was land area, we selected the data from 2° latitude bands between 11°S and 30°N as validation datasets (809 points), while the rest of the data was all used as training data (2939 points). As suggested by Wang et al. (2020), the measurement data collected from the same cruise are highly intercorrelated, and using near-neighbor values to predict validation data may cause the model overfit. So, we selected the validation data manually rather than automatically.

Figure 1 displays the validation results for XGBoost model, which reproduced DMS concentrations with high correlation coefficients (R) of 0.75 and low root-mean-square error (RMSE) of 1.97 µmol m^{-3}. The validation statistics are comparable to other studies (R=0.73-0.81 and RMSE=1.92-2.00 µmol m^{-3}) that used nonlinear/multilinear models to predict sea-surface DMS concentrations over the global ocean (Galí et al., 2018;Wang et al., 2020). Model performance for predicting DMS concentration in each season was illustrated Table S2. Predicted DMS concentrations were slightly underestimated in

comparison with validation datasets, with mean bias (MB) of -0.59 to -0.21 μmol m^{-3} and normalized mean bias (NMB) of -19.36 to -6.51% across the four seasons. A lower RMSE of 1.81 μmol m^{-3} was observed in spring. The MB and NMB in spring were smaller than those in other seasons, which indicated that model performed best in spring. Most of available validation datasets were concentrated in the spring (about 67.9%). Thus, the imbalanced data may leaded to less ideal performance in other seasons.

The advantage of utilizing a machine learning method is the ability to capture nonlinear relationships between DMS and its affecting parameters to estimate DMS concentrations with a plausible underlying basis in spatial-temporal variability. A shortcoming of the traditional geographical interpolation method is that relatively sparse data is typically interpolated to the entire ocean, which has been highlighted by previous studies (Galí et al., 2015;Galí et al., 2018;Wang et al., 2020). In this study, the advantage of the machine learning method is also demonstrated by comparing two different model simulations (see Section 3.2). In the implementation phase of the machining learning algorithm to regional ocean, the remotely sensed datasets used to predict DMS concentrations are all from MODIS-Aqua products in 2017, and monthly climatologies were interpolated to the 8-day or monthly periods for the remotely sensed data; then, we trained XGBoost model to obtain grid values that did not have DMS measurements. Finally, estimated DMS concentration were temporally averaged to a seasonal period and spatially binned to 0.5°×0.5° grid for Asia region (see Section 3.1).

Decision-tree-based machine learning model have a high interpretability. The SHapley Additive exPlanation regression (SHAP) (Lundberg et al., 2020) can provide a deeper understanding of model predictions, which allows for individualized feature attribution for every decision. Stirnberg et al. (2021) quantified the impact of various meteorological derivers on PM$_1$ concentrations by using SHAP analysis, and Silva et al. (2022) used SHAP to explore the errors in the prediction of lightning occurrence in a widely used earth system model. In this study, SHAP was applied to investigate the importance of each predictor on model predicted DMS concentrations.

3 Results

3.1 Spatial and temporal patterns of the sea-water DMS

Regional DMS maps for sea surface DMS concentrations predicted by XGBoost in four seasons are displayed in Figure 2. The data show distinct seasonal variations. The highest regional mean DMS concentrations were observed in the MAM, that is 2.52 μmol m^{-3}, which was approximately 1.15, 1.24, and 1.31 times higher than those in Jun-Jul-Aug (JJA), Sep-Oct-Nov (SON), and Dec-Jan-Feb (DJF) (Table S23), respectively. However, according to the previous studies (Lana et al., 2011;Galí et al., 2018;Wang et al., 2020), the highest DMS concentrations usually occurred in JJA, mainly due to adequate solar irradiation and warm temperature being favourable for primary production. We assumed that this difference was caused by the fact that we examined a different statistical region, compared to previous results that were based on global scale estimates. For comparative purposes, we extracted corresponding simulation domain (Figure 2) estimates values from global scale estimates results; they are listed in Table S23. Across the Asian Seas, all the highest seasonal mean DMS concentrations

occurred in MAM, demonstrating that our estimates agreed well with the estimates of 2.21-2.33 µmol m^{-3} reported in previous studies (Wang et al., 2020;Lana et al., 2011). As shown in Figure S2, zonal mean DMS concentrations between 10°S and 30°N latitude areas of simulation domain were higher in MAM than in JJA, but those between the 30°N and 50°N latitude band were higher in JJA than in MAM. As mentioned in Section 2.3, most of ocean area is concentrated in 10°S and 30°N latitude band of the entire simulation domain (11°S to 55°N, 60–150°E), which leads to the highest regional mean DMS concentrations being observed in MAM. This is most likely due to the seasonal variation of solar irradiation, because most of ocean area (11°S to 30°N) in the simulation domain was influenced more by the solar irradiation in the MAM than in JJA. A similar result can be found in monthly Hovmöller diagrams of DMS climatologies, depicted by Galí et al. (2018). Throughout the four seasons, there were some high concentrations of DMS (higher than 4.3 µmol m^{-3}) that appeared in different coastal areas, which is probably relevant to high nutrient and chlorophyll concentrations over the coastal areas. Galí et al. (2015) also found that most of the coastal regions have higher dimethylsulfoniopropionate (DMSP) concentrations compared to the global ocean, and DMS in the sea water was generated from the breakdown of DMSP.

Figure S3 summarizes the ranked mean SHAP values of each predictor across all prediction cases. The line ranges represent interquartile range across the distribution. Larger SHAP value magnitudes are interpreted to as more important for the prediction task as they have larger contribution from that variable to that prediction. In our study, the most important environmental parameter to predict DMS concentrations was Chl, followed by MLD, PAR, POC, and salinity. Above all, the SHAP value of Chl is more than double its value of MLD and PAR, and much larger than all others. This is consistent with known importance of Chl in developing predicting models of surface water DMS concentrations, because of its biogenic origin (Simó and Dachs, 2002;Galí et al., 2015;Wang et al., 2020;Deng et al., 2021).

We calculated regional sea-air DMS fluxes using the N00 gas transfer velocity and DMS concentrations predicted by XGBoost (Figure 3a). We estimated annual DMS emission fluxes of 1.25 Tg (S), which corresponds to 15.4% of the anthropogenic sulfur emissions over the entire simulation domain (covers most of Asia) in 2017. The higher estimated values of DMS fluxes (higher than 250 tonnes (S)/grid) occurred over some coastal waters, which generally agreed well with the estimated sea surface DMS concentration distribution. The highest emission fluxes occurred over the Chinese Seas (reach up to 450 tonnes (S)/grid). These high fluxes can be attributed to local DMS observations dataset in the Chinese Seas (red point in Figure S1) that were included in the machine learning estimates. Our previous studies (Li et al., 2020a;Li et al., 2020b) have reported that DMS emissions fluxes calculated with the local dataset is 3 times higher than default global-database (Lana et al., 2011) over most area of the Chinese Seas. The highest positive changes of DMS emissions fluxes were mainly in the areas of East China Sea (up to 200 tonnes (S)/grid), and some coastal regions (Figure 3b). However, there were more negative changes of DMS emissions fluxes than positive changes in the sea water, which suggested that sea-air DMS flux estimated in this study generally lower than those from Lana et al. (2011). Similar results can be found in Wang et al. (2020).

3.2 Model evaluation

3.2.1 Model performance of DMS and its oxidation product MSA

Modelled atmospheric DMS concentrations were compared to observations from 2017 Cruise Survey Experiment (CSE) 1-3 (Figure 4). Due to the discontinuities in time and gaps in observations, we averaged the whole period of each CSE observation for our comparisons. The results in Table S~~3~~4 demonstrate moderate improvements in the model performance of DMS predictions when using updated DMS emissions relative to default DMS emissions, i.e., the difference between the observations and predictions (observation - prediction) became smaller (from -16.34 to 6.68 pptv for CSE 1, -21.11 to -16.17 pptv for CSE 2, and -121.57 to 117.39 pptv for CSE 3, respectively). CSE 3 had much higher DMS concentrations, because most of the measurements were from the mouth of the Changjiang River, and it is difficult for a coarse model grid ($0.5° \times 0.625°$) to represent the high values that occur off coastal areas. MSA is a tracer of DMS, because it is formed exclusively from DMS (Gondwe et al., 2003). We also evaluated the model performances for MSA by comparing the model simulations with long-term online measurement site data (Zhou et al., 2021) from Hua Niao Island (Figure 4). Figure 5 displays time series of daily mean MSA values of predictions (XG and LANA) and observations. The simulated MSA concentrations from XG and LANA are both within the range of observed values, and the trends of the MSA concentrations were relatively well reproduced, with mean values of 0.014, 0.020, and 0.023 μg m^{-3} for LANA, XG, and observations. ~~During the period of Jun 21 to Jun 25 and Jun 28 to July 3, LANA simulation results were closer to the observations, and XG simulations underpredicted the measurements in those two periods. Overall, the simulation results of XG in other periods were closer to the observations than those of LANA simulation results.~~ Although in some periods, LANA simulation results were closer to the observations, and XG simulations underpredicted, e.g., RMSE of 0.013 and 0.006 μg m^{-3} for LANA, 0.021 and 0.010 μg m^{-3} for XG during the period of Jun 21 to Jun 25 and Jun 28 to July 3, respectively. However, in the whole, the simulation results of XG in other periods were closer to the observations than those of LANA simulation results, with RMSE of 0.024 and 0.018μg m^{-3} for LANA and XG, respectively.

3.2.2 Model performance evaluation for PM$_{2.5}$, AOD, and CCN

The magnitude and distributions of PM$_{2.5}$, AOD, and CCN directly influence DRF and IRF estimates. To evaluate whether GC-TOMAS can reproduce the spatial distribution and temporal trends of these parameters over the simulation area, we evaluated model performance by comparing simulation results for XG with ground observations and satellite-retrieved estimates. Since DMS impacts PM$_{2.5}$ and CCN over the ocean and some coastal areas (see Section 3.3), and the ground observational data is all over land areas, so we only used one of the simulations for model evaluation.

Boylan and Russell (2006) suggested that model predictions can be regarded as sufficiently accurate when the model has a mean fractional bias(MFB) $\leq \pm 30\%$ and mean fractional error (MFE) $\leq \pm 50\%$. Figure S~~3~~4 presents the distributions of simulated annual mean PM$_{2.5}$ concentrations and observations at 366 city sites from China National Environmental Monitoring Center (CNEMC). The model performed well against PM$_{2.5}$ observations for the year 2017, with MFB of 5.5% and MFE of

23.1%, which both within the range suggested by Boylan and Russell (2006), and had a Pearson's correlation coefficient (R) of 0.62. Simulated $PM_{2.5}$ concentrations were slightly underpredicted with a MB of -1.3 μg m^{-3}, which is probably ascribed to underpredict $PM_{2.5}$ in some parts northern China. Further, uncertainties in land-based emission inventories tend to cause 295 different model performance in different regions.

Table S45 summarizes the collected in-situ measurements of CCN concentrations in other previous studies and corresponding annual-mean simulated CCN concentrations which were used for evaluation. The MFB and MFE was 28.17% and 34.16%, which met the suggested benchmark; however, the model estimates underpredicted the measurements in most areas. Liu et al., (2020) adopted a satellite-based method to retrieve CCN concentrations from 2013 to 2019 and reported that 300 they could reasonably reproduce the spatial pattern of CCN in East Asia. In this study, monthly mean GC-TOMAS CCN concentrations were compared to satellite-retrieved CCN concentrations at supersaturation levels of approximately 0.2% from Liu et al., (2020). A total of 8 months of satellite-retrieved CCN concentrations were averaged on the MERRA-2 grid (corresponding 667 simulation grids) for comparison (Figure S45). The simulated CCN concentrations presented generally similar monthly variations when compared to the satellite-retrieved concentrations. The modeled concentrations had a MFB 305 of 17.23% and MFE of 37.28%, both meeting the criteria suggested by Boylan and Russell (2006). The GC-TOMAS CCN concentrations (430 cm^{-3}) for 8 months underestimated the satellite-retrieved concentrations (587 cm^{-3}). This underestimation is more apparent in July, August, September, and November. However, for other months (February, April, May, and June) the simulated CCN concentrations only slightly underpredicted observations with mean bias (MB) of -75 cm^{-3}. This difference is more likely attributable to differences in model performance in different regions. For example, the underpredictions of CCN 310 in May were mainly distributed in eastern coastal area of China, the Korean Peninsula, and Japan. But in August and September, the underprediction of model estimates discrepancies were mainly in the southern and northern part of China, respectively. Due to the limited CCN monitoring data in our domain during simulation period, we compared predicted results with satellite-retrieved CCN. However, as Liu et al., (2020) indicated that errors in retrieved data and the CCN counters might cause inaccuracy of satellite CCN inversion results. Thus, we note that the satellite derived CCN cannot be treated as true as in-situ 315 observations when validating model results.

For AOD, monthly averages from the Aerosol Robotic Network (AERONET) Version 3 spectral deconvolution algorithm (SDA) level 2.0 measurements (Giles et al., 2019) were used to validate the model estimations. In total, there were 79 measurements within the simulation domain. Figure S56 displays annual-mean model estimates and AERONET measurements AOD at 550nm (the AERONET AODs at 500nm are converted to 550nm using Ångström exponents at 500nm). The model 320 estimates compared well with measurements with a Pearson's R of 0.84 and only a slightly underprediction of AOD with MBs of -0.123. The respective MFB and MFE were -28.64% and 13.45%, which all meet the benchmark suggested in Boylan and Russel et al, 2006.

3.3 Seasonal variations of DMS impacts to SO_4^{2-}, CCN, and radiative forcing

325   By updated the DMS emissions in GEOS-Chem (XG-ND), we find an enhancement of near-surface SO_4^{2-} concentrations of 0.1-0.3 μg m^{-3} over most areas of seawater (Figure 6(a)). The highest impacts (approximately 0.3 μg m^{-3}) occurred in MAM around the South China Sea area due to highest regional mean DMS concentrations in MAM. However, the spatial distributions of SO_4^{2-} concentrations enhanced by addition of DMS emissions in the four seasons did not exactly follow the spatial and temporal pattern of seawater DMS concentrations (Figure 2). Sea surface wind speed has noticeable impacts on the sea-air

330   DMS flux and followed atmospheric DMS concentrations, which caused higher atmospheric DMS concentrations over the India Ocean in the MAM. Ambient oxidant level also plays an important role in the subsequent DMS oxidation phase. For example, higher atmospheric DMS (300-400pptv) and SO_2 (0.2-0.3μg m^{-3}) concentrations contributed by DMS can be found around the areas of East China Sea (Figure S6̶7 and S7̶8) in MAM and JJA. However, a higher contribution of DMS emissions to near-surface SO_4^{2-} concentrations occurred over south China Sea in DJF and MAM. The spatial disparities might be due to

335   the roles of oxidants in the conversion of SO_2 into SO_4^{2-} in different seasons. Additionally, cloud cover could also affect aqueous conversion.

