Current model capabilities for simulating black carbon and sulfate concentrations in the Arctic atmosphere: a multi-model evaluation using a comprehensive measurement data set

S. Eckhardt[1], B. Quennehen[2*], D.J.L. Olivié[3], T.K. Berntsen[4], R. Cherian[5], J. H. Christensen[6], W. Collins[7,8], S. Crepinsek[9,10], N. Daskalakis[11], M. Flanner[12], A. Herber[13], C. Heyes[14], Ø. Hodnebrog[4], L. Huang[15], M. Kanakidou[11], Z. Klimont[14], J. Langner[16], K.S. Law[2], M. T. Lund[4], R. Mahmood[19,20], A. Massling[6], S. Myriokefalitakis[11], I. E. Nielsen[6], J. K. Nøjgaard[6], J. Quaas[5], P. K. Quinn[17], J.-C. Raut[2], S. T. Rumbold[7,8], M. Schulz[3], S. Sharma[15], R.B. Skeie[4], H. Skov[6], T. Uttal[10], K. von Salzen[18] and A. Stohl[1]

[1]{NILU - Norwegian Institute for Air Research, Kjeller, Norway}

[2]{Sorbonne Universités, UPMC Univ. Paris 06 ; Université Versailles St-Quentin ; CNRS/INSU ; LATMOS-IPSL, UMR 8190, Paris, France}

[3]{Norwegian Meteorological Institute, Oslo, Norway}

[4]{Center for International Climate and Environmental Research – Oslo (CICERO), Oslo, Norway}

[5]{Institute for Meteorology, Universität Leipzig, Germany}

[6]{ENVS Department of Environmental Science, Aarhus University, Roskilde, Denmark}

[7]{Met Office Hadley Centre, Exeter, UK}

[8]{University of Reading, Reading, UK}

[9] {Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA}

[10]. NOAA Earth System Research Laboratory Physical Sciences Division / Polar Observations & Processes. Boulder, Colorado, USA

[11]{Environmental Chemical Processes Laboratory, Department of Chemistry, University of Crete, Heraklion, Crete, and ICE-HT/FORTH, Patras, Greece }

[12]{Department of Atmospheric, Oceanic, and Space Sciences, University of Michigan, Ann Arbor, MI, USA}

[13]{Alfred Wegener Institut, Helmholtz Centre for Polar- and Marine Research, Bremerhaven, Germany}

[14]{International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria}

[15] {Climate Research Division, Atmospheric Sci. & Tech. Directorate,S & T, Environment Canada Toronto, Ontario, Canada}

[16]{Swedish Meteorological and Hydrological Institute (SMHI), SE-60176 Norrköping, Sweden}

[17] {National Oceanic and Atmospheric Administration Pacific Marine Environmental Laboratory, Seattle, WA, USA}

[18]{Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia, Canada}

[19]{School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada}

[20]{Department of Meteorology, COMSATS Institute of Information Technology, Islamabad, Pakistan}

* - now at Univ. Grenoble Alpes/CNRS, Laboratoire de Glaciologie et Géophysique de l'Environnement (LGGE), 38041 Grenoble, France

Correspondence to: S. Eckhardt (sabine.eckhardt@nilu.no)

Abstract

The concentrations of sulfate, black carbon (BC) and other aerosols in the Arctic are characterized by high values in late winter and spring (so-called Arctic Haze) and low values in summer. Models have long been struggling to capture this seasonality and especially the high concentrations associated with Arctic Haze. In this study, we evaluate sulfate and BC concentrations from eleven different models driven with the same emission inventory against a comprehensive pan-Arctic measurement data set over a time period of two years (2008-2009). The set of models consisted of one Lagrangian particle dispersion model, four chemistry-transport models (CTMs), one atmospheric chemistry-weather forecast model and five chemistry-climate models (CCMs), of which two were nudged to meteorological analyses and three were running freely. The measurement data set consisted of surface measurements of equivalent BC (eBC) from five stations (Alert, Barrow, Pallas, Tiksi and Zeppelin), elemental carbon (EC) from Station Nord and Alert and aircraft measurements of refractory BC (rBC) from six different campaigns. We find that the models generally captured the measured eBC or rBC and sulfate concentrations quite well, compared to previous comparisons. However, the aerosol seasonality at the surface is still too weak in most models. Concentrations of eBC and sulfate averaged over three surface sites are underestimated in winter/spring in all but one model (model means for January-March underestimated by 59% and 37% for BC and sulfate, respectively), whereas concentrations in summer are overestimated in the model mean (by 88% and 44% for July-September), but with over- as well as underestimates present in individual models. The most pronounced eBC underestimates, not included in the above multi-site average, are found for the station Tiksi in Siberia where the measured annual mean eBC concentration is three times higher than the average annual mean for all other stations. This suggests an underestimate of BC sources in Russia in the emission inventory used. Based on the campaign data, biomass burning was identified as another cause of the modelling problems. For sulfate, very large differences were found in the model ensemble, with an apparent anti-correlation between modeled surface concentrations and total atmospheric columns. There is a strong correlation between observed sulfate and eBC concentrations with consistent sulfate/eBC slopes found for all Arctic stations, indicating that the sources contributing to sulfate and BC are similar throughout the Arctic and that the aerosols are internally mixed and undergo similar removal. However, only three models reproduced this finding, whereas sulfate and BC are weakly correlated in the other models. Overall, no class of models (e.g., CTMs, CCMs) performed better than the others and differences are independent of model resolution.

1 Introduction

Aerosols are important climate forcers (Ramanathan and Carmichael 2008; Myhre et al., 2013), but the magnitude of their forcing is highly uncertain and depends on altitude, position relative to clouds, the surface albedo and the optical properties of the aerosol as well as cloud indirect effects. While absorbing aerosols such as black carbon (BC) are likely to increase climate warming (Shindell and Faluvegi, 2009), scattering aerosols such as sulfate have a cooling effect (Myhre et al., 2013). In addition to atmospheric radiative forcing, deposition of absorbing aerosols on snow or ice reduces the albedo and can thus induce faster melting and efficient surface warming (Jacobson, 2004; Flanner et al., 2009). The highly reflective surfaces of snow and ice as well as strong feedback processes make the Arctic a region of particular interest for aerosol research (Quinn et al., 2008).

The Arctic aerosol consists of a varying mixture of sulfate and organic carbon (OC), as well as ammonium, nitrate, BC and mineral dust (Quinn et al., 2007; Brock et al., 2011). Aerosols in the Arctic feature a strong annual cycle with a late winter/spring peak (the so-called Arctic Haze) and a summer minimum. Increased transport during the cold season (Stohl, 2006) and increased removal by wet deposition during the warm season can explain this annual variation (Shaw, 1995; Law and Stohl, 2007) and also shape the aerosol size distribution (Tunved et al., 2013).

Models have for a long time struggled to capture the distribution of aerosols in the Arctic (Shindell et al., 2008; Koch et al., 2009). The concentrations of BC during the Arctic Haze season in particular were underestimated, in some cases by more than an order of magnitude (Shindell et al., 2008), whereas summer concentrations were sometimes overestimated. The simulated aerosol seasonality is strongly dependent on the model treatment of aerosol removal processes. For instance, changes in the calculation of aerosol microphysical properties, size distribution and removal can change simulated concentrations by more than an order of magnitude in remote regions such as the Arctic (Vignati et al., 2010) and the calculated Arctic BC mass concentrations are very sensitive to parameterizations of BC aging (conversion from hydrophobic to hydrophilic properties) and wet scavenging (Liu et al., 2011; Huang et al., 2010).