The magnitude of the all-sky sulfate DRF at TOA contributed by DMS ranged from -200 to -20 mW m^{-2} in four seasons (Figure 6(b)). The spatial patterns of DRF are highly consistent with those of SO_4^{2-} concentrations, with the stronger negative DRF (-200 to -120 mW m^{-2}) in the areas with higher SO_4^{2-} concentrations contributed by DMS such as the South China Sea,

340   Philippine Sea, and Japan Sea. It should be noted that DRF calculation is from the whole column of the atmosphere whereas Figure 6(a) just shown the surface layer concentrations, yet the spatial results are still qualitatively similar. As reported by some previous studies (Khan et al., 2016;Chen et al., 2018;Zhao et al., 2021), DMS mainly exists in the lower atmosphere, and the impacts of DMS to SO_2 and SO_4^{2-} concentrations are limited to the lower troposphere. So, the magnitude of sulfate DRF at TOA shown in Figure 6(b) is mostly caused by lower-altitude SO_4^{2-} from DMS. SO_4^{2-} aerosols are non-absorbing

345   aerosols that primarily scatter incoming radiation and the increase in reflected solar radiation flux at TOA and almost equally reduce the radiation at the surface (Ramanathan et al., 2001). Thus, for sulfate aerosol, the magnitude of the cooling effect can be estimated from the aerosol radiative forcing at the TOA. The seasonal mean sulfate DRF has a contribution of -22.24, -18.79, -21.58, and -17.43 mW m^{-2} from DMS over the simulation domain in DJF, MAM, JJA, and SON, respectively. The magnitude of the DMS-induced sulfate DRF in DJF and JJA is higher than other seasons, but the highest impacts of DMS

350   emissions on SO_4^{2-} concentrations occurred in MAM followed by DJF. The all-sky DRF was calculated based on the RRTMG model using aerosol mass concentrations (whole column) and optical parameters along with surface albedo and cloud fractions from MERRA-2 assimilated meteorological data. Hence, aerosol mass concentrations as well as other parameters can impact the magnitude and spatial distributions of the DRF. For clear sky condition, aerosol scatter more of incoming solar radiation than in all sky condition, which lead to aerosol DRF at TOA and surface increases compared to all sky conditions.

355   Figure 7(a) shows the changes in seasonal mean CCN surface concentrations at 0.2% supersaturation (CCN (0.2%)) between the XG - ND simulations. Updating the DMS emissions lead to an increase in CCN concentrations by 3 - 42 cm^{-3} over

most areas of seawater and an increase of 6 - 16 cm^{-3} in some coastal regions. The highest increases occurred in DJF, followed by MAM. The impacts of DMS on CCN concentrations are shown in Figure 6(a). The modeled DMS-induced cloud-albedo IRF ranged from -900 to -100 mW m^{-2} in four seasons (Figure 7(b)), which is much higher relative to that of the sulfate DRF attributable to DMS. The seasonal mean sulfate IRF had a contribution of -43.29, -45.04, -43.60, and -33.03 mW m^{-2} from DMS in our domain in DJF, MAM, JJA, and SON, respectively. There are some similarities in the spatial distribution of the effects of DMS on IRF and CCN. However, the strong negative IRF was mainly over remote oceans (-900 to -600 mW m^{-2}), while the higher contributions to CCN were concentrated within coastal waters. One explanation for these differences was that strong anthropogenic emissions in Asia leaded to an intense competition for water vapor during cloud-droplet activation, which further decreased the maximum supersaturation achieved in updrafts and limits droplet activation (Kodros et al., 2016). Also, the clouds are not necessarily at a height where CCN changes are affected by DMS.

3.4 Annual DMS impacts to SO$_4^{2-}$, CCN, and radiative forcing

3.4.1 Annual DMS impacts to SO$_4^{2-}$, CCN, and radiative forcing between XG and ND simulation

Figure 8 (a) shows the annual-mean percent changes and absolute changes in SO$_4^{2-}$ and CCN between the XG and ND simulations. Oceanic DMS emissions increased the near-surface SO$_4^{2-}$ and CCN concentrations by 0.1-0.3 μg m^{-3} and 3 - 42 cm^{-3} over most areas of seawater across the four seasons. Due to heavy amounts of anthropogenic pollutants from the continent, DMS contributed 88% of SO$_4^{2-}$ and 42% of CCN in remote oceanic areas. More than 40% the SO$_4^{2-}$ and 20% of the CCN contributed by DMS emissions were also found in the Philippine Sea and India Ocean, respectively. DMS had a moderate impact of 0.1-0.18 μg m^{-3} for SO$_4^{2-}$ and 10 -22 cm^{-3} for CCN when considering all coastal regions of simulation domain. Yang et al. (2017) indicated that DMS emissions only have 20-40% of contributions to SO$_4^{2-}$ concentrations over downwind ocean areas of East Asia, which was much lower than 40-70% contribution estimated in this study. This discrepancy is mainly ascribed to a significant reduction (62%) in SO$_2$ emissions in China from 2010 to 2017 (Zheng et al., 2018).

The modeled all sky DRF of DMS induced sulfate here range from -100 to -10 mW m^{-2} (Figure 8 (b)). The sulfate DRF was the strongest (-100 to -60 mW m^{-2}) over the South China Sea, which is consistent with the distributions of SO$_4^{2-}$ concentrations contributed by DMS emissions. The DMS induced cloud-albedo IRF (-700 to -100 mW m^{-2}) here was higher than the all sky DRF estimate. A relatively strong cooling IRF (-700 to -400 mW m^{-2}) induced by DMS emissions can be seen in the vicinity of equatorial belt in India Ocean and northwest Pacific Ocean. The simulated annual mean sulfate DRF and IRF is -20.01 and -41.26 mW m^{-2} over the simulation domain, respectively. Li et al. (2019) estimated the annual mean all sky DRF of -100 mW m^{-2} from DMS emissions over the East China Sea. Our estimates (-20.01 mW m^{-2}) were lower than their result, which is likely attributable to discrepancies in the DMS emissions used to drive the model.

3.4.2 Annual DMS impacts to SO$_4^{2-}$, CCN, and radiative forcing between XG and LANA simulation

To quantify the impacts of DMS emissions changes on SO$_4^{2-}$, CCN, and radiative forcing, we compared the XG and LANA simulations (Figure S89(a)). Increases in SO$_4^{2-}$ and CCN can be found in the areas of Indonesia and northwest Pacific Ocean (Figure S89 (a)), which was generally consistent with the changes in DMS emissions fluxes between XG and LANA (Figure 3b). DMS emissions changes (between XG and LANA) accounted for 4-20% and 6-18% of SO$_4^{2-}$ and CCN concentrations over areas of Indonesia, and 2-10% and 3-6% of those concentrations over the northwest Pacific Ocean, respectively. The largest decreases were seen in the vicinity of the India Ocean, which was -0.06 μg m^{-3} for SO$_4^{2-}$ and -10 cm^{-3} for CCN. Due to the higher background concentrations contributed by anthropogenic sources, the relative percent change was smaller over that area, where DMS emissions changes only accounted for -8 to -4 % for SO$_4^{2-}$ and -6 to -3 % for CCN. Also, changes of DMS fluxes around the equatorial belt in western Pacific Ocean (Figure 3b) did not directly link to negative changes in SO$_4^{2-}$ and CCN, which was most likely offset by large scale transport of sulfate caused by DMS from East China Sea. The changes in annual mean DRF and IRF from XG-LANA simulation as shown in Figure S89 (b). Decreases in DRF (-20 to -5 mW m^{-2}) mainly were concentrated over the northwest Pacific Ocean. The largest increase in DRF (up to 40 mW m^{-2}) was found in the areas of the Japan Sea, and most of the increases in DRF (5 to 20 mW m^{-2}) were mainly distributed in the region of the Indian Ocean and land areas of India. Their spatial patterns were consistent with distributions of absolute changes of SO$_4^{2-}$ concentrations. The largest changes in IRF were found in areas of the northwest Pacific Ocean and the Sea of Okhotsk, with changes up to -200 and 200 mW m^{-2}. The decreases of IRF from the XG-LANA simulation can span most Pacific Ocean over the simulation domain and some continental regions, and increases of IRF are more concentrated within the India Ocean and Sea of Okhotsk. Generally, our estimated sea-air DMS fluxed are lower than those from Lana et al. (2011) over the most of the ocean areas, but the DMS-caused changes to SO$_4^{2-}$, CCN, and radiative forcing were more varied, with the increases over the northwest Pacific Ocean for SO$_4^{2-}$ and CCN, and decreases in the regions of India Ocean, oppositely, increases in the regions of India Ocean for DRF and IRF, and decreases over the northwest Pacific Ocean.

3.5 Limitations of this study

We found several limitations in our emission estimates and modeling study. We try to use machine learning estimates of DMS concentrations to fill the regions without observations. ~~However, the primary weakness of the machine learning method is that the training process is not interpretable and not transparent (Reichstein et al., 2019;Wang et al., 2020). the relationship between training parameters should have a minimal physical interpretation, which should be done in future work to give not only accurate but also credible predictions.~~ While the recently measured 1022 seawater DMS observations over Chinese Seas included a training period, for some months (January, November, etc.) there were still not enough data to create a monthly mean. Hence, we temporally averaged input parameters to a seasonal period rather than use monthly data, which is a limitation of this study, but as shown in Section 3.1, the estimated results showed distinct seasonal variations, and the results are

comparable with other studies. Due to the limited continuous measurements of atmospheric DMS and MSA concentrations, we only presented the averaged each cruise survey observations for DMS model evaluation and temporal variation of MSA prediction performance evaluated only from a single observation site. We acknowledged that this is an important limitation of this study, which prevents us from giving comprehensive estimates (in temporal and spatial scale) of the advantage of our updated DMS emissions. More marine and atmospheric observational data are necessary for further model evaluation.

In addition, due to the limited high temporal resolution monitoring data of CCN for the simulation year 2017 in our domain, we verified the model performance of the CCN simulation by comparing the modeled results with the collected mean annual observed concentrations of CCN in other previous studies and satellite-retrieved CCN concentrations. We acknowledge that the CCN model-measurement comparisons listed in Table S45 are not the exact times where CCN simulated, and satellite-retrieved CCN (given the uncertainty in water uptake and size distributions) are not necessarily accurate enough to represent true atmospheric CCN concentrations in 2017.

Modeled AOD may be biased during cloudy conditions when AERONET measurements cannot be made. Hence, there would be an uncertainty in using monthly averaged measurements and model predictions for comparison (Schutgens et al., 2016).

Different chemical mechanisms of various chemical-transport models and the treatment of aerosol optical properties can also lead to differences in simulation results. Globally, the annual-mean DRF and IRF contributed by DMS reported by other studies (as listed in Table S56) varied from -0.23 to -0.074 W m^{-2} and -6.55 to -0.3 W m^{-2}, respectively. Aerosol-cloud interactions are a major source of uncertainty in the prediction of climate change, impacting radiative forcing estimates, especially the IRF calculation. Differences in aerosol nucleation schemes, activation parameterizations, and emissions between models can contribute to large discrepancies in their simulation results (Carslaw et al., 2013). However, we did not explore impact of different nucleation schemes on radiative forcing. We recommend that this should be done in the future work to minimize uncertainties in future modeling studies.

4 Conclusions

In this study, we utilized the XGBoost machine learning algorithm to estimate sea-water DMS concentrations by training 12 ocean environmental parameters on newly updated DMS measurements. 1022 recent seawater DMS measurements over Chinese Sea were included in our training data, and we used the machine learning method to fill the gap at times and in locations without observations. The DMS model-measurement validation results showed that our XGBoost estimates could capture the observed DMS concentrations with a correlation coefficient of 0.75. Zonal mean DMS concentrations between 10°S and 30°N latitude areas of simulation domain were higher in MAM than in JJA, and most of ocean area was concentrated in 10°S and 30°N latitude band, which leaded to the highest regional mean DMS concentrations observed in MAM. We estimated annual DMS emission fluxes of 1.25 Tg (S), which accounted for 15.4% of anthropogenic sulfur emissions over the entire simulation domain (covers most of Asia) in 2017. Comparative analysis revealed that the sea-air DMS flux estimated in this study (from XG estimates) was generally lower than those from global-database DMS emissions (Lana et al., 2011). The

model estimates of DMS and MSA from XG simulation, were evaluated by comparing with cruise survey experiments and long-term online measurement site data. In general, the improvement in model performance can be observed by comparing

455   XG with the LANA simulation, which uses the global-database DMS emissions.

The modeled DMS-induced sulfate DRF and IRF ranged from -200 to -20 mW m^{-2} and -900 to -100 mW m^{-2} across the four seasons, respectively. The stronger negative DRF (-120 to -200 mW m^{-2}) was in the areas where with higher SO_4^{2-} concentrations contributed by DMS, such as the South China Sea, Philippine Sea, and Japan Sea. However, the strong negative IRF was mainly over remote oceans (-900 to -600 mW m^{-2}), which did not match with the spatial distributions of contributions

460   of DMS to CCN concentrations due to the role of clouds in the IRF. Annually, DMS-induced sulfate IRF (-700 to -100 mW m^{-2}) here obviously higher than those all sky DRF (-100 to -10 mW m^{-2}). By adding our updated DMS emissions to a simulation with no DMS (XG-ND), we predict the enhancement of near-surface SO_4^{2-} and CCN concentrations by 0.1-0.3 μg m^{-3} and 3 - 42 cm^{-3}, respectively, over most oceanic areas in all four seasons. We found higher contributions from DMS emissions to SO_4^{2-} and CCN in MAM and DJF than JJA and SON.

465   In this work, we quantified the contributions of DMS to atmospheric SO_4^{2-} and CCN aerosol concentrations along with their radiative effect over a modeled Asian domain (covers most of Asia). This work provides better insights into the source strength of DMS and its impact on climate, addressing knowledge gaps related to factors controlling aerosols in the marine boundary layer and their climate impacts. As discussed in Section 3.6, there are several limitations that need to be improved upon in the future work. More marine and atmospheric observational data are necessary for further DMS emission estimates

470   and model evaluation to explore the interactions of DMS with aerosols and radiative forcing. In the future work, we also need to explore the impact of different aerosol nucleation schemes on radiative forcing, to more completely quantify the uncertainties our modeling study.

Data availability. The GEOS-Chem model is available at https://doi.org/10.5281/zenodo.3974569 (last access: 25 March 2021).

475   The LANA DMS climatology can be downloaded from http://saga.pmel.noaa.gov/dms/ (last access: 1 May 2021), and the regional sea surface DMS concentrations developed in this study can be provided by the authors on request.

Competing interests. The authors declare that they have no conflict of interest.

480   *Acknowledgements.* This work was supported by the National Key Research and Development Program of China (Grant 2016YFA060130X), and the National Natural Science Foundation of China (42077195), the Natural Science Foundation of Shanghai, China (22ZR1407700).

References

Abdul-Razzak, H., and Ghan, S. J.: A parameterization of aerosol activation 3. Sectional representation, Journal of Geophysical Research: Atmospheres, 107, AAC 1-1-AAC 1-6, https://doi.org/10.1029/2001JD000483, 2002.

Adams, P. J., and Seinfeld, J. H.: Predicting global aerosol size distributions in general circulation models, Journal of Geophysical Research: Atmospheres, 107, AAC 4-1-AAC 4-23, https://doi.org/10.1029/2001JD001010, 2002.