The seasonal decrease of aerosol concentrations from winter to summer in the Arctic is likely also due to the different efficiency of scavenging by different types of clouds. There is a

transition from inefficient ice-phase cloud scavenging in winter to more efficient warm cloud scavenging in summer, and there is also the appearance of warm drizzling cloud in the late spring and summer boundary layer. Including these processes in one model clearly improved its performance both in terms of absolute concentrations as well as seasonality for sulfate and BC (Browse et al., 2012). This result is in agreement with the observation-based findings that scavenging efficiencies are increased in summer both for light-scattering (of which sulfate is an important component) as well as for light-absorbing (of which BC is an important component) aerosols (Garrett et al., 2010, 2011). Another modeling problem may be excessive convective transport and underestimation of the associated wet scavenging in convective clouds, which can lead to model overestimates of BC in the upper troposphere and lower stratosphere (Allen and Landuyt, 2014; Wang et al., 2014). Despite remaining difficulties, simulations of Arctic aerosols with many models have improved considerably in the last few years by updating the model treatment of some or all of the above mentioned processes (Fisher et al., 2011; Breider et al., 2014; Sharma et al., 2013; Lund and Berntsen, 2012; Allen and Landuyt, 2014).

Remaining problems may also be due to missing emission sources or incorrect spatial or temporal distribution of emissions in the inventories used for the modeling. The main sources of BC are biomass burning and incomplete combustion of fossil fuels and biofuels (Bond et al., 2004). Sulfate aerosols are formed by sea spray or originate from natural sources such as oxidation of dimethyl sulfide (DMS) or volcanoes. It is also produced from oxidation of SO_2 emitted when sulfur-containing fossil fuels are burned or by metal smelting. Studies based on observed surface concentrations repeatedly suggest that the main source regions for Arctic BC and sulfate are located in high-latitude Eurasia (e.g., Sharma et al., 2006, Eleftheriadis 2009, Hirdman et al., 2010). Stohl et al. (2013) suggested that gas flaring in high-latitude Russia is an important source of BC which is missing from most inventories. In their simulations, BC emissions from gas flaring accounted for 42% of the annual mean BC surface concentrations in the Arctic. However, they also noted the large uncertainty of the gas flaring emissions.

The radiative effects of aerosols are not so much determined by the surface concentrations but by the column loadings as well as the altitude distribution of the aerosol (Samset et al., 2013; Samset and Myhre, 2011). Nevertheless, in the past, model results for the Arctic were evaluated mainly against surface measurements due to their availability over long time periods. However, surface concentrations are not representative of concentrations aloft, which are controlled, at

least in part, by different source regions and different processes. It is therefore important to evaluate models not only against surface measurements but also using vertical profile information.

The purpose of this study is to explore the capabilities of a range of chemistry transport models (CTMs) and chemistry climate models (CCMs) widely used to simulate the Arctic aerosol concentrations. The models use a common emission inventory, which includes gas flaring emissions and provides monthly resolution of the domestic burning emissions. Differences between their modeled aerosol concentrations are therefore solely due to differences in the simulated transport, aerosol processing (e.g., sulfate formation, BC aging) and removal. We concentrate our investigations on BC and sulfate, for which we collected data from six surface stations and five aircraft campaigns in the Arctic.

2 Methods

2.1 Measurement data

We have collected measurements of BC performed with different types of instruments, and these measurements may not always be directly comparable. Following the nomenclature of Petzold et al. (2013), we refer to measurements based on light absorption as equivalent BC (eBC), measurements based on thermal-optical methods as elemental carbon (EC) and measurements based on refractory methods as refractory BC (rBC). All these data are compared to each other as far as possible and to modeled BC values.

Aerosol light absorption data were obtained from five sites in different parts of the Arctic: Alert, Canada (62.3°W, 82.5°N; 210 m above sea level (asl)), Zeppelin/Ny Ålesund, Spitsbergen, Norway (11.9°E, 78.9°N; 478 m asl), Tiksi, Russia (128.9° E, 71.6°N; 1 m asl), Barrow, Alaska (156.6°W, 71.3°N; 11 m asl) and Pallas, Finland (24.12°E, 67.97°N; 565 m asl). The locations of these measurement stations are shown in Fig. 1. Different types of particle soot absorption photometers (PSAPs) were used for the measurements at Barrow, Alert, and Zeppelin, a multi-angle absorption photometer was used at Pallas (Hyvärinen et al., 2011), and an aethalometer was used at Tiksi. All these instruments measure the particle light absorption coefficient σ_{ap}, each at its own specific wavelength (typically at around 530–550 nm), and for different size fractions of the aerosol (typically particles smaller than 1, 2.5 or 10 μm are sampled at different humidities). Conversion of σ_{ap} to eBC mass concentrations is not straightforward and requires certain assumptions (Petzold et al., 2013). The mass absorption efficiency used for conversion

can be specific to a site, the instrument and the wavelength used and is uncertain by at least a factor of two. For Tiksi, the conversion is done internally by the aethalometer. For the other sites, a mass absorption efficiency of $10 m^2 g^{-1}$, typical of aged BC aerosol (Bond and Bergstrom, 2006), was used.

For Barrow, Alert, Pallas and Zeppelin eBC data were available for the years 2008-2009 and could be compared directly with model data which were available for the same period. At Tiksi, the measurements started only in 2009 and thus measured values for the period July 2009 to June 2010 were compared with modeled values for the year 2009.

Barrow and Alert data are routinely subject to data cleaning, which should remove the influence from local sources. The Tiksi data has been quality controlled as well and episodes of local pollution have been removed. Zeppelin generally is not strongly influenced by local emissions; however, summer values are enhanced by some 11% due to local cruise ship emissions (Eckhardt et al., 2013). Thermo-optical measurements of EC were available from Station Nord, Greenland (16.67°W, 81.6°N; 30 m asl) and from Alert. At Station Nord, weekly aerosol samples were collected during 2008-2009 and the EC/OC filter samples at Alert were collected as bi –weekly integrated samples For Station Nord a Digitel DHA 80 high volume sampler (HVS, Digitel/Riemer Messtechnik, Germany) was used for PM10. Both stations' samples were analyzed with a thermal-optical Lab OC/EC instrument from Sunset Laboratory Inc (Tigard, OR, USA). Punches of 2.5 cm^2 were cut from the filters sampled at Station Nord and analyzed according to the EUSAAR-2 protocol (Cavalli et al., 2010). The samples from Alert were analyzed by using EnCan-total-900 thermal method originally developed by carbon isotope analysis for OC/EC (Huang et al., 2006) and further optimized (Chan et al., 2010)."

Sulfate measurement data were available from the stations Pallas, Zeppelin, Barrow, Nord and Alert. The sulfate data were obtained on open face filters and cations and anions were subsequently quantified by ion chromatography. Non-sea salt (nss) sulfate concentrations were obtained by subtracting the sea salt contribution via analysis of Na^+ and Cl^- data, thus making the sulfate data directly comparable to the modeled nss-sulfate values. For Station Nord, the contribution from sea salt is only minor (Heidam 2004), no correction was applied there. Samples were taken with daily to weekly resolution, depending on station and season.

Aircraft data were obtained from several campaigns. In the framework of POLARCAT (Polar Study using Aircraft, Remote Sensing, Surface Measurements, and Models of Climate

Chemistry, Aerosols, and Transport; Law et al., 2014), two ARCTAS (Arctic Research of the Composition of the Troposphere from Aircraft and Satellites) campaigns in April and June/July 2008 with a DC-8 aircraft covered mainly the North American Arctic (Jacob et al., 2010). The ARCPAC (Aerosol, Radiation, and Cloud Processes affecting Arctic Climate; Brock et al., 2011) campaign was conducted from Alaska together with ARCTAS in April 2008. The PAMARCMiP (Polar Airborne Measurements and Arctic Regional Climate Model Simulation Project) campaign covered the entire western Arctic in April 2009 (Stone et al., 2010). Two HIPPO (High-Performance Instrumented Airborne Platform for Environmental Research Pole-to-Pole Observations; Schwarz et al., 2010; Schwarz et al., 2013; Wofsy et al., 2011) campaigns during January 2009 and October 2009 explored the North American Arctic. Flight legs north of 70°N for all of these campaigns are shown in Fig. 1. Refractory BC (rBC) was measured during these campaigns with single particle soot photometer (SP2) instruments (Kondo et al., 2001; Schwarz et al., 2006). Observations of submicrometer aerosol sulfate mass during ARCTAS were made with a particle-into liquid-Sampler (PILS) (Sullivan et al, 2006) coupled to an ion chromatograh. Sulfate measurements during ARCPAC were made with a compact time-of-flight aerosol mass spectrometer (Bahreini et al., 2008).

During April 2008 agricultural and boreal biomass burning influence was widespread throughout the Arctic (Warneke et al., 2010; Brock et al., 2011) and ARCTAS and ARCPAC often targeted these fire plumes. Anthropogenic pollution from Asia was also sampled by these campaigns in the western Arctic, particularly in the mid-upper troposphere (see Law et al., 2014 and references therein). Pollution from Europe also made a significant contribution in the lower troposphere. In contrast, PAMARCMiP and HIPPO sampled the Arctic atmosphere at times with little influence from biomass burning and also did not target pollution plumes. Thus, the higher mean rBC concentrations found during ARCTAS and ARCPAC than during PAMARCMiP a year later are caused both by the sampling strategy of these campaigns as well as the early start of the biomass burning season in 2008. Even though all available rBC and sulfate data from several campaigns were used for model evaluation, the data coverage and representativity for the Arctic as a whole must still be considered as rather poor. The Eastern Arctic, in particular, was not sampled by any campaign.

ARCTAS-B was the only summertime POLARCAT campaign to make detailed measurements of BC and sulfate (Jacob et al., 2010). These flights focused mainly on boreal fires over Canada in July 2008 but several flights into the high Arctic sampled, for example Asian pollution close

to the North Pole (Sodemann et al., 2010). Plumes of Asian origin were also sampled in the upper troposphere over Canada (Singh et al., 2010).

2.2 Emissions

All models made use of an identical emission dataset, the ECLIPSE (Evaluating the Climate and Air Quality Impacts of Short-Lived Pollutants) emission inventory version V4a (Klimont et al., 2015a, 2015b). The ECLIPSE inventory was created using the GAINS (Greenhouse gas – Air pollution Interactions and Synergies) model (Amann et al., 2011), which provides emissions of long-lived greenhouse gases and shorter-lived species in a consistent framework. The proxies used in GAINS are consistent with those applied within the RCP (Representative Concentration Pathways) projections as described in Lamarque et al. (2010) and as further developed within the Global Energy Assessment project (GEA, 2012). They were, however, modified to accommodate more recent information where available, e.g., on population distribution and open biomass burning, effectively making them year specific (Riahi et al., 2012; Klimont et al., 2013). Emissions for the years 2008 and 2009 were lumped into the following source categories: industrial combustion, residential combustion, energy production, transport, agriculture, waste treatment, shipping, agricultural waste burning and gas flaring. All emission data were gridded consistently to a resolution of 0.5°x0.5°. Monthly disaggregation factors were provided for the domestic heating emissions, based on ambient air temperatures. For a more detailed description of the ECLIPSE emission data set, see Klimont et al. (2015a, 2015b). A detailed description of the high-latitude emissions in the ECLIPSE inventory and comparisons with other emission inventories can be found in AMAP (2015).

Non-agricultural biomass burning emissions were not available through GAINS and were therefore taken from the Global Fire Emission Database (GFED), version 3.1 (van der Werf et al., 2010). No attempt was made to harmonize sulfur emissions from volcanic sources or the ocean, which could explain some differences in simulated sulfate concentrations.

2.3 Models

We show results of 11 different models, whose main characteristics and references are summarized in Table 1. In principle we are using two types of atmospheric models: off-line models and on-line models. Both model types have certain advantages and disadvantages. Off-line models based on meteorological re-analysis data can capture actual meteorological

situations, thus facilitating a direct comparison of measured and modelled aerosol quantities. Often, they also have higher resolution than the on-line global models. However, off-line models cannot be used for predictions and the off-line coupling can also cause inaccuracies in the treatment of transport, chemistry and removal processes. The global on-line models in our study are free-running and thus produce their own model climate, which means that they cannot reproduce a given meteorological situation. Nevertheless, their modelled climate should correspond to the current climatic conditions and, thus, seasonally averaged quantities (i.e., averages over many different meteorological situations) should be comparable to measured quantities. The main advantage of the on-line models is that they can also be used for predictions.

Further, there were two different types of off-line models used, namely Eulerian chemistry transport models (CTMs) and one Lagrangian particle dispersion model (LPDM). Our on-line models were climate chemistry models (CCMs), where a climate model is coupled with a chemistry and aerosol module. We also use one global climate model coupled with an aerosol module which, however, does not simulate atmospheric chemistry. We refer to this as an aerosol climate model (ACM) to distinguish it from the CCMs. Furthermore, we use one regional weather forecast model coupled on-line with a chemistry model (WRF-Chem). This model is similar to the CCMs, but only used for regional simulations and it is designed for short-term simulations rather than simulations over climate time scales. WRF-Chem is also nudged towards re-analysis data and therefore can capture actual meteorological situations, similarly to the off-line models.

The horizontal resolution of the individual models ranges from about 0.6°x0.8° to 2.8°x2.8°. We use one Lagrangian particle transport model, FLEXPART (Flexible Particle Dispersion Model), which is run in backward mode for 30 days (thus, older source contributions are not accounted for). The simulation is driven by 1°x1° operational analyses from the European Centre for Medium Range Weather Forecast (ECMWF). The OsloCTM2, TM4-ECPL (Tracer Model version 4 - Environmental Chemical Processes Laboratory) and SMHI MATCH (Swedish Meteorological and Hydrological Institute Multi-scale Atmospheric Transport and Chemistry Model) are CTMs and also use meteorological data from ECMWF (for details see table 1). The DEHM (Danish Eulerian Hemispheric Model) CTM is driven by NCEP (National Centers for Environmental Prediction) meteorological data. WRF-Chem (Weather Research and Forecasting Model coupled with Chemistry) is an on-line atmospheric chemistry-weather

forecast model which was nudged to NCEP FNL (final analysis) data for this study. The aerosol-climate model (ACM) ECHAM6-HAM2 (for brevity, referred to as ECHAM6 in figures) is the European Centre for Medium-Range Weather Forecasts Hamburg model version 6 (Stevens et al., 2013) extended with the Hamburg aerosol module version 2 (HAM2) (Zhang et al., 2012). ECHAM6-HAM2 and the CCMs including HadGEM3 (Met Office Hadley Centre Climate Model, version 3) and CanAM4.2 (Canadian Atmospheric model, version 4.2) were nudged to ECMWF data. CESM1-CAM5.2 (Community Earth System Model version 1 – Community Atmosphere model version 5.2) and NorESM1-M (Norwegian Earth System Model version 1 with intermediate resolution and used here in a version where aerosols are fully coupled with a tropospheric gas-phase chemistry scheme, hereafter referred to as NorESM) are also CCMs but were running freely, thus producing their own meteorological data. These latter models cannot be compared point-to-point with the measurement data because they produced meteorological conditions which were different from the actual ones; however, longer-term (e.g., seasonal) medians should still be comparable with the measurements, especially since sea surface temperatures (SST) and sea-ice extent were prescribed and specific to the years 2008-2009. All models were sampled exactly at the locations of the measurement stations and along the flight tracks at the highest possible (mostly hourly) temporal resolution. Notice that not all models simulated the full 2008-2009 period and FLEXPART only simulated BC.

3. Simulated BC and sulfate concentrations

Figure 2 shows the simulated BC and sulfate column mass loadings as a function of latitude for the time periods of the Arctic Haze (March) and the much cleaner summer (July) in the Arctic, for the models for which this information was available. For BC in March, most models show a maximum near 20°N, with some models extending this maximum to 40°N. This approximately covers the latitude range with the highest global emissions where the models agree at least within a factor of two in their simulated column loadings. In contrast, larger differences between the models are found in the Arctic, where column mass loadings vary by more than an order of magnitude. Similar results are also found for sulfate in March, for which most models also show a maximum around 20-40°N; however, compared to BC, the models show a less pronounced decrease towards higher latitudes and two models even simulate increasing sulfate burdens with latitude. The relatively good agreement between the models in

the BC and sulfate source region latitudes is not surprising, given that they all use the same emission data set. In contrast, the differences between the atmospheric column loadings in the Arctic must mainly be due to differences in the aerosol processing and removal and hence aerosol lifetimes, and probably differences in atmospheric transport. Most models with relatively low BC column loadings in the Arctic also have low sulfate loadings there, indicating similarities in the simulated removal of these two types of aerosols. A notable exception, however, is HadGEM3, which has moderately low BC but the highest sulfate loadings in the Arctic.