Alexander, B., Park, R. J., Jacob, D. J., Li, Q. B., Yantosca, R. M., Savarino, J., Lee, C. C. W., and Thiemens, M. H.: Sulfate formation in sea-salt aerosols: Constraints from oxygen isotopes, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/10.1029/2004JD005659, 2005.

Amos, H. M., Jacob, D. J., Holmes, C. D., Fisher, J. A., Wang, Q., Yantosca, R. M., Corbitt, E. S., Galarneau, E., Rutter, A. P., Gustin, M. S., Steffen, A., Schauer, J. J., Graydon, J. A., Louis, V. L. S., Talbot, R. W., Edgerton, E. S., Zhang, Y., and Sunderland, E. M.: Gas-particle partitioning of atmospheric Hg(II) and its effect on global mercury deposition, Atmos. Chem. Phys., 12, 591-603, 10.5194/acp-12-591-2012, 2012.

Andreae, M. O., and Rosenfeld, D.: Aerosol–cloud–precipitation interactions. Part 1. The nature and sources of cloud-active aerosols, Earth-Science Reviews, 89, 13-41, https://doi.org/10.1016/j.earscirev.2008.03.001, 2008.

Bell, T. G., Porter, J. G., Wang, W.-L., Lawler, M. J., Boss, E., Behrenfeld, M. J., and Saltzman, E. S.: Predictability of Seawater DMS During the North Atlantic Aerosol and Marine Ecosystem Study (NAAMES), 7, 10.3389/fmars.2020.596763, 2021.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, Journal of Geophysical Research: Atmospheres, 106, 23073-23095, https://doi.org/10.1029/2001JD000807, 2001.

Bilsback, K. R., Baumgartner, J., Cheeseman, M., Ford, B., Kodros, J. K., Li, X., Ramnarine, E., Tao, S., Zhang, Y., Carter, E., and Pierce, J. R.: Estimated Aerosol Health and Radiative Effects of the Residential Coal Ban in the Beijing-Tianjin-Hebei Region of China, Aerosol and Air Quality Research, 20, 2332-2346, 10.4209/aaqr.2019.11.0565, 2020a.

Bilsback, K. R., Kerry, D., Croft, B., Ford, B., Jathar, S. H., Carter, E., Martin, R. V., and Pierce, J. R.: Beyond SOx reductions from shipping: assessing the impact of NOx and carbonaceous-particle controls on human health and climate, Environmental Research Letters, 15, 124046, 10.1088/1748-9326/abc718, 2020b.

Bohren, C. F., and Huffman, D. R.: Absorption and scattering of light by small particles, 1983.

Boylan, J. W., and Russell, A. G.: PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models, Atmospheric Environment, 40, 4946-4959, https://doi.org/10.1016/j.atmosenv.2005.09.087, 2006.

Cao, D., Ma, Y., Sun, L., and Gao, L.: Fast observation simulation method based on XGBoost for visible bands over the ocean surface under clear-sky conditions, Remote Sensing Letters, 12, 674-683, 10.1080/2150704X.2021.1925371, 2021.

Carslaw, K. S., Lee, L. A., Reddington, C. L., Pringle, K. J., Rap, A., Forster, P. M., Mann, G. W., Spracklen, D. V., Woodhouse, M. T., Regayre, L. A., and Pierce, J. R.: Large contribution of natural aerosols to uncertainty in indirect forcing, Nature, 503, 67-71, 10.1038/nature12674, 2013.

Charlson, R. J., Lovelock, J. E., Andreae, M. O., and Warren, S. G.: Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate, Nature, 326, 655-661, 10.1038/326655a0, 1987.

Chen, Q., Sherwen, T., Evans, M., and 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, 10.5194/acp-18-13617-2018, 2018.

Chen, T., and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, 2016.

Choi, Y.-N., Song, S.-K., Lee, S. H., and Moon, J.-H.: Estimation of marine dimethyl sulfide emissions from East Asian seas and their impact on natural direct radiative forcing, Atmospheric Environment, 222, 117165, https://doi.org/10.1016/j.atmosenv.2019.117165, 2020.

D'Andrea, S. D., Häkkinen, S. A. K., Westervelt, D. M., Kuang, C., Levin, E. J. T., Kanawade, V. P., Leaitch, W. R., Spracklen, D. V., Riipinen, I., and Pierce, J. R.: Understanding global secondary organic aerosol amount and size-resolved condensational behavior, Atmos. Chem. Phys., 13, 11519-11534, 10.5194/acp-13-11519-2013, 2013.

Duncan Fairlie, T., Jacob, D. J., and Park, R. J.: The impact of transpacific transport of mineral dust in the United States, Atmospheric Environment, 41, 1251-1266, https://doi.org/10.1016/j.atmosenv.2006.09.048, 2007.

Fountoukis, C., and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K^+–Ca^{2+}–Mg^{2+}–NH_4^+–Na^+–SO_4^{2-}–NO_3^-–Cl^-–H_2O aerosols, Atmos. Chem. Phys., 7, 4639-4659, 10.5194/acp-7-4639-2007, 2007.

Gade, K.: A Non-singular Horizontal Position Representation, Journal of Navigation, 63, 395-417, 10.1017/S0373463309990415, 2010.

Galí, M., Devred, E., Levasseur, M., Royer, S.-J., and Babin, M.: A remote sensing algorithm for planktonic dimethylsulfoniopropionate (DMSP) and an analysis of global patterns, Remote Sensing of Environment, 171, 171-184, https://doi.org/10.1016/j.rse.2015.10.012, 2015.

Galí, M., Levasseur, M., Devred, E., Simó, R., and Babin, M.: Sea-surface dimethylsulfide (DMS) concentration from satellite data at global and regional scales, Biogeosciences, 15, 3497-3519, 10.5194/bg-15-3497-2018, 2018.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Journal of Climate, 30, 5419-5454, 10.1175/JCLI-D-16-0758.1, 2017.

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., Welton, E. J., Korkin, S. V., and Lyapustin, A. I.: Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements, Atmos. Meas. Tech., 12, 169-209, 10.5194/amt-12-169-2019, 2019.

Gondwe, M., Krol, M., Gieskes, W., Klaassen, W., and de Baar, H.: The contribution of ocean-leaving DMS to the global atmospheric burdens of DMS, MSA, SO2, and NSS SO4=, Global Biogeochemical Cycles, 17, 10.1029/2002GB001937, 2003.

Gregor, L., Kok, S., and Monteiro, P. M. S.: Empirical methods for the estimation of Southern Ocean CO2: support vector and random forest regression, Biogeosciences, 14, 5551-5569, 10.5194/bg-14-5551-2017, 2017.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geosci. Model Dev., 5, 1471-1492, 10.5194/gmd-5-1471-2012, 2012.

Hodshire, A. L., Campuzano-Jost, P., Kodros, J. K., Croft, B., Nault, B. A., Schroder, J. C., Jimenez, J. L., and Pierce, J. R.: The potential role of methanesulfonic acid (MSA) in aerosol formation and growth and the associated radiative forcings, Atmos. Chem. Phys., 19, 3137-3160, 10.5194/acp-19-3137-2019, 2019.

Hudman, R. C., Moore, N. E., Mebust, A. K., Martin, R. V., Russell, A. R., Valin, L. C., and Cohen, R. C.: Steps towards a mechanistic model of global soil nitric oxide emissions: implementation and space based-constraints, Atmos. Chem. Phys., 12, 7779-7795, 10.5194/acp-12-7779-2012, 2012.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2008JD009944, 2008.

Ivatt, P. D., and Evans, M. J.: Improving the prediction of an atmospheric chemistry transport model using gradient-boosted regression trees, Atmos. Chem. Phys., 20, 8063-8082, 10.5194/acp-20-8063-2020, 2020.

Jathar, S. H., Sharma, N., Bilsback, K. R., Pierce, J. R., Vanhanen, J., Gordon, T. D., and Volckens, J.: Emissions and radiative impacts of sub-10 nm particles from biofuel and fossil fuel cookstoves, Aerosol Science and Technology, 54, 1231-1243, 10.1080/02786826.2020.1769837, 2020.

Jian, S., Zhang, H.-H., Yang, G.-P., and Li, G.-L.: Variation of biogenic dimethylated sulfur compounds in the Changjiang River Estuary and the coastal East China Sea during spring and summer, Journal of Marine Systems, 199, 103222, https://doi.org/10.1016/j.jmarsys.2019.103222, 2019.

Jin, Q., Grandey, B. S., Rothenberg, D., Avramov, A., and Wang, C.: Impacts on cloud radiative effects induced by coexisting aerosols converted from international shipping and maritime DMS emissions, Atmos. Chem. Phys., 18, 16793-16808, 10.5194/acp-18-16793-2018, 2018.

Kettle, A. J., Andreae, M. O., Amouroux, D., Andreae, T. W., Bates, T. S., Berresheim, H., Bingemer, H., Boniforti, R., Curran, M. A. J., DiTullio, G. R., Helas, G., Jones, G. B., Keller, M. D., Kiene, R. P., Leck, C., Levasseur, M., Malin, G., Maspero, M., Matrai, P., McTaggart, A. R., Mihalopoulos, N., Nguyen, B. C., Novo, A., Putaud, J. P., Rapsomanikis, S., Roberts, G., Schebeske, G., Sharma, S., Simó, R., Staubes, R., Turner, S., and Uher, G.: 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, Global Biogeochemical Cycles, 13, 399-444, 10.1029/1999GB900004, 1999.

Khan, M. A. H., Gillespie, S. M. P., Razis, B., Xiao, P., Davies-Coleman, M. T., Percival, C. J., Derwent, R. G., Dyke, J. M., Ghosh, M. V., Lee, E. P. F., and Shallcross, D. E.: A modelling study of the atmospheric chemistry of DMS using the global model, STOCHEM-CRI, Atmospheric Environment, 127, 69-79, https://doi.org/10.1016/j.atmosenv.2015.12.028, 2016.

Kodros, J. K., Cucinotta, R., Ridley, D. A., Wiedinmyer, C., and Pierce, J. R.: The aerosol radiative effects of uncontrolled combustion of domestic waste, Atmos. Chem. Phys., 16, 6771-6784, 10.5194/acp-16-6771-2016, 2016.

Kodros, J. K., and Pierce, J. R.: Important global and regional differences in aerosol cloud-albedo effect estimates between simulations with and without prognostic aerosol microphysics, Journal of Geophysical Research: Atmospheres, 122, 4003-4018, https://doi.org/10.1002/2016JD025886, 2017.

Koepke, P., Hess, M., Schult, I., and Shettle, E.: Global aerosol data set, in, report, 1997.

Kulmala, M., Petäjä, T., Ehn, M., Thornton, J., Sipilä, M., Worsnop, D. R., and Kerminen, V. M.: Chemistry of Atmospheric Nucleation: On the Recent Advances on Precursor Characterization and Atmospheric Cluster Composition in Connection with Atmospheric New Particle Formation, Annual Review of Physical Chemistry, 65, 21-37, 10.1146/annurev-physchem-040412-110014, 2014.

Lana, A., Bell, T. G., Simó, R., Vallina, S. M., Ballabrera-Poy, J., Kettle, A. J., Dachs, J., Bopp, L., Saltzman, E. S., Stefels, J., Johnson, J. E., and Liss, P. S.: An updated climatology of surface dimethylsulfide concentrations and emission fluxes in the global ocean, Global Biogeochemical Cycles, 25, 10.1029/2010GB003850, 2011.

Lee, Y. H., Chen, K., and Adams, P. J.: Development of a global model of mineral dust aerosol microphysics, Atmos. Chem. Phys., 9, 2441-2458, 10.5194/acp-9-2441-2009, 2009.

Lee, Y. H., and Adams, P. J.: A Fast and Efficient Version of the TwO-Moment Aerosol Sectional (TOMAS) Global Aerosol Microphysics Model, Aerosol Science and Technology, 46, 678-689, 10.1080/02786826.2011.643259, 2012.

Lee, Y. H., Pierce, J. R., and Adams, P. J.: Representation of nucleation mode microphysics in a global aerosol model with sectional microphysics, Geosci. Model Dev., 6, 1221-1232, 10.5194/gmd-6-1221-2013, 2013.

Li, J., Han, Z., Yao, X., Xie, Z., and Tan, S.: The distributions and direct radiative effects of marine aerosols over East Asia in springtime, Science of The Total Environment, 651, 1913-1925, https://doi.org/10.1016/j.scitotenv.2018.09.368, 2019.

Li, S., Sarwar, G., Zhao, J., Zhang, Y., Zhou, S., Chen, Y., Yang, G., and Saiz-Lopez, A.: Modeling the Impact of Marine DMS Emissions on Summertime Air Quality Over the Coastal East China Seas, Earth and Space Science, 7, e2020EA001220, https://doi.org/10.1029/2020EA001220, 2020a.

Li, S., Zhang, Y., Zhao, J., Sarwar, G., Zhou, S., Chen, Y., Yang, G., and Saiz-Lopez, A.: Regional and Urban-Scale Environmental Influences of Oceanic DMS Emissions over Coastal China Seas, Atmosphere, 11, 10.3390/atmos11080849, 2020b.

Liu, C., Wang, T., Rosenfeld, D., Zhu, Y., Yue, Z., Yu, X., Xie, X., Li, S., Zhuang, B., Cheng, T., and Niu, S.: Anthropogenic Effects on Cloud Condensation Nuclei Distribution and Rain Initiation in East Asia, Geophysical Research Letters, 47, e2019GL086184, https://doi.org/10.1029/2019GL086184, 2020.

Liu, H., Jacob, D. J., Bey, I., and Yantosca, R. M.: Constraints from 210Pb and 7Be on wet deposition and transport in a global three-dimensional chemical tracer model driven by assimilated meteorological fields, Journal of Geophysical Research: Atmospheres, 106, 12109-12128, https://doi.org/10.1029/2000JD900839, 2001.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I.: From local explanations to global understanding with explainable AI for trees, Nature Machine Intelligence, 2, 56-67, 10.1038/s42256-019-0138-9, 2020.

Mahajan, A. S., Fadnavis, S., Thomas, M. A., Pozzoli, L., Gupta, S., Royer, S.-J., Saiz-Lopez, A., and Simó, R.: Quantifying the impacts of an updated global dimethyl sulfide climatology on cloud microphysics and aerosol radiative forcing, Journal of Geophysical Research: Atmospheres, 120, 2524-2536, https://doi.org/10.1002/2014JD022687, 2015.

Mao, S.-H., Zhuang, G.-C., Liu, X.-W., Jin, N., Zhang, H.-H., Montgomery, A., Liu, X.-T., and Yang, G.-P.: Seasonality of dimethylated sulfur compounds cycling in north China marginal seas, Marine Pollution Bulletin, 170, 112635, https://doi.org/10.1016/j.marpolbul.2021.112635, 2021.

McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E. A., Zheng, B., Crippa, M., Brauer, M., and Martin, R. V.: A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): an application of the Community Emissions Data System (CEDS), Earth Syst. Sci. Data, 12, 3413-3442, 10.5194/essd-12-3413-2020, 2020.