In July, the BC column loadings show a double peak in the southern tropics and northern subtropics. The southern tropical peak is due to the migration of the inter-tropical convergence zone (ITCZ) into the northern hemisphere, which leads to less efficient wet removal and dry conditions favoring biomass burning in the southern tropics. On the other hand, BC concentrations near 10°N show a deep minimum, due to the efficient wet removal near the ITCZ. Most models show a third peak in BC loading near 60°N, which results from open vegetation fires in the boreal region. North of 60°N, the BC loadings decline rapidly towards the North Pole. The sulfate column loading distribution in July lacks the peaks in the southern tropics and the boreal region because biomass burning is not a strong source of sulfate. HadGEM3 stands out against the other models even more than in spring, as its polar sulfate loadings are more than a factor of five higher than those of all other models, which show a smooth decrease with latitude north of 40°N.

In the simulated surface BC and sulfate mass mixing ratios the same basic patterns are found as in the column loadings, but with enhanced gradients between source areas and remote regions (Fig. 3). When looking at individual models, there are, however, notable differences for sulfate. ECHAM6-HAM2 has the highest sulfate surface mass mixing ratios of all models, especially in the northern hemisphere subtropics and mid-latitudes. Combined with the rather "normal" column sulfate loadings of this model, this indicates that ECHAM6-HAM2 does not transport sulfate away from the surface as quickly as the other models. On the other hand, HadGEM3, which has by far the largest sulfate column loadings, has the smallest surface concentrations. This deficiency was due to the implementation of the Global Model of Aerosol Processes (GLOMAP; Mann et al., 2010), which in this HadGEM3 version resulted in too little removal of the sulfate precursor SO_2 during the venting from the boundary layer to the free troposphere. The longer sulfate lifetime there explains the high column loadings.

In summary, we find that the Arctic is a region with particularly large relative differences between the models, both for the surface mass mixing ratios (with differences of more than an order of magnitude) as well as for the column loadings, and both for BC and sulfate. This result must be related to differences in aerosol removal and lifetimes in the different models. We also found that, especially for sulfate, there can be an anticorrelation between simulated surface concentrations and column loadings. Hence there is a strong motivation to evaluate the models' performance in the Arctic, based on measurements taken both at the surface and aloft.

4 Observed and simulated BC and sulfate seasonality at Arctic surface measurement stations

We start our discussion of the annual cycles of aerosol concentrations with the example of BC at the Zeppelin station in Spitsbergen (Fig. 4). Monthly medians as well as the 25th and 75th percentile are calculated for every month based on hourly data for the two years 2008 and 2009. Maximum median eBC concentrations of 46 and 53 ng/m^3 occur in March and April, while summer median values are only 2 to 3 ng/m^3. Some of the models reproduce this seasonality with high winter/spring values and much lower summer values quite well, although in most of these models BC reaches its highest values already in January. Only the CanAM4.2 model seems to capture the observed spring maximum. All models except WRF-Chem capture that summer is having the lowest values of the year. OsloCTM2, TM4-ECPL and NorESM have smaller annual variation than observed. HadGEM3, which we have seen to produce lower BC surface concentrations than the other models in Fig. 3, strongly underestimates the measured eBC concentrations throughout the year. The variability of the modeled values within a month (described by the height of the bars) shows clear differences between the models. For instance, CESM1-CAM5.2 simulates much less variable BC concentrations than CanAM4.2 and DEHM, or the measurements.

The eBC mass concentrations at the three other sites in the western Arctic (Alert, Barrow, Pallas) are quite comparable to those at Zeppelin station, with monthly median values of about 20-80 ng/m^3 in late winter/early spring and of less than 10 ng/m^3 in summer/early fall (see Fig. 5). One exception is EC measured at Station Nord, which in summer is higher than eBC measured at the other sites. At Alert, where both eBC and EC data are available, EC values in

summer are also somewhat higher than eBC values (although lower than the Station Nord EC values), probably due to systematic differences in measurement techniques.

At the Tiksi station, which is closer to the main source regions of Arctic BC in high-latitude Eurasia (Hirdman et al., 2010), higher monthly median eBC values were measured (more than 100 ng /m^3 in winter/spring, about 20-40 ng /m^3 in summer) and the annual mean (81 ng/m^3) is 2.5 times higher than the average for the other stations (31 ng/m^3). The seasonality of measured eBC is strongest at Alert where the summer concentrations are very low, but the winter/spring concentrations are similar to the other sites in the western Arctic. This result points to a deepening of the seasonal minimum with latitude. While the aerosol concentrations in the Arctic during late winter/early spring are comparable to remote regions further south, the concentrations in summer/early fall are lower because of the effective cleansing of the atmosphere (Garrett et al., 2010, 2011; Browse et al., 2012; Tunved et al., 2013) and less efficient transport from source regions (Stohl, 2006). The highest eBC concentrations were observed in January (Alert), February (Barrow), March (Pallas, Tiksi) or April (Zeppelin), with no clear dependence of the time of the maximum on latitude; however, the maximum occurred earlier at the two North American sites than at the other sites.

The models capture the Arctic BC concentrations with variable success (Fig. 5). Most models capture the much higher concentrations in winter/spring than summer/fall, and some models can approximately reproduce the concentrations reached during the Arctic Haze season (see also Breider et al., 2014). However, as already seen for the Zeppelin station (Fig. 4) and the annual mean surface mass mixing ratios (Fig. 3), there is a large variability between individual models, with seasonal median values varying by about an order of magnitude both in spring and summer even when excluding the most extreme models (see also Table 2). Seasonal mean concentrations during January to March are underestimated by up to a factor of 27 for individual models and by more than a factor 2 for the mean over all models, and only one model slightly overestimates the measured concentrations (Table 2). Nevertheless, this indicates clear progress since earlier studies (e.g. Shindell et al., 2008; Koch et al., 2009; AMAP, 2011), where it was reported that most models had a completely wrong seasonality and systematically underpredicted the Arctic Haze concentrations. For instance, in Shindell et al. (2008), none of their models came close to the measured concentrations at Barrow and Alert during winter and spring, with a model-mean underestimate of about one order of magnitude (their Fig. 7). It is also important to keep in mind that the eBC measurements are uncertain and could be biased

high. However, EC and eBC values at Alert are very similar and we find a similar model underestimate of measured EC at Station Nord as well.

Our finding that Arctic BC concentrations in the spring tend to be underestimated by our models implies that these models would also underestimate radiative forcing by BC in the Arctic. This is particularly important because spring is the season when both aerosol concentrations are large and solar radiation is abundant. Furthermore, it is the season when feedback processes, e.g., via ice and snow melting, are most important (Quinn et al., 2008). The concentrations of BC in summer are much lower than in spring, so even with more abundant solar radiation modelling problems in summer would have a relatively small effect on radiative forcing.

In contrast, five models overpredict the low concentrations in summer, the most extreme model by an order of magnitude (Table 2). Some models (e.g., HadGEM3) underpredict strongly throughout the year. For the sites in the western Arctic, the model deficiencies become worse with increasing latitude. For instance, at the northernmost site, Alert (82.5°N), all models underpredict for the full duration of the Arctic Haze season from January until April.