Nightingale, P. D., Malin, G., Law, C. S., Watson, A. J., Liss, P. S., Liddicoat, M. I., Boutin, J., and Upstill-Goddard, R. C.: In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers, Global Biogeochemical Cycles, 14, 373-387, 10.1029/1999GB900091, 2000.

Pan, B.: Application of XGBoost algorithm in hourly PM2.5 concentration prediction, IOP Conference Series: Earth and Environmental Science, 113, 012127, 10.1088/1755-1315/113/1/012127, 2018.

Park, K. T., Jang, S., Lee, K., Yoon, Y. J., Kim, M. S., Park, K., Cho, H. J., Kang, J. H., Udisti, R., Lee, B. Y., and Shin, K. H.: Observational evidence for the formation of DMS-derived aerosols during Arctic phytoplankton blooms, Atmos. Chem. Phys., 17, 9665-9675, 10.5194/acp-17-9665-2017, 2017.

Park, R. J., Jacob, D. J., Field, B. D., Yantosca, R. M., and Chin, M.: Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/10.1029/2003JD004473, 2004.

Pierce, J. R., and Adams, P. J.: Global evaluation of CCN formation by direct emission of sea salt and growth of ultrafine sea salt, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2005JD006186, 2006.

Pierce, J. R., Chen, K., and Adams, P. J.: Contribution of primary carbonaceous aerosol to cloud condensation nuclei: processes and uncertainties evaluated with a global aerosol microphysics model, Atmos. Chem. Phys., 7, 5447-5466, 10.5194/acp-7-5447-2007, 2007.

Price, C., and Rind, D.: A simple lightning parameterization for calculating global lightning distributions, Journal of Geophysical Research: Atmospheres, 97, 9919-9933, https://doi.org/10.1029/92JD00719, 1992.

Pye, H. O. T., Liao, H., Wu, S., Mickley, L. J., Jacob, D. J., Henze, D. K., and Seinfeld, J. H.: Effect of changes in climate and emissions on future sulfate-nitrate-ammonium aerosol levels in the United States, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/10.1029/2008JD010701, 2009.

Qian, Q. F., Jia, X. J., and Lin, H.: Machine Learning Models for the Seasonal Forecast of Winter Surface Air Temperature in North America, Earth and Space Science, 7, e2020EA001140, https://doi.org/10.1029/2020EA001140, 2020.

Quinn, P. K., and Bates, T. S.: The case against climate regulation via oceanic phytoplankton sulphur emissions, Nature, 480, 51-56, 10.1038/nature10580, 2011.

Quinn, P. K., Coffman, D. J., Johnson, J. E., Upchurch, L. M., and Bates, T. S.: Small fraction of marine cloud condensation nuclei made up of sea spray aerosol, Nature Geoscience, 10, 674-679, 10.1038/ngeo3003, 2017.

Ramanathan, V., Crutzen, P. J., Kiehl, J. T., and Rosenfeld, D.: Aerosols, Climate, and the Hydrological Cycle, Science, 294, 2119-2124, 10.1126/science.1064034, 2001.

Rap, A., Scott, C. E., Spracklen, D. V., Bellouin, N., Forster, P. M., Carslaw, K. S., Schmidt, A., and Mann, G.: Natural aerosol direct and indirect radiative effects, Geophysical Research Letters, 40, 3297-3301, https://doi.org/10.1002/grl.50441, 2013.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195-204, 10.1038/s41586-019-0912-1, 2019.

Schutgens, N. A. J., Partridge, D. G., and Stier, P.: The importance of temporal collocation for the evaluation of aerosol models with observations, Atmos. Chem. Phys., 16, 1065-1079, 10.5194/acp-16-1065-2016, 2016.

Scott, C. E., Rap, A., Spracklen, D. V., Forster, P. M., Carslaw, K. S., Mann, G. W., Pringle, K. J., Kivekäs, N., Kulmala, M., Lihavainen, H., and Tunved, P.: The direct and indirect radiative effects of biogenic secondary organic aerosol, Atmos. Chem. Phys., 14, 447-470, 10.5194/acp-14-447-2014, 2014.

Shwartz-Ziv, R., and Armon, A.: Tabular data: Deep learning is not all you need, Information Fusion, 81, 84-90, https://doi.org/10.1016/j.inffus.2021.11.011, 2022.

Silva, S. J., Keller, C. A., and Hardin, J.: Using an Explainable Machine Learning Approach to Characterize Earth System Model Errors: Application of SHAP Analysis to Modeling Lightning Flash Occurrence, Journal of Advances in Modeling Earth Systems, 14, e2021MS002881, https://doi.org/10.1029/2021MS002881, 2022.

Stirnberg, R., Cermak, J., Kotthaus, S., Haeffelin, M., Andersen, H., Fuchs, J., Kim, M., Petit, J. E., and Favez, O.: Meteorology-driven variability of air pollution (PM1) revealed with explainable machine learning, Atmos. Chem. Phys., 21, 3919-3948, 10.5194/acp-21-3919-2021, 2021.

Sun, Y., Yin, H., Lu, X., Notholt, J., Palm, M., Liu, C., Tian, Y., and Zheng, B.: The drivers and health risks of unexpected surface ozone enhancements over the Sichuan Basin, China, in 2020, Atmos. Chem. Phys., 21, 18589-18608, 10.5194/acp-21-18589-2021, 2021.

Thomas, M. A., Suntharalingam, P., Pozzoli, L., Rast, S., Devasthale, A., Kloster, S., Feichter, J., and Lenton, T. M.: Quantification of DMS aerosol-cloud-climate interactions using the ECHAM5-HAMMOZ model in a current climate scenario, Atmos. Chem. Phys., 10, 7425-7438, 10.5194/acp-10-7425-2010, 2010.

Trivitayanurak, W., Adams, P. J., Spracklen, D. V., and Carslaw, K. S.: Tropospheric aerosol microphysics simulation with assimilated meteorology: model description and intermodel comparison, Atmos. Chem. Phys., 8, 3149-3168, 10.5194/acp-8-3149-2008, 2008.

Vallina, S. M., and Simó, R.: Strong Relationship Between DMS and the Solar Radiation Dose over the Global Surface Ocean, Science, 315, 506, 10.1126/science.1133680, 2007.

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, Earth Syst. Sci. Data, 9, 697-720, 10.5194/essd-9-697-2017, 2017.

Wang, S., Elliott, S., Maltrud, M., and Cameron-Smith, P.: Influence of explicit Phaeocystis parameterizations on the global distribution of marine dimethyl sulfide, Journal of Geophysical Research: Biogeosciences, 120, 2158-2177, https://doi.org/10.1002/2015JG003017, 2015.

Wang, W. L., Song, G., Primeau, F., Saltzman, E. S., Bell, T. G., and Moore, J. K.: Global ocean dimethyl sulfide climatology estimated from observations and an artificial neural network, Biogeosciences, 17, 5335-5354, 10.5194/bg-17-5335-2020, 2020.

Wang, Y., Jacob, D. J., and Logan, J. A.: Global simulation of tropospheric O3-NO x -hydrocarbon chemistry: 1. Model formulation, Journal of Geophysical Research: Atmospheres, 103, 10713-10725, https://doi.org/10.1029/98JD00158, 1998.

Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, Atmospheric Environment, 41, 52-63, https://doi.org/10.1016/j.atmosenv.2007.10.058, 2007.

Westervelt, D. M., Pierce, J. R., Riipinen, I., Trivitayanurak, W., Hamed, A., Kulmala, M., Laaksonen, A., Decesari, S., and Adams, P. J.: Formation and growth of nucleated particles into cloud condensation nuclei: model–measurement comparison, Atmos. Chem. Phys., 13, 7645-7663, 10.5194/acp-13-7645-2013, 2013.

Wu, X., Li, P. F., Zhang, H. H., Zhu, M. X., Liu, C. Y., and Yang, G. P.: Acrylic acid and related dimethylated sulfur compounds in the Bohai and Yellow seas during summer and winter, Biogeosciences, 17, 1991-2008, 10.5194/bg-17-1991-2020, 2020.

Xu, F., Yan, S.-B., Zhang, H.-H., Wu, Y.-C., Ma, Q.-Y., Song, Y.-C., Zhuang, G.-C., and Yang, G.-P.: Occurrence and cycle of dimethyl sulfide in the western Pacific Ocean, Limnology and Oceanography, n/a, https://doi.org/10.1002/lno.11797, 2021.

Yang, G.-P., Song, Y.-Z., Zhang, H.-H., Li, C.-X., and Wu, G.-W.: Seasonal variation and biogeochemical cycling of dimethylsulfide (DMS) and dimethylsulfoniopropionate (DMSP) in the Yellow Sea and Bohai Sea, Journal of Geophysical Research: Oceans, 119, 8897-8915, https://doi.org/10.1002/2014JC010373, 2014.

Yang, G.-P., Zhang, S.-H., Zhang, H.-H., Yang, J., and Liu, C.-Y.: Distribution of biogenic sulfur in the Bohai Sea and northern Yellow Sea and its contribution to atmospheric sulfate aerosol in the late fall, Marine Chemistry, 169, 23-32, https://doi.org/10.1016/j.marchem.2014.12.008, 2015a.

Yang, J., Yang, G., Zhang, H., and Zhang, S.: Spatial distribution of dimethylsulfide and dimethylsulfoniopropionate in the Yellow Sea and Bohai Sea during summer, Chinese Journal of Oceanology and Limnology, 33, 1020-1038, 10.1007/s00343-015-4188-5, 2015b.

705   Yang, Y., Wang, H., Smith, S. J., Easter, R., Ma, P. L., Qian, Y., Yu, H., Li, C., and Rasch, P. J.: Global source attribution of sulfate concentration and direct and indirect radiative forcing, Atmos. Chem. Phys., 17, 8903-8922, 10.5194/acp-17-8903-2017, 2017.

Yu, J., Tian, J. Y., Zhang, Z. Y., Yang, G. P., Chen, H. J., Xu, R., and Chen, R.: Role of Calanus sinicus (Copepoda, Calanoida) on Dimethylsulfide and Dimethylsulfoniopropionate Production in Jiaozhou Bay, Journal of Geophysical Research: Biogeosciences, 124, 2481-2498, https://doi.org/10.1029/2018JG004721, 2019.

710   Zender, C. S., Bian, H., and Newman, D.: Mineral Dust Entrainment and Deposition (DEAD) model: Description and 1990s dust climatology, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002JD002775, 2003.

Zhai, X., Song, Y.-C., Li, J.-L., Yang, J., Zhang, H.-H., and Yang, G.-P.: Distribution Characteristics of Dimethylated Sulfur Compounds and Turnover of Dimethylsulfide in the Northern South China Sea During Summer, Journal of Geophysical Research: Biogeosciences, 125, e2019JG005363, https://doi.org/10.1029/2019JG005363, 2020.

715   Zhao, J., Sarwar, G., Gantt, B., Foley, K., Henderson, B. H., Pye, H. O. T., Fahey, K. M., Kang, D., Mathur, R., Zhang, Y., Li, Q., and Saiz-Lopez, A.: Impact of dimethylsulfide chemistry on air quality over the Northern Hemisphere, Atmospheric Environment, 244, 117961, https://doi.org/10.1016/j.atmosenv.2020.117961, 2021.

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K., and Zhang, Q.: Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions, Atmos. Chem. Phys., 18,
720   14095-14111, 10.5194/acp-18-14095-2018, 2018.

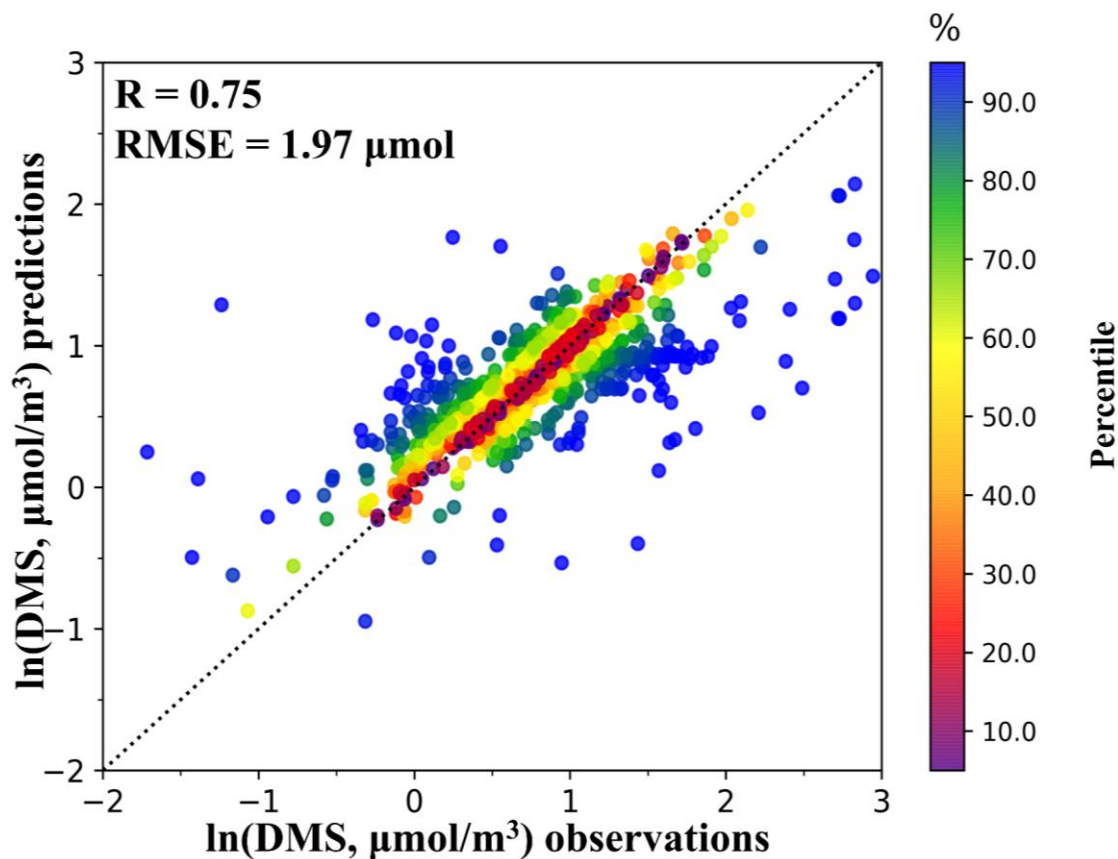
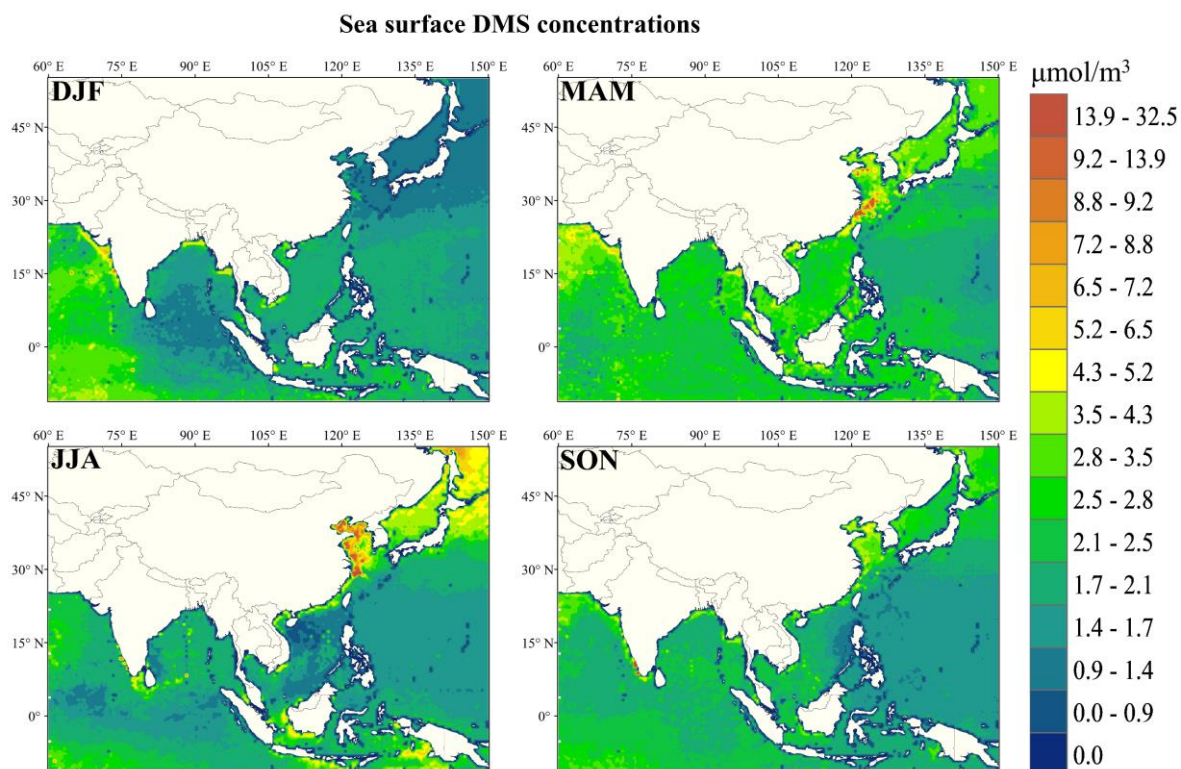
Zhou, S., Chen, Y., Paytan, A., Li, H., Wang, F., Zhu, Y., Yang, T., Zhang, Y., and Zhang, R.: Non-Marine Sources Contribute to Aerosol Methanesulfonate Over Coastal Seas, Journal of Geophysical Research: Atmospheres, 126, e2021JD034960, https://doi.org/10.1029/2021JD034960, 2021.