For Tiksi, the data comparison is less direct as measurement data from July 2009-June 2010 were used. Nevertheless, it is clear that except for CanAM4.2 (which produces the highest modeled values at most sites) the models strongly underpredict for this site, especially in winter/spring. The most likely explanation for this is that the BC emissions in high-latitude Russia are underestimated in the ECLIPSE inventory. It is difficult to know where exactly the missing sources are located. However, we find that in the ECLIPSE inventory the BC emissions in Norilsk (88.2°E, 69.3°N; population 170000) are zero. We do not suggest that Norilsk emissions are responsible for the strong underestimation of BC concentrations at Tiksi, but these discrepancies (and others for sulfur emissions discussed later) suggest that the high-latitude Russian pollutant emissions are underestimated and/or wrongly placed in the ECLIPSE inventory. Similar problems likely occur with most other global emission inventories. For instance, AMAP (2015) compared the ECLIPSE emission data set with 10 other inventories and found that the differences between the different inventories grow with latitude and are largest north of 70ºN (i.e., high-latitude Eurasian emissions).

The seasonal cycle of sulfate at the monitoring stations is similar to that of eBC, with a clear maximum during the Arctic Haze season and a minimum in summer/early fall (Fig. 6). However, the seasonal cycle at the northernmost stations is less strong than for eBC, with about a factor of 5 difference between spring and summer, compared to a factor of 15 for eBC (Table

2). This is probably due to the influence of biogenic sources of sulfate in summer (Quinn et al., 2002) and/or a weaker seasonality in the emissions (e.g., smelter emissions of SO_2 are probably relatively constant throughout the year).

The models have similar difficulties capturing the sulfate seasonality as they have for BC. Again, there is up to more than an order of magnitude difference between simulated seasonal median concentrations from different models, both in summer and in winter (Table 2). The model differences in summer are in fact even larger than for BC, probably related to different treatment of natural sources, especially dimethyl sulfide emissions from the Arctic Ocean. There is a tendency for models that strongly underestimate BC concentrations to also underestimate sulfate (e.g., HadGEM3 model) but the correlation between the two simulated species from the different models is quite low, especially in summer. For instance, ECHAM6-HAM2 underestimates BC by factors of 26 and 1.6 in winter and summer, but underestimates sulfate only by about 13% in winter and even overestimates sulfate by a factor of 3.8 in summer (see Table 2). As seen in Fig. 2 and 3, ECHAM6-HAM2 simulates relatively high surface concentrations of sulfate but low total column loadings, both at source and Arctic latitudes.

The models generally underpredict sulfate most strongly at the northernmost station (Alert), which is consistent with the BC results (compare Figs. 5 and 6). The CanAM4.2 model, which had some of the highest BC concentrations, also gives the highest sulfate values (Table 2). It is the only model that matches the high measured sulfate values at Alert and Station Nord in spring. The reason why CanAM4.2 captures the spring peak better might be that this model has a less efficient removal through wet deposition under stratiform condition compared to the other models (Mahmood et al., 2015 submitted).

At Pallas, the lowest-latitude station in this comparison, most models severely underestimate sulfate throughout the year (Fig. 6), although they tend to overestimate BC in spring there. One likely reason for the sulfate underestimation is the proximity of the Pallas station to the Kola peninsula, where metal smelters are a strong source of sulfur. According to AMAP (2006), SO_2 emissions in Nikel, Zapolyarnyy and Monchegorsk together were about 170 kt/yr in the year 2002. In the ECLIPSE version 4a inventory used for this study the SO_2 emissions in these areas are only about 33 kt/yr in total for the year 2005. Similar deficiencies were in fact reported also for other emission inventories for this region (Prank et al., 2010). Strong underestimation of the SO_2 emissions from metal smelting in the Kola peninsula is therefore a likely explanation for why almost all models underestimate sulfate at Pallas so strongly. Similar discrepancies were

in fact found for SO_2 emissions in Norilsk, prompting a regridding of the ECLIPSE emissions (now available version 5a) using better location information for the metal smelting industry.

5 Vertical Profiles

Figure 7 summarizes all rBC data from the ARCTAS and ARCPAC campaigns in spring 2008. Median concentrations are shown as a function of latitude (binned into 10° intervals) both for lower (<3 km) and higher (>3 km) altitudes, and as a function of altitude both for the high Arctic (>70°N) and lower latitudes. As the campaigns focused on the Arctic, data south of 60°N are scarce and limited to North America. The models were sampled in their grid box containing a measurement location and at the time of a measurement and were subsequently binned in the same way as the measurement data to allow a direct comparison. For the free-running climate models, the same procedure was used, albeit with the caveat that the simulated meteorological situation at the measurement time does not correspond to the real conditions.

For the low-altitude (<3 km) bin, the highest median rBC values were measured (see 2nd from top row of panels in Fig. 7) at 35°N and 55°N, with a substantial concentration drop towards higher latitudes. The mid-latitude maximum reflects the location of the BC sources in North America, where ARCTAS and ARCPAC were conducted. Above 3 km (top row of panels in Fig. 7), the highest median rBC concentrations were measured further north, at 60°N, and the concentrations drop less strongly towards the North Pole than at lower altitudes. This is due to quasi-isentropic lifting occurring together with northward transport (Stohl, 2006). All models, except CanAM4.2, systematically underestimate the measured values for both altitude bins and for all latitudes, and they also underestimate the measured rBC variability. However, most of the models simulate a decrease of the concentrations with latitude that is consistent with the measured latitude dependence.

When plotted as a function of altitude (two bottom panel rows in Fig. 7), the measured values peak in the 4-5 km altitude bin, both for sub-Arctic and Arctic latitudes. The models, except for CanAM4.2, underestimate the measured median values throughout the entire depth of the profile. Some of the models, mainly those driven by observed meteorology, capture the rBC maximum in the mid-troposphere in the Arctic. However, the lower-latitude 4-5 km maximum is hardly reproduced by any of the models. One likely reason for the modeling problems is the strong biomass burning activity during spring 2008, which influenced a substantial fraction of the measurement data (Warneke et al., 2010; Brock et al., 2011). Even though this should be reflected in the GFED emission data for 2008, it seems possible that the GFED emissions are

underestimated. Furthermore, as some of the flights targeted biomass burning plumes specifically, the influence of the biomass burning may be enhanced in the measurement data compared to the models, especially if the models did not capture the plume transport well enough and thus potentially simulated the biomass burning plumes at other locations than observed. This sampling bias is particularly strong for the CCMs which are not driven by observed meteorological fields.

Comparisons like those shown in Fig. 7 were also performed for the other aircraft campaigns. For the sake of brevity, we further aggregate the data and only show results for latitudes north of 70°N and for median values below and above 3 km altitude (Fig. 8). For spring 2008, the aggregate plots for BC (Fig. 8e-f) show even more clearly than Fig. 7 that all models except CanAM4.2 underestimate the measured rBC concentrations both at low and high altitudes. The spring 2009 PAMARCMiP campaign, however, shows a different picture (Fig. 8c-d). This campaign was influenced very little by biomass burning. The measured median rBC mass concentrations at low (high) altitudes were about a factor two (three) lower than for the spring 2008 campaigns. Most models also simulated lower median BC concentrations than a year earlier but the modeled reductions were less pronounced than the measured ones and, thus, about half of the models under- and the other half overestimated the measured median values. The vertical gradient of measured BC was also different in 2008 and in 2009. While in spring 2008, the concentrations above 3 km were higher than those below, the opposite was true in spring 2009, likely because of the weaker biomass burning influence in 2009. This feature can be seen very clearly in the vertical profiles shown in Fig. 9 and it is not well captured by the models, most of which showed a relatively flat vertical BC distribution.

The concentrations measured by the ARCTAS summer campaign in 2008 are much lower than those measured in spring 2008 and 2009, both at low and high altitudes (Fig. 8g-h), which is in agreement with the seasonality seen at the surface stations. Some of the models under- and others overestimate the measured concentrations, with the majority of the models overestimating, especially below 3 km. The mean values, averaged over all models, are about two (three) times as high as the measurements for altitudes above (below) 3 km. Some of the models reproduce the measured rBC maximum at 6 km (Fig. 9).

The HIPPO campaign in fall 2009 (Fig. 8i-j) was conducted about one month after the seasonal minimum at most surface sites and measured very low rBC mass concentrations, which is

consistent with the surface observations. Most of the models overestimate the measured concentrations throughout the entire vertical profile (Fig. 9).