725

Figure 1. The scatter plot compares model predictions and observations of DMS. The colour represents the percentile of distribution of absolute difference between predicted and observation data.

Sea surface DMS concentrations

730 **Figure 2**. Sea-surface DMS concentrations predicted by XGBoost model by season.

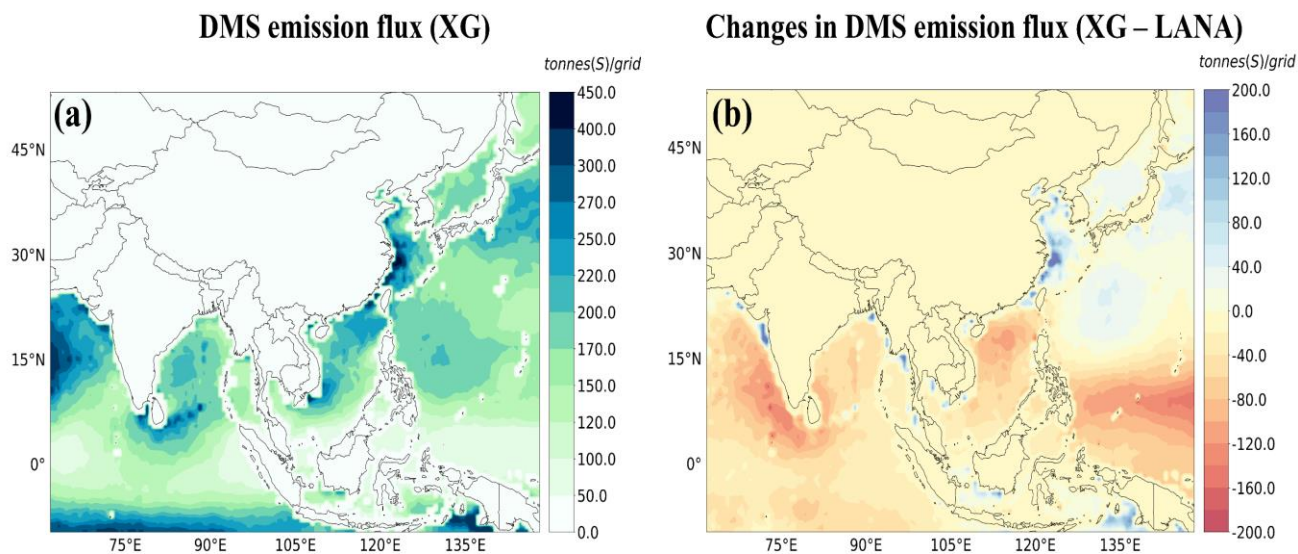



Figure 3. Panel (a) presents Annual DMS emission flux calculated based on N00 flux parameterization (Nightingale et al., 2000) from XG sea surface concentrations, panel (b) presents changes between DMS emission flux from updated (XG) and default climatology (LANA).

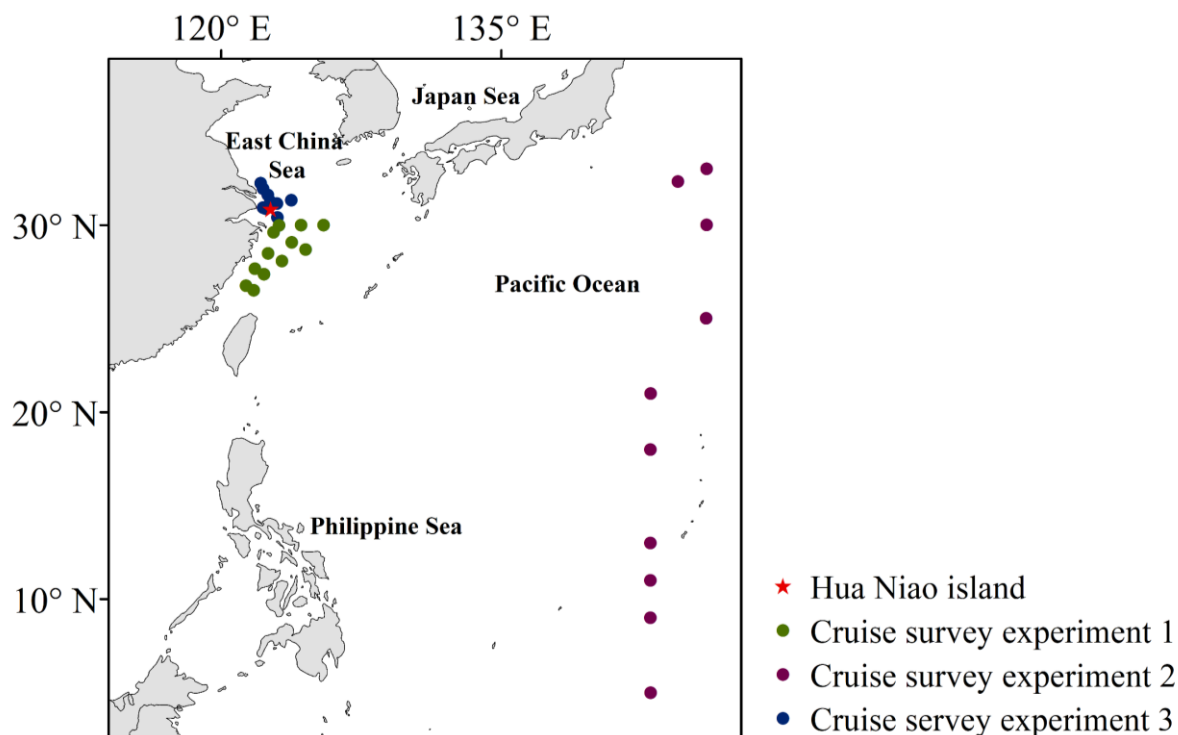



Figure 4. Locations of atmospheric DMS observations from cruise survey experiments 1-3, and the MSA observation site in Hua Niao Island.

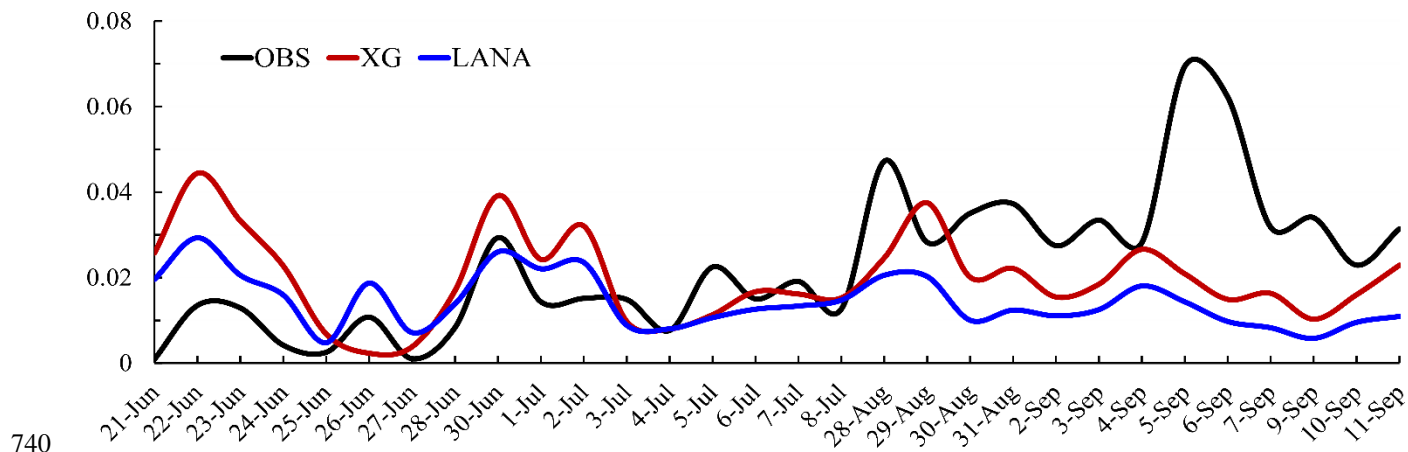
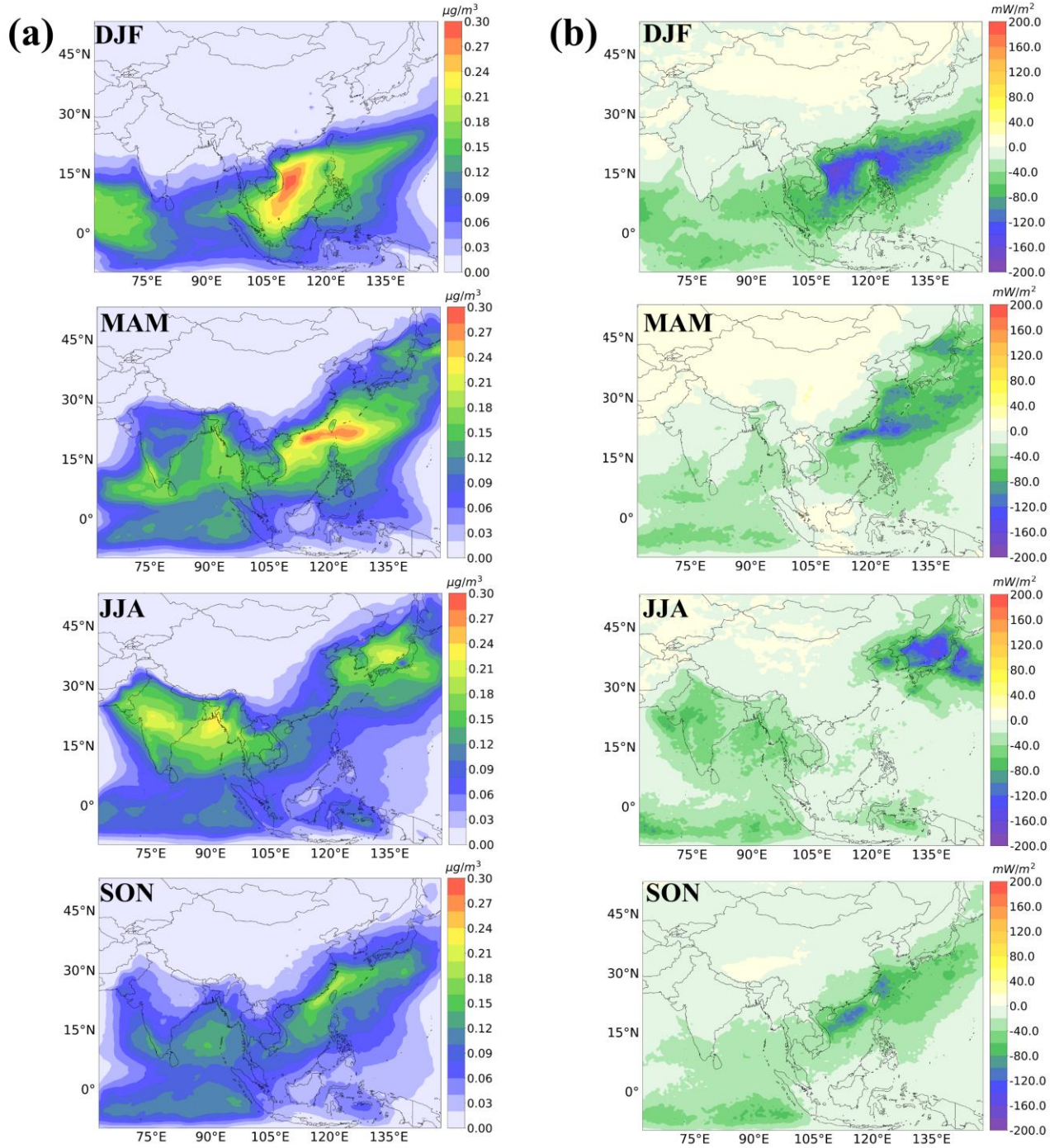



Figure 5. A comparison of simulated daily concentrations of MSA with observations at the Hua Niao Island site (Units: µg m^{-3}).

Absolute contribution of DMS to SO$_4$ (XG − ND) **DMS-derived DRF (XG − ND)**

745 **Figure 6**. Spatial pattern of the seasonal mean absolute changes in surface SO_4^{2-} (first column) and all-sky DRF (second column) between the XG and ND (no DMS) simulations.

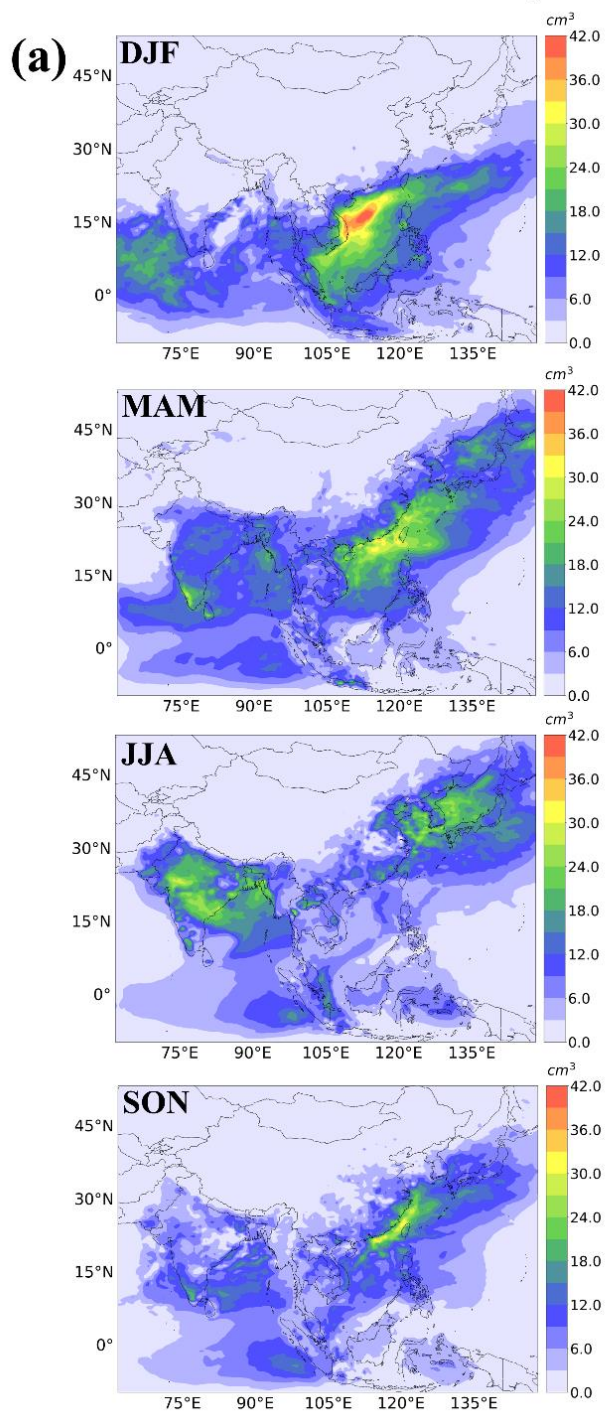
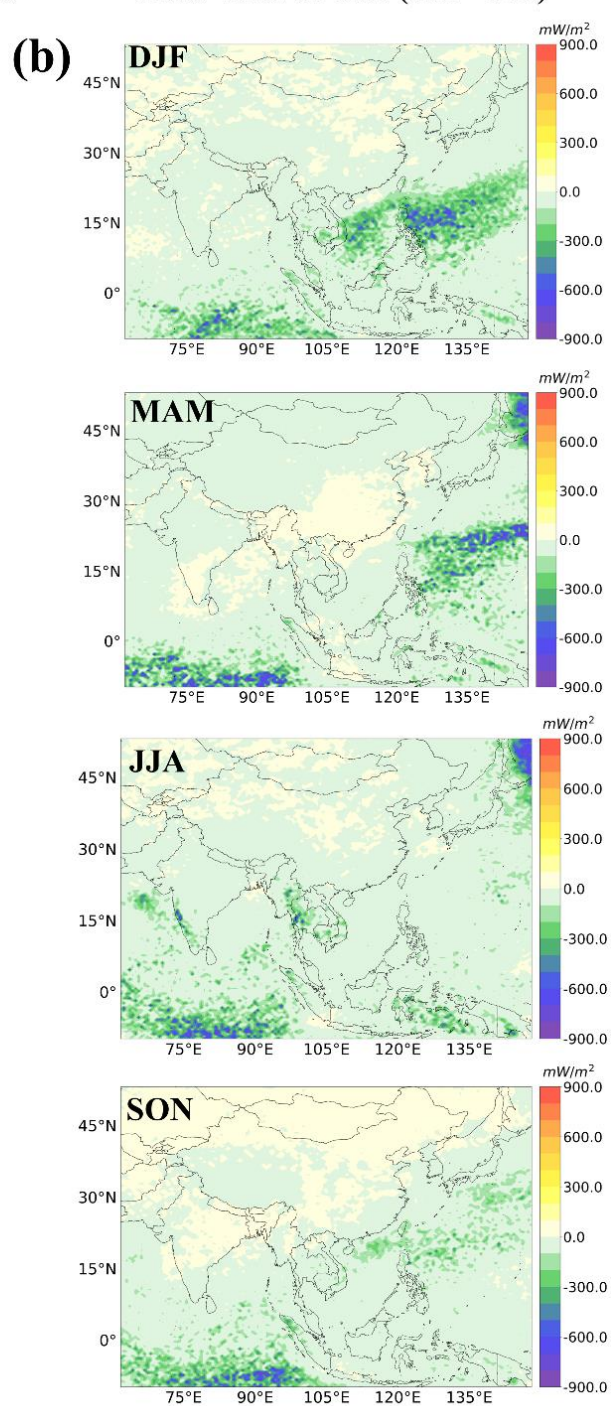



Figure 7. Spatial pattern of the seasonal mean absolute changes in surface CCN (0.2%) (first column) and cloud-albedo IRF (second column) between the XG and ND (no DMS) simulations.

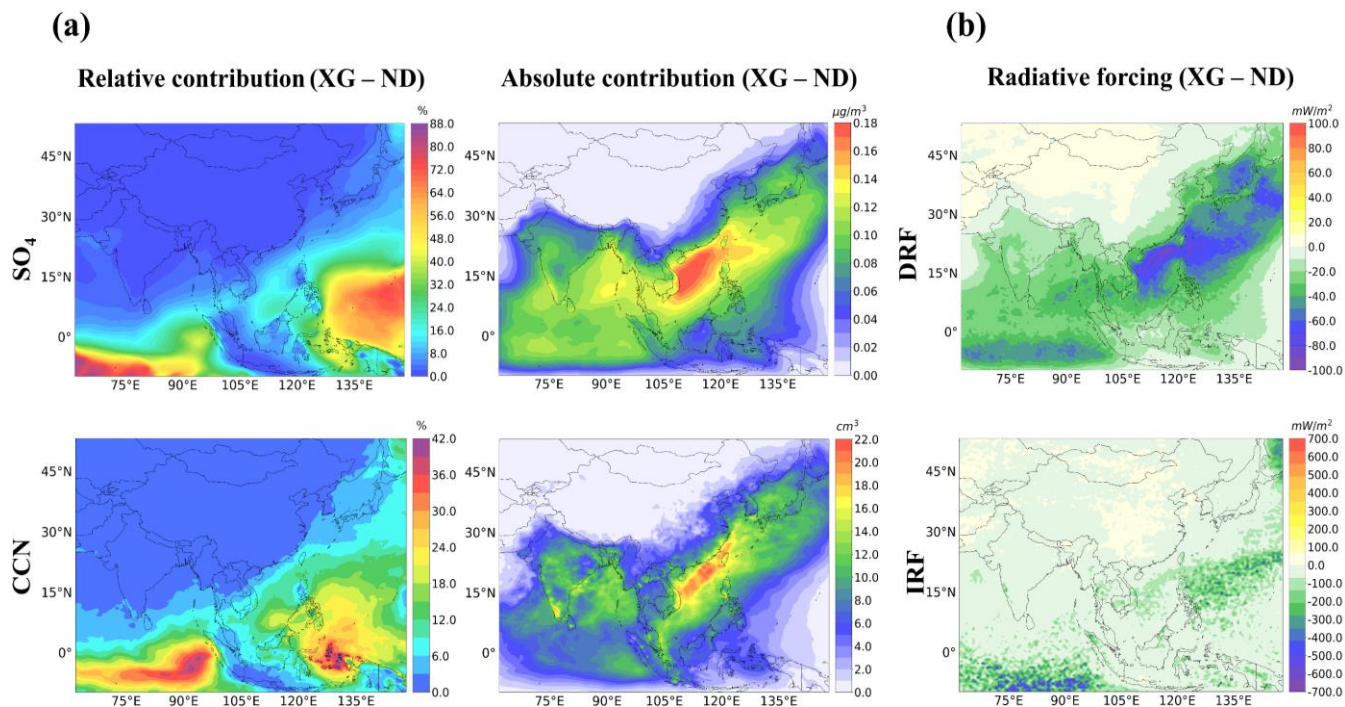



Figure 8. Panel (a) presents the spatial distributions of annual mean percent changes and absolute changes in surface SO_4^{2-} and CCN, and panel (b) presents the spatial distributions of annual mean all-sky DRF and cloud-albedo IRF between XG and ND (no DMS) simulations.

Table 1. Description of simulation

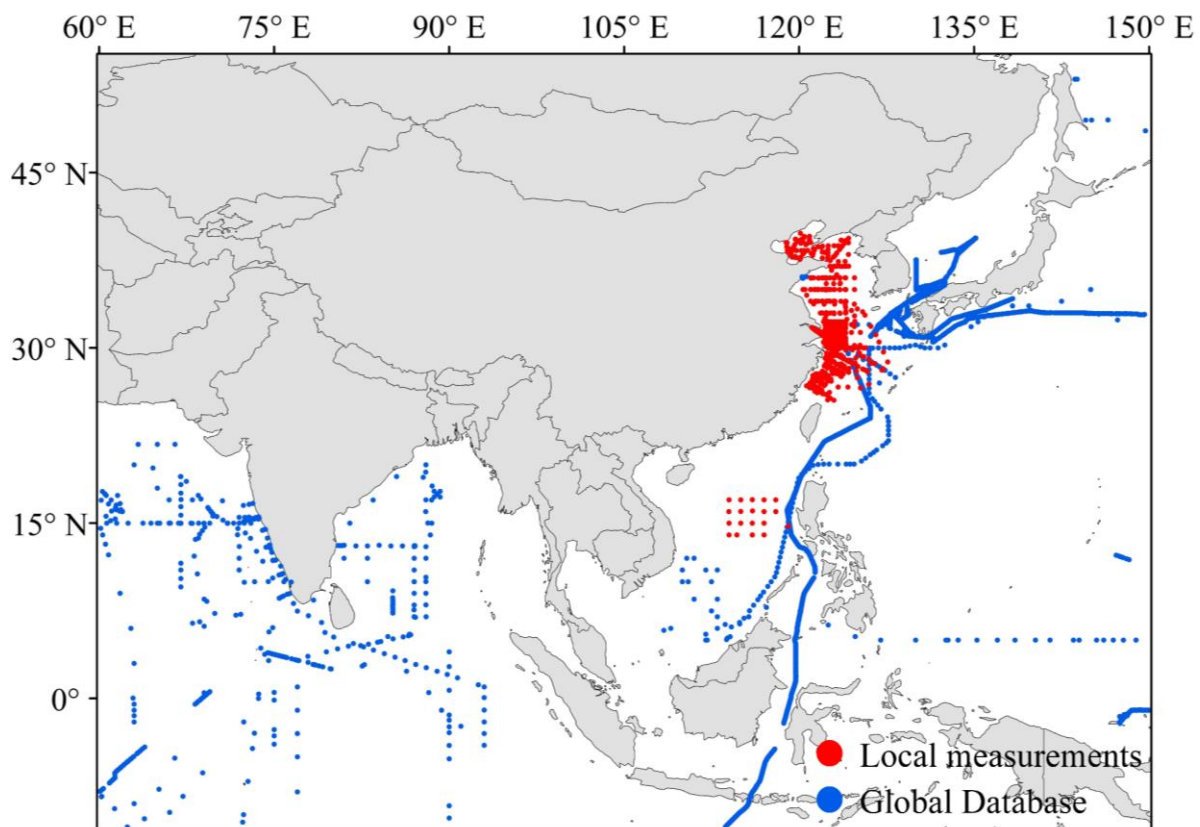
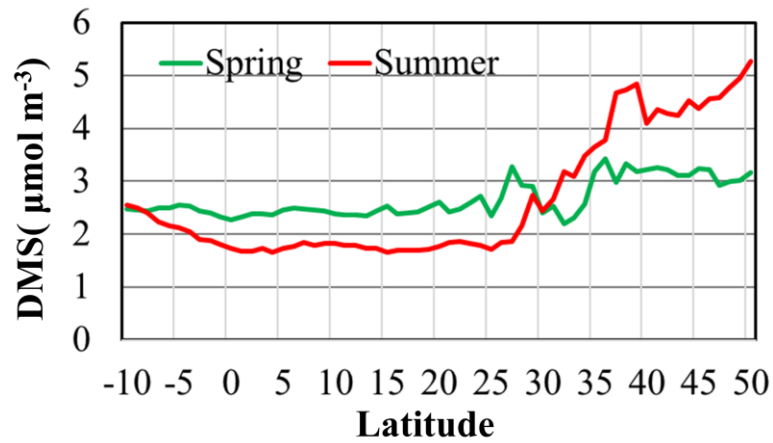
Simulation	Description
XG	DMS emissions on with updated DMS emissions predicted by XGBoost model.
LANA	DMS emissions on with default DMS emissions from Lana et al. (2011).
ND	DMS emissions turned off.

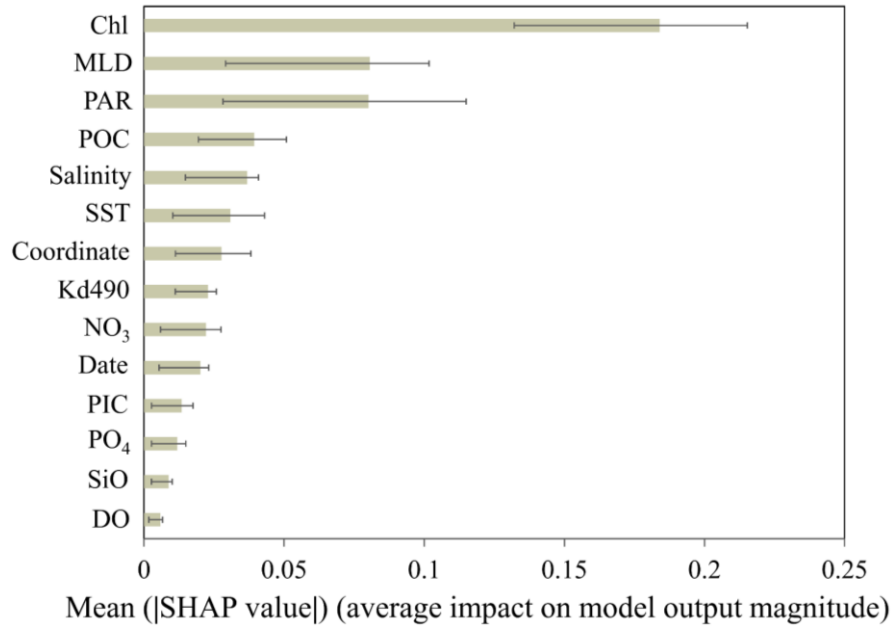
1 *Supplement of*

2 **Simulating the radiative forcing of oceanic dimethylsulfide (DMS) in Asia based on Machine learning**

3 **estimates**

6 **Junri Zhao et al.**

7 *Correspondence to*: Yan Zhang (yan_zhang@fudan.edu.cn)

10 **Figure S1.** Spatial distribution of DMS measurements in simulation domain (11°S to 55°N, 60–150°E).