The HIPPO campaign in January 2009 (Fig. 8a-b) measured strong altitude differences: moderately high rBC mass concentrations up to 3 km, but the lowest concentrations of all campaigns above. This feature is well captured by some of the models (Fig. 9). The lack of high concentrations aloft is likely related to the minimal influence of biomass burning at this time of the year.

Overall, the aircraft measurements confirm the BC seasonality measured at the surface stations. They also confirm that most models underestimate the concentrations in spring (at least for the year 2008) but many models overestimate the concentrations in summer and fall. It thus seems that models produce a too weak BC seasonality throughout the depth of the troposphere. However, for the year as a whole there is a tendency towards model overestimates, in contrast to the surface sites. Even stronger model overestimates downwind of Asia over the Pacific, especially in the upper troposphere, were recently reported by Samset et al. (2014) who suggested that the BC lifetime in the models is too long. However, a uniform reduction of BC lifetime in our models would lead to strong underestimates of the BC concentrations at the Arctic measurement stations. Even our Arctic aircraft comparisons only support at most a very moderate BC lifetime reduction. Of course, regional and/or vertical differences in the model lifetime biases or excessive convective uplift could explain the contrasting findings of our study and Samset et al. (2014).

For sulfate, measured median concentrations in the Arctic during spring 2008 were lower above 3 km than below 3 km (Fig. 10a-b). All models, except CanAM4.2, strongly underestimate the measured sulfate concentrations, some models by more than an order of magnitude. This is consistent with the findings from the surface station comparisons (Fig. 6, Table 2). The models also do not give a consistent picture of the vertical distribution of sulfate, with some models correctly simulating lower concentrations above 3 km than below but others giving the opposite result. The model underestimates for sulfate are likely not related to a sampling bias towards frequent encounters of biomass burning plumes, as biomass burning plumes are relatively poor in sulfate (e.g. Brock et al., 2011). Instead, the underestimation suggests other missing sulfur sources or a too quick removal of sulfate from the atmosphere. Indeed, the latter would be consistent with the suggestion of Kristiansen et al. (2012) that sulfate lifetimes in models are too short in spring.

During summer 2008 (Fig. 10c-d), the measured median sulfate concentrations were about a factor of 4-6 lower than in spring 2008, consistent with the seasonality measured at surface sites. Median concentrations above and below 3 km are very similar. The models have very large differences in their simulated sulfate concentrations, with some models over- and others underestimating the measured concentrations in summer. This is again consistent with the findings from the surface site comparison (Fig. 6, Table 2).

6 Station vs. low-altitude aircraft measurements

Contrary to the year-round station measurement programs, the aircraft campaigns sample the atmosphere only during limited time periods and their representativeness with regard to climatological means may be questioned. Furthermore, from the aircraft measurements we have seen that spring 2008 and 2009 had very different measured rBC concentrations, and modeling problems were larger for spring 2008 when there was intensive biomass burning influence in the Arctic. A valid question is therefore whether the surface measurements show the same differences between 2008 and 2009.

To investigate how consistent a picture the aircraft campaigns give vis a vis the station measurements, we compare all aircraft data from the lowest 3 km and lowest 1 km to the values obtained from the surface stations for the same months (Fig. 11). Selecting data only for even lower altitudes is problematic as the data coverage becomes very poor. In Fig. 11, we also show the station measurements obtained for the years 2008 and 2009 separately. For eBC, the measurements obtained for the same month at the different stations and during different years are (with a few exceptions such as Barrow in January 2008) quite comparable with each other. In particular, April 2008 did not show higher eBC values than April 2009. This is consistent with the finding that the biomass burning layers in 2008 did not extend to the surface (Brock et al., 2011). At Alert, the EC values are similar to the eBC values, whereas the Station Nord EC values are in summer and fall higher than eBC values at other stations. The aircraft rBC measurements for all campaigns show consistently lower values than the eBC or EC measurements at the ground, except for the HIPPO campaign in January 2009 where, however, the data coverage particularly below 1 km is poor. It is possible that the BC concentrations show a strong gradient in the lowest 1 km and that surface concentrations are indeed systematically higher than concentrations just aloft. However, an alternative explanation could be that the rBC measurements are biased low against the eBC or EC measurements, given the

different measurement techniques used. A direct comparison of all three measurement techniques at the Alert station also suggests a low bias of rBC against eBC and EC concentrations (S. Sharma, personal communication).For sulfate (Fig. 12) the measurements show a much larger variability than for BC, both between stations and between the two different years. For instance, the 25th percentile of the sulfate concentrations at Alert in January 2009 is higher than the 75th percentile of the other stations and also of Alert in January 2008. On the other hand, the sulfate concentrations measured during the two available flight campaigns in spring and summer 2008 are not systematically different from those measured at the stations, although the median concentration in summer 2008 is somewhat lower than at the stations. This is consistent with the eBC or rBC differences.

7 Sulfate/BC correlations

In this section, we perform a correlation analysis of BC and sulfate. Such an analysis allows some insights into the mixing state of the Arctic aerosol. BC and sulfate largely originate from different sources (although some sulfate is co-emitted with BC by combustion processes). A poor correlation between BC and sulfate means that BC and sulfate either arrive at the measurement stations in distinct air masses or that at least the different aerosol types (even if the air masses mix) remain externally mixed and thus are affected to a different and varying extent by removal processes. On the other hand, a strong correlation implies that BC and sulfate arrive in air masses where contributions from their different emission sources are mixed and that, furthermore, also the aerosol must be internally mixed, as otherwise different removal efficiency for BC and sulfate would lead to decorrelation between the two species. Such a correlation analysis has in fact recently also been performed with measurement data from Station Nord (Massling et al., 2015). In our case, we can furthermore compare measured and modeled correlations, allowing some insights into how models treat the mixing of different aerosol types compared to reality.

Figure 13 shows correlation plots between monthly mean sulfate and eBC for the measurements and the models sampled at the different stations. In the observations, sulfate and eBC correlations for Alert, Pallas and Zeppelin are statistically significant at the 99.9% level (Table 3). The slopes of the regression lines shown in Fig. 13 are reported in Table 3. For the observations, they are very similar: 10.1, 8.4 and 8.9 ng[SO_4]/m^3 / ng[eBC] m^3 for Alert, Pallas and Zeppelin, respectively. For Barrow, where the correlation is not significant because of two

eBC-rich outlier data points, the slope is smaller (6.4 ng[SO_4]/ m^3 / ng[eBC]/m^3). The strong correlation between sulfate and eBC and the similarity of the slopes suggests that the sources contributing to the measurements at the different stations are similar and that the removal of sulfate and eBC is highly correlated, which would be expected for internally mixed aged aerosol as is typical for the Arctic.

Most of the models, on the other hand, show much weaker correlation between sulfate and BC and some of the models have no significant correlation at all. Exceptions are DEHM, CESM1-CAM5.2 and WRF-Chem which show mainly significant correlations, and slopes that are comparable at the different stations and which are also quite similar to the observed slopes. This suggests that, with the given emissions, it is possible to reproduce the observed correlations. The lack of correlation between sulfate and BC in the other models – in disagreement with the observations – therefore suggests that they treat the two species differently, probably having a too large fraction of the aerosol as externally mixed. Correlations could also be degraded by a too strong influence of biogenic (dimethyl sulfide) emissions from the oceans or factors influencing SO_2 to sulfate conversion such as the level of oxidants in the models. This could lead to varying fractions of sulfur present as SO_2 and maybe these fractions are more variable in the models than in reality.

Based on the ECLIPSE inventory which is available for BC and for SO2, we estimated ratios between those two substances under the assumption that all SO_2 is converted to sulfate. The SO_2 to BC emission ratio of anthropogenic emissions in the ECLIPSE inventory is 25 globally and 40 north of 50°N. For the GFED biomass burning emissions the emission ratio is only 1.7 globally and 2.5 north of 50°N, and for the sum of anthropogenic and biomass burning emissions, we obtain ratios of 19 globally and 25 north of 50°N. The mean observed slopes of the observations (9.1 ng[SO_4]/m^3 / ng[eBC]/m^3) and the slopes modeled by DEHM (5.4 ng[SO_4]/m^3 / ng[BC]/m^3), CESM1-CAM5.2 (9.9 ng[SO_4]/m^3 / ng[BC]/m^3) and WRF-Chem (8.5 ng[SO_4]/m^3 / ng[BC]/m^3) are much lower than the emission ratio of anthropogenic emissions in the ECLIPSE inventory and they are also lower than the emission ratio for mixed anthropogenic and biomass burning emissions. This suggests that biomass burning emissions are relatively more important in the Arctic than elsewhere, that there are missing BC sources, that sulfur emissions are overestimated (although this is not so likely, given the too low SO_2 emissions in high latitude Russia in the ECLIPSE version 4a inventory used here), and/or that

there exists a mechanism that enriches aerosols in BC relative to sulfate in the Arctic atmosphere. The latter could be related to the hydrophobic nature of freshly emitted BC.

8 Conclusions

Based on our comprehensive study of measured and modelled BC and sulfate in the Arctic, we can draw the following conclusions:

- The simulation of BC concentrations in the Arctic has improved compared to earlier studies (e.g. Shindell et al., 2008; Koch et al., 2009; AMAP, 2011). For instance, our model-mean underestimate of Arctic eBC at Barrow and Alert is about a factor of 2, compared to one order of magnitude reported in Shindell et al. (2008). Nevertheless, the aerosol seasonality at the surface is still too weak in most models. Concentrations of eBC and sulfate averaged over three surface sites in the western Arctic are underestimated in winter/spring in all but one model (model means for January-March underestimated by 59% and 37% for BC and sulfate), whereas concentrations in summer are overestimated in the model mean (by 88% and 44% for July-September), but with over- as well as underestimates present in individual models.

- For the aircraft campaigns, the models overestimated measured rBC during all seasons except for spring and throughout the depth of the troposphere. In spring 2009, no overestimate was found, and in spring 2008 the models underestimated both rBC and sulfate strongly. For rBC, this could have been due to underestimation of the strong influence of biomass burning emissions observed during that campaign. The largest eBC underestimates are found for the station Tiksi, which is closest to potential Russian source regions and where the annual mean eBC concentration is three times higher than the average annual mean for all other stations. This suggests an underestimate of BC sources in Russia in the emission inventory used, even though this inventory contains gas flaring as an important BC source there.

- We found a strong correlation between observed sulfate and eBC with consistent sulfate/eBC slopes for all Arctic stations. This confirms earlier studies that the source regions contributing to sulfate and BC throughout the Arctic are similar (e.g., Hirdman et al., 2010) and that the aerosols are internally mixed and undergo similar removal (e.g., Quinn et al., 2007). However, only three models reproduced this finding, whereas sulfate and BC are weakly correlated in the other models.

- We found that overall, no class of models (e.g., CTMs, CCMs) performed substantially better than the others and model performance did also not depend on resolution. Therefore, differences are largely due to the treatment of aerosol removal in the models.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 282688 – ECLIPSE. Some of the work was conducted for and funded by the Arctic Monitoring and Assessment Programme (AMAP). French authors also acknowledge support from the CLIMSLIP-ANR project and computer resources provided by IDRIS HPC resources under the allocation 2014-017141 under GENCI. Contributions by SMHI were funded by the Swedish Environmental Protection Agency under contract NV-09414-12 and through the Swedish Climate and Clean Air research program, SCAC. Simulations with CanAM4.2 were supported by the Network on Climate and Aerosols: Addressing Key Uncertainties in Remote Canadian Environments (NETCARE), with partial funding from the Natural Sciences and Engineering Research Council of Canada (NSERC). This is PMEL contribution number 4276. ECMWF gave access to their meteorological data. Environment Canada provided the sulfate data and eBC data. Shao-Meng Li (Environment Canada) provided PAMARCMIP BC Dataset obtained by the EC system (SP2). We thank Stockholm University (P. Tunved) for eBC data from Zeppelin, and all contributors to the ARCTAS, ARCPAC, HIPPO, and PAMARCMiP campaigns. HIPPO data products were downloaded from http://hippo.ornl.gov/dataaccess. Julia Schmale is acknowledged for valuable discussion. We thank the two anonymous reviewers for their comments and suggestions.

References

Allen, R. J., and Landuyt, W.: The vertical distribution of black carbon in CMIP5models: Comparison to observations and the importance of convective transport, Journal of Geophysical Research-Atmospheres, 119, 4808-4835, 10.1002/2014jd021595, 2014.

Amann M, Bertok I, Borken-Kleefeld J, Cofala J, Heyes C, Hoeglund-Isaksson L, Klimont Z, Nguyen B, Posch M, Rafaj P, Sandler R, Schoepp W, Wagner F, Winiwarter W: Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications. Environmental Modelling & Software, 26(12):1489-1501, 2011.

AMAP: AMAP Assessment 2006: Acidifying Pollutants, Arctic Haze, and Acidification in the Arctic. Arctic Monitoring and Assessment Programme (AMAP), Oslo, Norway. xii+112 pp, 2006.

AMAP: AMAP Assessment 2015: Black Carbon and Ozone as Arctic Climate Forcers. Arctic Monitoring and Assessment Programme (AMAP), Oslo, Norway, in press, 2015.

Andersson, C., Langner, J., and Bergstrom, R.: Interannual variation and trends in air pollution over Europe due to climate variability during 1958-2001 simulated with a regional CTM coupled to the ERA40 reanalysis, Tellus Series B-Chemical and Physical Meteorology, 59, 77-98, 10.1111/j.1600-0889.2006.00196.x, 2007.

Bahreini, R., Dunlea, E. J., Matthew, B. M., Simons, C., Docherty, K. S., DeCarlo, P. F., Jimenez, J. L., Brock, C. A., and Middlebrook, A. M.: Design and operation of a pressure-controlled inlet for airborne sampling with an aerodynamic aerosol lens, Aerosol Science and Technology, 42, 465-471, 10.1080/02786820802178514, 2008.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevag, A., Seland, O., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjansson, J. E.: The Norwegian Earth System Model, NorESM1-M - Part 1: Description and basic evaluation of the physical climate, Geoscientific Model Development, 6, 687-720, 10.5194/gmd-6-687-2013, 2013.

Bond, T. C., Streets, D. G., Yarber, K. F., Nelson, S. M., Woo, J. H., and Klimont, Z.: A technology-based global inventory of black and organic carbon emissions from combustion, Journal of Geophysical Research-Atmospheres, 109, 10.1029/2003jd003697, 2004.

Bond, T. C., and Bergstrom, R. W.: Light absorption by carbonaceous particles: An investigative review, Aerosol Science and Technology, 40, 27-67, 10.1080/02786820500421521, 2006.

Brandt, J., Silver, J. D., Frohn, L. M., Geels, C., Gross, A., Hansen, A. B., Hansen, K. M., Hedegaard, G. B., Skjoth, C. A., Villadsen, H., Zare, A., and Christensen, J. H.: An integrated model study for Europe and North America using the Danish Eulerian Hemispheric Model with focus on intercontinental transport of air pollution, Atmospheric Environment, 53, 156-176, 10.1016/j.atmosenv.2012.01.011, 2012.

Breider, T. J., Mickley, L. J., Jacob, D. J., Wang, Q., Fisher, J. A., Chang, R. Y. W., and Alexander, B.: Annual distributions and sources of Arctic aerosol components, aerosol optical depth, and aerosol absorption, Journal of Geophysical Research-Atmospheres, 119, 4107-4124, 2014.