13 **Figure S2.** Zonal mean DMS sea-surface concentrations predicted by XGBoost between -10 to 50° latitude band of

14 entire simulation domain (11°S to 55°N, 60–150°E).

17 **Figure S3.** The ranked mean SHAP values of each predictor across all prediction cases, and the line ranges represent

18 the upper and lower quantile (25th–75th percentile) across the distribution.

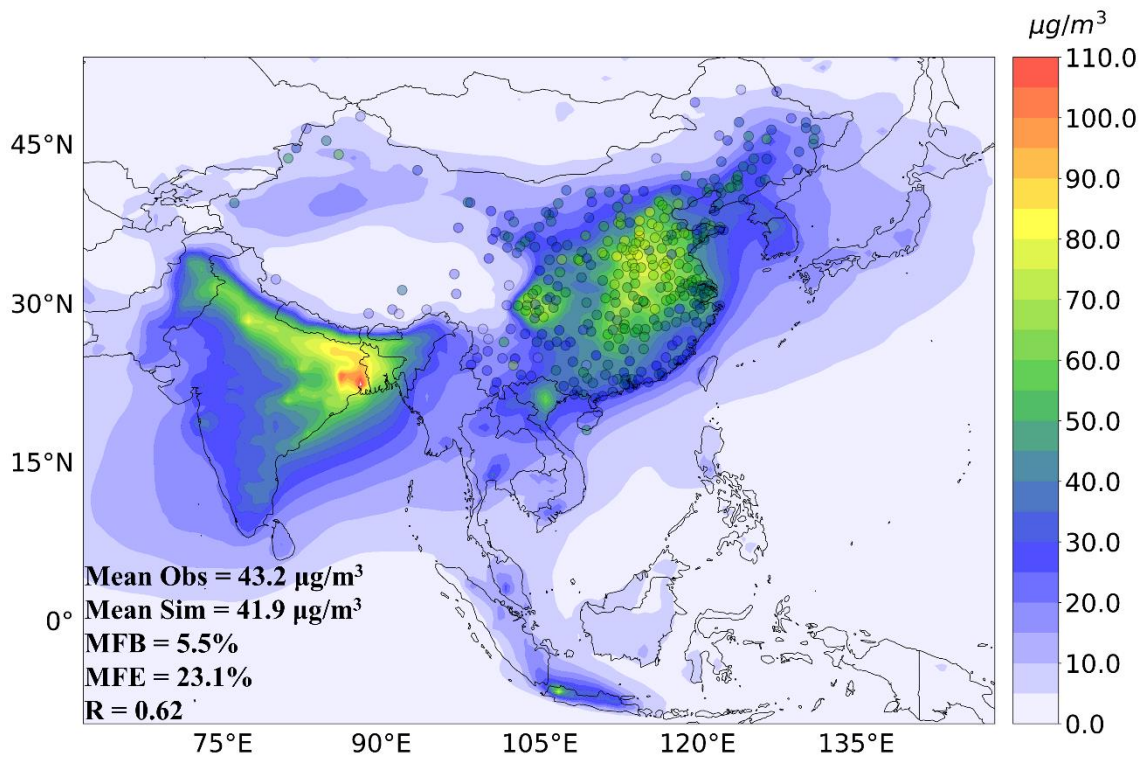



Figure S34. Spatial plots of annual mean concentrations from XG simulation (background plots) and observed (circles) values of PM₂.₅. The mean observation (Mean Obs), mean simulation (Mean Sim), mean fractional error (MFE), and mean fractional bias (MFB) displayed at lower left corner are the statistic result of averaged 365 grids.

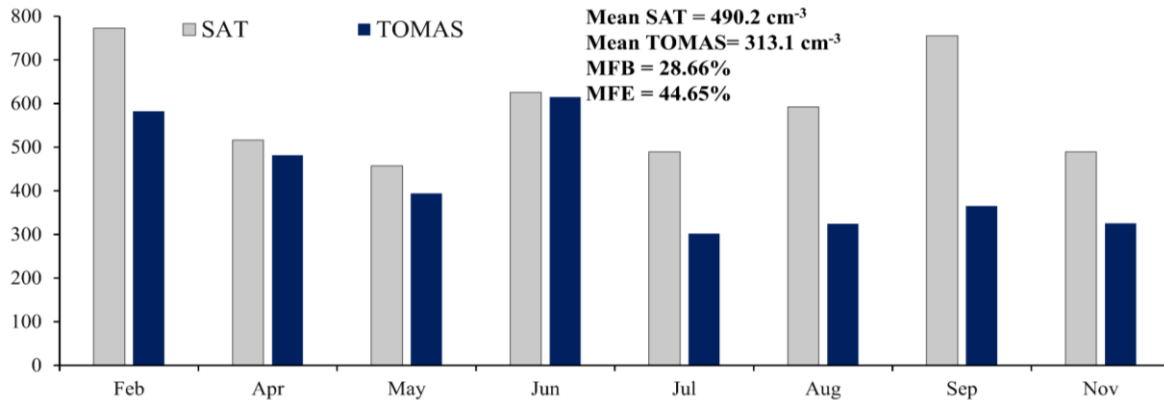



Figure S45. Monthly mean simulated CCN concentrations (TOMAS) from XG simulation and satellite-retrieved CCN concentrations (SAT) from Liu et al., (2020). The mean satellite-retrieved (Mean SAT) CCN concentrations, mean simulation (Mean TOMAS), mean fractional error (MFE), and mean fractional bias (MFB) displayed at upper centre are the statistic result of averaged on the simulation grid for corresponding satellite-retrieved CCN geographical location (in total 667 grids).

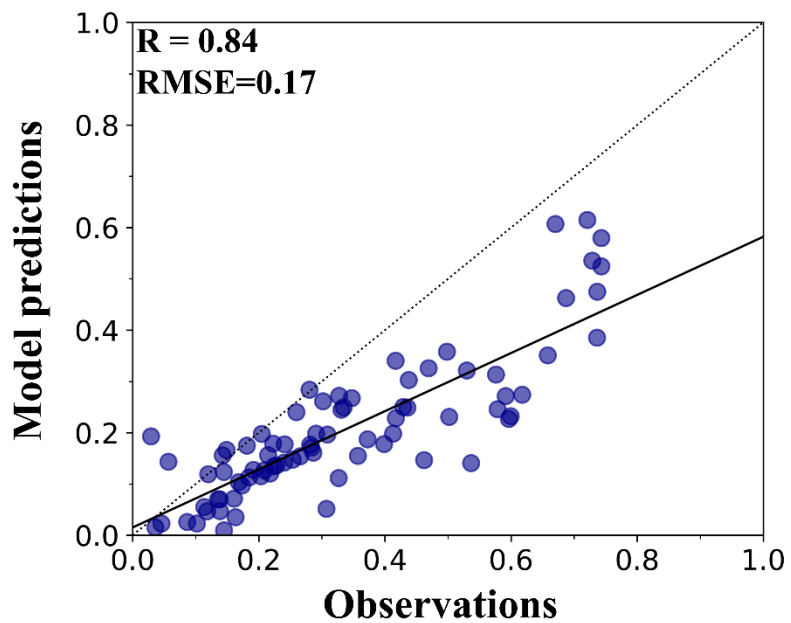
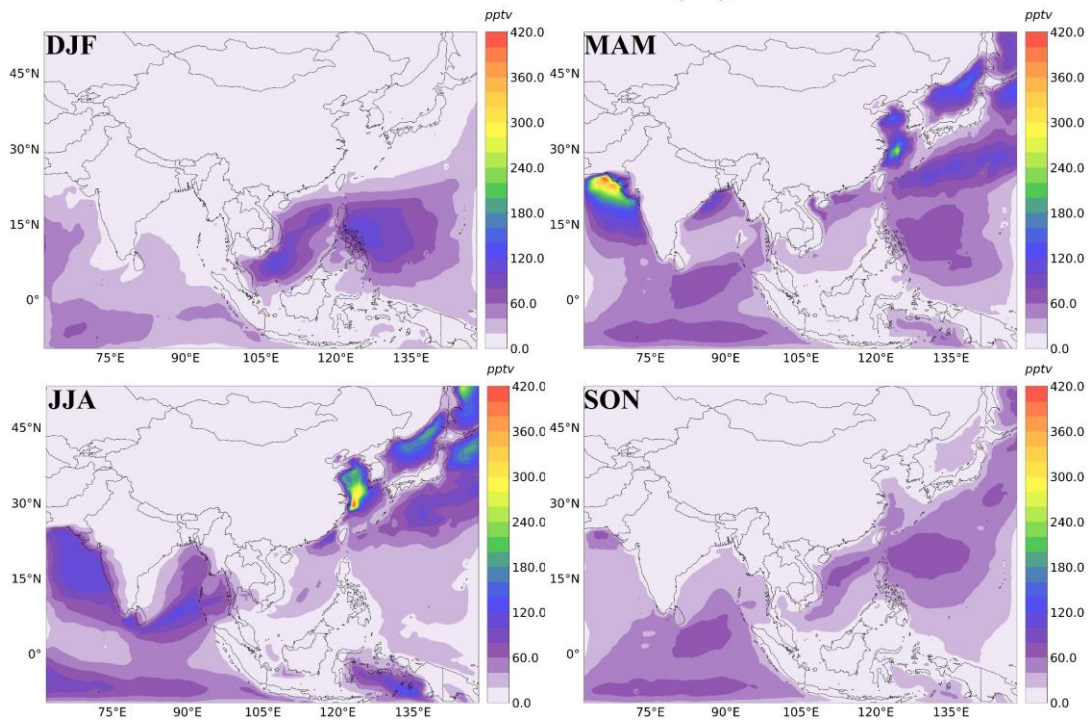
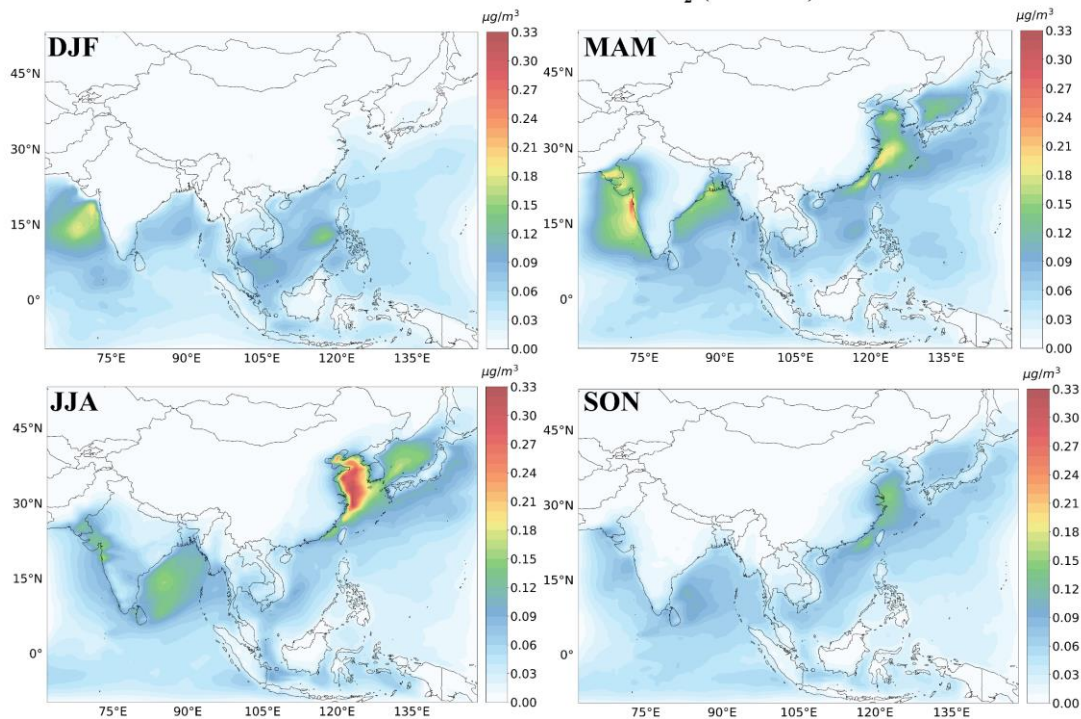



Figure S56. The scatter plot of AOD compares annual mean model predictions (XG simulation) and observations from
AERONET. The black solid line and dashed line indicate trendline and 1:1 line, respectively.

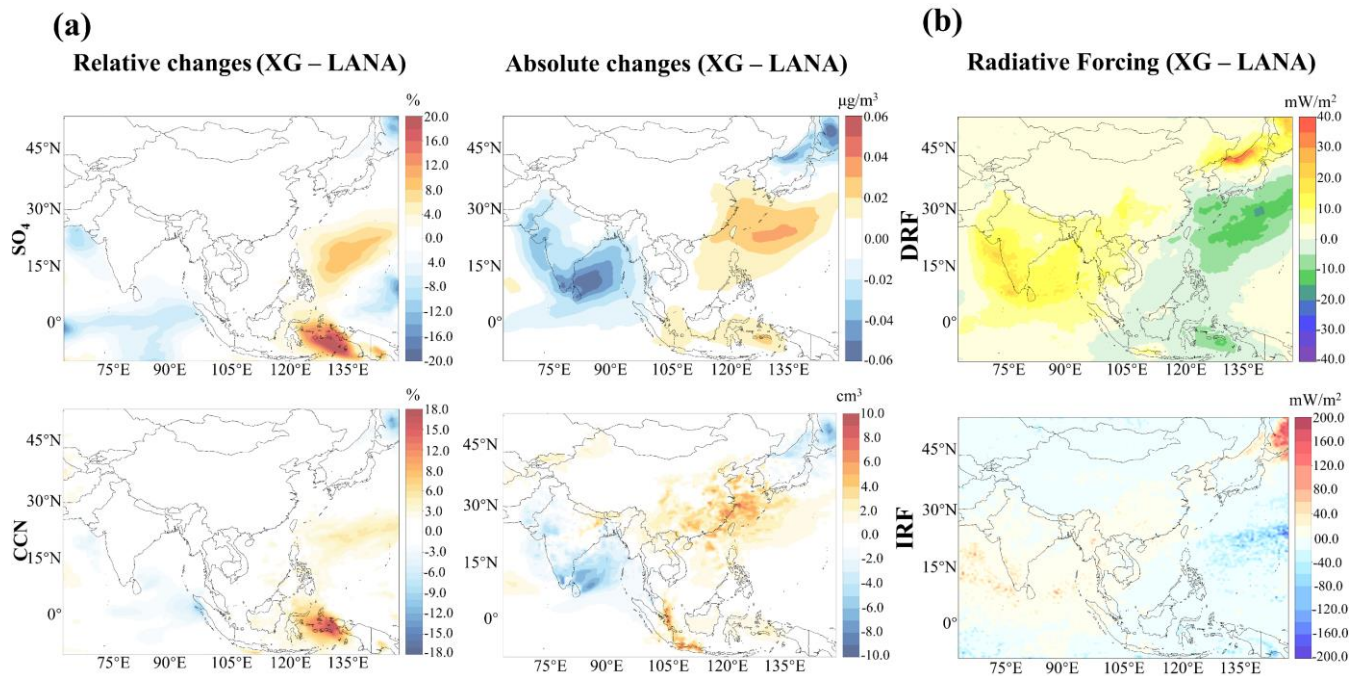
Surface DMS concentrations (XG)

37

38 **Figure S67.** Seasonal variations of atmospheric surface DMS concentrations estimated by the XG simulation.

39

Absolute contribution of DMS to SO₂ (XG – ND)

40

41 **Figure S78.** Seasonal variations of absolute changes in surface SO₂ between the XG and ND (no DMS) simulations.

42

43

44   **Figure S89**. Panel (a) presents the spatial distributions of annual mean percent changes and absolute changes in surface

45   SO_4^{2-} and CCN, and panel (b) presents the spatial distributions of annual mean all-sky DRF and cloud-albedo IRF between

46   the XG and LANA simulations.

47

48   **Table S1.** Training dataset participated in machine learning estimates.

Variables	units	Filtering threshold	sources
DMS measurements in China Seas	$\mu mol\ m^{-3}$	Remove values < 0.1 and values >100	Local measurements from Ocean University of China (Yang et al., 2015a;Yang et al., 2014;Yang et al., 2015b;Li et al., 2020b;Li et al., 2020a)
DMS measurements from global database	$\mu mol\ m^{-3}$	Remove values < 0.1 and values >100	http://saga.pmel.noaa.gov/dms/ (Kettle et al., 1999;Lana et al., 2011)
Salinity	unitless	Remove values < 30 and sampling depth >10m	https://www.ncei.noaa.gov/products/world-ocean-atlas (Garcia et al., 2019)
Sea surface temperature (SST)	°C	-	https://www.ncei.noaa.gov/products/world-ocean-atlas (Garcia et al., 2019)
Dissolved Oxgyen (DO)	$\mu mol\ kg^{-1}$	Remove values < 0.01	https://www.ncei.noaa.gov/products/world-ocean-atlas (Garcia et al., 2019)
Silicate	$\mu mol\ kg^{-1}$	Remove values < 0.01	https://www.ncei.noaa.gov/products/world-ocean-atlas (Garcia et al., 2019)
Phosphate	$\mu mol\ kg^{-1}$	Remove values < 0.01	https://www.ncei.noaa.gov/products/world-ocean-atlas (Garcia et al., 2019)
Nitrate	$\mu mol\ kg^{-1}$	Remove values < 0.01	https://www.ncei.noaa.gov/products/world-ocean-atlas (Garcia et al., 2019)
Mixed layer depth (MLD)	m	Remove values >150	https://www.pmel.noaa.gov/mimoc/ (Schmidtko et al., 2013)
Chlorophyll concentration (Chl)	$mg\ m^{-3}$	Remove values < 0.01	https://oceancolor.gsfc.nasa.gov/l3
Diffuse attenuation coefficient at 490m (kd490)	m^{-1}	-	https://oceancolor.gsfc.nasa.gov/l3
Particulate Inorganic Carbon (PIC)	$mg\ m^{-3}$	-	https://oceancolor.gsfc.nasa.gov/l3
Particulate organic Carbon (POC)	$mg\ m^{-3}$	-	https://oceancolor.gsfc.nasa.gov/l3
Photosynthetically available Radiation (PAR)	Einstein $m^{-2}\ d^{-1}$	-	https://oceancolor.gsfc.nasa.gov/atbd/par/

49

50   **Table S2**. Model performance for sea surface DMS concentration predictions.

	OBS	PRE	MB	RMSE	NMB
Winter	1.99	1.65	-0.34	1.69	-16.87

Spring	2.92	2.71	-0.21	1.81	-6.51
Summer	3.93	3.58	-0.35	2.19	-11.82
Fall	2.87	2.28	-0.59	2.08	-19.36
Annual average	2.57	2.31	-0.26	1.97	-10.28

Notes: OBS: mean observation value; PRE: mean prediction value; MB: mean bias; RMSE: root mean squared error; NMB: normalized mean bias.

Table S23. Comparison of seasonal mean DMS concentrations over Asia region from other studies ($\mu mol\ m^{-3}$).

Season	This study	Lana et al. (2011)	Wang et al. (2020)
Winter	1.93	1.78	1.30
Spring	2.52	2.33	2.21
Summer	2.19	2.01	1.81
Fall	2.03	1.87	1.71

Table S34. Comparison of regional mean modelled atmospheric DMS concentrations with observations from Cruise Survey Experiment (CSE) 1-3(Units: pptv).

	CSE 1	CSE 2	CSE 3
OBS	39.74	57.82	217.51
XG	46.42	41.65	334.90
LANA	23.40	36.71	95.94
OBS-XG	6.68	-16.17	117.39
OBS-LANA	-16.34	-21.11	-121.57

Notes: OBS refer to observations; XG and LANA refer to simulation values; OBS-XG and OBS-LANA refer to differences between observations and simulation values (XG and LANA).

Table S45. Comparison of mean annual CCN predictions from other studies (Units: cm^{-3})

City	Longitude	Latitude	CCN_obs	CCN_sim	Sources
Nanjing	118.7°E	32.06°N	1540.5	919.1	(Liu et al., 2020)
Nanjing	118.7°E	32.21°N	1669.2	994.3	(Liu et al., 2020)

Shanghai	121.29°E	31.18°N	2783.0	983.3	(Liu et al., 2020)
Guangzhou	113.21°E	23.07°N	1299.3	796.9	(Liu et al., 2020)
Shenzhen	114.56°E	22.48°N	1289.0	501.9	(Liu et al., 2020)
Taiwan	121.54°E	25.3°N	1010.5	325.4	(Cheung et al., 2020)
Indo-Gangetic Plain	85.81°E	20.24°N	2250.0	948.7	(Jayachandran et al., 2020)
Indo-Gangetic Plain	82.85°E	25.45°N	1650.0	1432.4	(Jayachandran et al., 2020)
Indo-Gangetic Plain	73.04°E	26.25°N	1375.0	602.2	(Jayachandran et al., 2020)
Eastern Himalaya	88.2°E	27°N	1800.0	1787.7	(Roy et al., 2017)
Central Himalaya	79.5°E	29.4°N	1000.0	1232.0	(Dumka et al., 2015)
Hyderabad	78.38°E	17.45°N	1100.0	1564.2	(Varghese et al., 2016)
Mahabaleshwar	73.4°E	17.6°N	1500.0	670.0	(Leena et al., 2016)

62

63 **Table S56.** Comparison of modelled annual mean DRF and IRF for DMS induced sulfate reported in previous studies.

Annual Mean DRF (W m^{-2})	Annual Mean IRF (W m^{-2})	Model	Sources
-0.23	-0.76	GLOMAP-mode global aerosol microphysics model	(Rap et al., 2013)
-0.074	-0.23	CESM-CAM5	(Yang et al., 2017)
--	-6.55	CESM-CAM5	(Jin et al., 2018)
--	-2.0	ECHAM5-HAMMOZ	(Thomas et al., 2010)
--	-1.79	ECHAM5-HAMMOZ	(Mahajan et al., 2015)

64 **Notes:** -- represents corresponding values that have not reported in their studies.

65

66 **References:**

67 Cheung, H. C., Chou, C. C. K., Lee, C. S. L., Kuo, W. C., and Chang, S. C.: Hygroscopic properties and cloud condensation
68 nuclei activity of atmospheric aerosols under the influences of Asian continental outflow and new particle formation at a
69 coastal site in eastern Asia, Atmos. Chem. Phys., 20, 5911-5922, 10.5194/acp-20-5911-2020, 2020.
70 Dumka, U. C., Bhattu, D., Tripathi, S. N., Kaskaoutis, D. G., and Madhavan, B. L.: Seasonal inhomogeneity in cloud
71 precursors over Gangetic Himalayan region during GVAX campaign, Atmospheric Research, 155, 158-175,
72 https://doi.org/10.1016/j.atmosres.2014.11.022, 2015.

Jayachandran, V. N., Suresh Babu, S. N., Vaishya, A., Gogoi, M. M., Nair, V. S., Satheesh, S. K., and Krishna Moorthy, K.: Altitude profiles of cloud condensation nuclei characteristics across the Indo-Gangetic Plain prior to the onset of the Indian summer monsoon, Atmos. Chem. Phys., 20, 561-576, 10.5194/acp-20-561-2020, 2020.

Jin, Q., Grandey, B. S., Rothenberg, D., Avramov, A., and Wang, C.: Impacts on cloud radiative effects induced by coexisting aerosols converted from international shipping and maritime DMS emissions, Atmos. Chem. Phys., 18, 16793-16808, 10.5194/acp-18-16793-2018, 2018.

Kettle, A. J., Andreae, M. O., Amouroux, D., Andreae, T. W., Bates, T. S., Berresheim, H., Bingemer, H., Boniforti, R., Curran, M. A. J., DiTullio, G. R., Helas, G., Jones, G. B., Keller, M. D., Kiene, R. P., Leck, C., Levasseur, M., Malin, G., Maspero, M., Matrai, P., McTaggart, A. R., Mihalopoulos, N., Nguyen, B. C., Novo, A., Putaud, J. P., Rapsomanikis, S., Roberts, G., Schebeske, G., Sharma, S., Simó, R., Staubes, R., Turner, S., and Uher, G.: 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, Global Biogeochemical Cycles, 13, 399-444, 10.1029/1999GB900004, 1999.

Lana, A., Bell, T. G., Simó, R., Vallina, S. M., Ballabrera-Poy, J., Kettle, A. J., Dachs, J., Bopp, L., Saltzman, E. S., Stefels, J., Johnson, J. E., and Liss, P. S.: An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean, Global Biogeochemical Cycles, 25, 10.1029/2010GB003850, 2011.

Leena, P. P., Pandithurai, G., Anilkumar, V., Murugavel, P., Sonbawne, S. M., and Dani, K. K.: Seasonal variability in aerosol, CCN and their relationship observed at a high altitude site in Western Ghats, Meteorology and Atmospheric Physics, 128, 143-153, 10.1007/s00703-015-0406-0, 2016.

Li, S., Sarwar, G., Zhao, J., Zhang, Y., Zhou, S., Chen, Y., Yang, G., and Saiz-Lopez, A.: Modeling the Impact of Marine DMS Emissions on Summertime Air Quality Over the Coastal East China Seas, Earth and Space Science, 7, e2020EA001220, https://doi.org/10.1029/2020EA001220, 2020a.

Li, S., Zhang, Y., Zhao, J., Sarwar, G., Zhou, S., Chen, Y., Yang, G., and Saiz-Lopez, A.: Regional and Urban-Scale Environmental Influences of Oceanic DMS Emissions over Coastal China Seas, Atmosphere, 11, 10.3390/atmos11080849, 2020b.

Liu, C., Wang, T., Rosenfeld, D., Zhu, Y., Yue, Z., Yu, X., Xie, X., Li, S., Zhuang, B., Cheng, T., and Niu, S.: Anthropogenic Effects on Cloud Condensation Nuclei Distribution and Rain Initiation in East Asia, Geophysical Research Letters, 47, e2019GL086184, https://doi.org/10.1029/2019GL086184, 2020.

Mahajan, A. S., Fadnavis, S., Thomas, M. A., Pozzoli, L., Gupta, S., Royer, S.-J., Saiz-Lopez, A., and Simó, R.: Quantifying the impacts of an updated global dimethyl sulfide climatology on cloud microphysics and aerosol radiative forcing, Journal of Geophysical Research: Atmospheres, 120, 2524-2536, https://doi.org/10.1002/2014JD022687, 2015.

Rap, A., Scott, C. E., Spracklen, D. V., Bellouin, N., Forster, P. M., Carslaw, K. S., Schmidt, A., and Mann, G.: Natural aerosol direct and indirect radiative effects, Geophysical Research Letters, 40, 3297-3301, https://doi.org/10.1002/grl.50441, 2013.

Roy, A., Chatterjee, A., Sarkar, C., Das, S. K., Ghosh, S. K., and Raha, S.: A study on aerosol-cloud condensation nuclei (CCN) activation over eastern Himalaya in India, Atmospheric Research, 189, 69-81, https://doi.org/10.1016/j.atmosres.2017.01.015, 2017.

Schmidtko, S., Johnson, G. C., and Lyman, J. M.: MIMOC: A global monthly isopycnal upper-ocean climatology with mixed layers, Journal of Geophysical Research: Oceans, 118, 1658-1672, https://doi.org/10.1002/jgrc.20122, 2013.

Thomas, M. A., Suntharalingam, P., Pozzoli, L., Rast, S., Devasthale, A., Kloster, S., Feichter, J., and Lenton, T. M.: Quantification of DMS aerosol-cloud-climate interactions using the ECHAM5-HAMMOZ model in a current climate scenario, Atmos. Chem. Phys., 10, 7425-7438, 10.5194/acp-10-7425-2010, 2010.

Varghese, M., Prabha, T. V., Malap, N., Resmi, E. A., Murugavel, P., Safai, P. D., Axisa, D., Pandithurai, G., and Dani, K.: Airborne and ground based CCN spectral characteristics: Inferences from CAIPEEX – 2011, Atmospheric Environment, 125, 324-336, https://doi.org/10.1016/j.atmosenv.2015.06.041, 2016.

Wang, W. L., Song, G., Primeau, F., Saltzman, E. S., Bell, T. G., and Moore, J. K.: Global ocean dimethyl sulfide climatology estimated from observations and an artificial neural network, Biogeosciences, 17, 5335-5354, 10.5194/bg-17-5335-2020, 2020.

Yang, G.-P., Song, Y.-Z., Zhang, H.-H., Li, C.-X., and Wu, G.-W.: Seasonal variation and biogeochemical cycling of dimethylsulfide (DMS) and dimethylsulfoniopropionate (DMSP) in the Yellow Sea and Bohai Sea, Journal of Geophysical Research: Oceans, 119, 8897-8915, https://doi.org/10.1002/2014JC010373, 2014.

123 Yang, G.-P., Zhang, S.-H., Zhang, H.-H., Yang, J., and Liu, C.-Y.: Distribution of biogenic sulfur in the Bohai Sea and
124 northern Yellow Sea and its contribution to atmospheric sulfate aerosol in the late fall, Marine Chemistry, 169, 23-32,
125 https://doi.org/10.1016/j.marchem.2014.12.008, 2015a.
126 Yang, J., Yang, G., Zhang, H., and Zhang, S.: Spatial distribution of dimethylsulfide and dimethylsulfoniopropionate in the
127 Yellow Sea and Bohai Sea during summer, Chinese Journal of Oceanology and Limnology, 33, 1020-1038, 10.1007/s00343-
128 015-4188-5, 2015b.
129 Yang, Y., Wang, H., Smith, S. J., Easter, R., Ma, P. L., Qian, Y., Yu, H., Li, C., and Rasch, P. J.: Global source attribution of
130 sulfate concentration and direct and indirect radiative forcing, Atmos. Chem. Phys., 17, 8903-8922, 10.5194/acp-17-8903-
131 2017, 2017.
132