Brock, C. A., Cozic, J., Bahreini, R., Froyd, K. D., Middlebrook, A. M., McComiskey, A., Brioude, J., Cooper, O. R., Stohl, A., Aikin, K. C., de Gouw, J. A., Fahey, D. W., Ferrare, R. A., Gao, R. S., Gore, W., Holloway, J. S., Hubler, G., Jefferson, A., Lack, D. A., Lance, S., Moore, R. H., Murphy, D. M., Nenes, A., Novelli, P. C., Nowak, J. B., Ogren, J. A., Peischl, J., Pierce, R. B., Pilewskie, P., Quinn, P. K., Ryerson, T. B., Schmidt, K. S., Schwarz, J. P., Sodemann, H., Spackman, J. R., Stark, H., Thomson, D. S., Thornberry, T., Veres, P., Watts, L. A., Warneke, C., and Wollny, A. G.: Characteristics, sources, and transport of aerosols measured in spring 2008 during the aerosol, radiation, and cloud processes affecting Arctic Climate (ARCPAC) Project, Atmospheric Chemistry and Physics, 11, 2423-2453, 10.5194/acp-11-2423-2011, 2011.

Browse, J., K. S. Carslaw, S. Arnold, K. J. Pringle and O. Boucher, The scavenging processes controlling the seasonal cycle in Arctic sulfate and black carbon aerosol, Atmospheric Chemistry and Physics 12, 6775 – 6798, 2012.

Cavalli, F., Viana, M., Yttri, K. E., Genberg, J.,Putaud, J.P., 2010, Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the EUSAAR protocol, Atmos. Meas. Tech. 3, 79-89.

Chan, T W., L. Huang, W. R. Leaitch, S. Sharma, J., R. Brook, J. Slowik, J. Abbatt: Determination of OM/OC ratios and specific attenuation coefficients in ambient fine PM at a rural site in Central Ontario, Canada, Atmospheric Chemistry and Physics, 10, 2393-2411, 2010

Christensen, J. H.: The Danish Eulerian hemispheric model - A three-dimensional air pollution model used for the Arctic, Atmospheric Environment, 31, 4169-4191, 10.1016/s1352-2310(97)00264-1, 1997.

Eleftheriadis, K., Vratolis, S., and Nyeki, S.: Aerosol black carbon in the European Arctic: Measurements at Zeppelin station, Ny-Alesund, Svalbard from 1998-2007, Geophysical Research Letters, 36, 5, 10.1029/2008gl035741, 2009.

Fisher, J. A., Jacob, D. J., Wang, Q., Bahreini, R., Carouge, C. C., Cubison, M. J., Dibb, J. E., Diehl, T., Jimenez, J. L., Leibensperger, E. M., Lu, Z., Meinders, M. B. J., Pye, H. O. T., Quinn, P. K., Sharma, S., Streets, D. G., van Donkelaar, A., and Yantosca, R. M.: Sources, distribution, and acidity of sulfate-ammonium aerosol in the Arctic in winter-spring, Atmospheric Environment, 45, 7301-7318, 10.1016/j.atmosenv.2011.08.030, 2011.

Flanner, M. G., Zender, C. S., Hess, P. G., Mahowald, N. M., Painter, T. H., Ramanathan, V., and Rasch, P. J.: Springtime warming and reduced snow cover from carbonaceous particles, Atmospheric Chemistry and Physics, 9, 2481-2497, 2009.

Gadhavi, H. S., Renuka, K., Kiran, V. R., Jayaraman, A., Stohl, A., Klimont, Z., and Beig, G.: Evaluation of black carbon emission inventories using a Lagrangian dispersion model - a case study over southern India. Atmos. Chem. Phys. 15, 1447-1461, doi:10.5194/acp-15-1447-2015, 2015.

Garrett, T. J., Zhao, C., and Novelli, P. C.: Assessing the relative contributions of transport efficiency and scavenging to seasonal variability in Arctic aerosol, Tellus Series B-Chemical and Physical Meteorology, 62, 190-196, 10.1111/j.1600-0889.2010.00453.x, 2010.

Garrett, T. J., Brattstrom, S., Sharma, S., Worthy, D. E. J., and Novelli, P.: The role of scavenging in the seasonal transport of black carbon and sulfate to the Arctic, Geophysical Research Letters, 38, 10.1029/2011gl048221, 2011.

GEA (2012), Global Energy Assessment: Toward a Sustainable Future, Cambridge University Press, UK.

Gong, S. L., Zhao, T. L., Sharma, S., Toom-Sauntry, D., Lavoue, D., Zhang, X. B., Leaitch, W. R., and Barrie, L. A.: Identification of trends and interannual variability of sulfate and black carbon in the Canadian High Arctic: 1981-2007, Journal of Geophysical Research-Atmospheres, 115, 9, 10.1029/2009jd012943, 2010.

Heidam, N. Z., Christensen, J., Wahlin, P., and Skov, H.: Arctic atmospheric contaminants in NE Greenland: levels, variations, origins, transport, transformations and trends 1990-2001, Science of the Total Environment, 331, 5-28, 10.1016/j.scitotenv.2004.03.033, 2004.

Hewitt, H. T., Copsey, D., Culverwell, I. D., Harris, C. M., Hill, R. S. R., Keen, A. B., McLaren, A. J., and Hunke, E. C.: Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, Geoscientific Model Development, 4, 223-253, 10.5194/gmd-4-223-2011, 2011.

Hirdman, D., Sodemann, H., Eckhardt, S., Burkhart, J. F., Jefferson, A., Mefford, T., Quinn, P. K., Sharma, S., Ström, J., and Stohl, A.: Source identification of short-lived air pollutants in the Arctic using statistical analysis of measurement data and particle dispersion model output, Atmos. Chem. Phys., 10, 669-693, doi:10.5194/acp-10-669-2010, 2010.

Huang, Lin., J.R. Brook, W. Zhang, S-M. Li, L. Graham, D. Ernst, A. Chivulescu and G. Lu. Stable isotope measurements of carbon fractions (OC/EC) in airborne particulate: A new dimension for source characterization and apportionment. Atmospheric Environment 40: 2690–2705, 2006

Huang, L., Gong, S. L., Jia, C. Q., and Lavoue, D.: Relative contributions of anthropogenic emissions to black carbon aerosol in the Arctic, Journal of Geophysical Research-Atmospheres, 115, 11, 10.1029/2009jd013592, 2010.

Jacob, D. J., Crawford, J. H., Maring, H., Clarke, A. D., Dibb, J. E., Emmons, L. K., Ferrare, R. A., Hostetler, C. A., Russell, P. B., Singh, H. B., Thompson, A. M., Shaw, G. E., McCauley, E., Pederson, J. R., and Fisher, J. A.: The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution, and first results, Atmospheric Chemistry and Physics, 10, 5191-5212, 10.5194/acp-10-5191-2010, 2010.

Jacobson, M. Z.: Climate response of fossil fuel and biofuel soot, accounting for soot's feedback to snow and sea ice albedo and emissivity, Journal of Geophysical Research-Atmospheres, 109, 10.1029/2004jd004945, 2004.

Kanakidou, M., Duce, R. A., Prospero, J. M., Baker, A. R., Benitez-Nelson, C., Dentener, F. J., Hunter, K. A., Liss, P. S., Mahowald, N., Okin, G. S., Sarin, M., Tsigaridis, K., Uematsu, M., Zamora, L. M., and Zhu, T.: Atmospheric fluxes of organic N and P to the global ocean, Global Biogeochemical Cycles, 26, 10.1029/2011gb004277, 2012.

Klimont, Z., Smith, S. J., and Cofala, J.: The last decade of global anthropogenic sulfur dioxide: 2000-2011 emissions, Environmental Research Letters, 8, 10.1088/1748-9326/8/1/014003, 2013.

Klimont, Z., Kupiainen, K., Heyes, Ch., Purohit, P., Cofala, J., Rafaj, P., Borken-Kleefeld, J., Schoepp, W. Global anthropogenic emissions of particulate matter. In preparation, 2015a.

Klimont, Z., Hoglund, L., Heyes, Ch., Rafaj, P., Schoepp, W., Cofala, J., Borken-Kleefeld, J., Purohit, P., Kupiainen, K., Winiwarter, W., Amann, M., Zhao, B., Wang, S.X., Bertok, I., and Sander, R. Global scenarios of air pollutants and methane: 1990-2050. In preparation, 2015b.

Koch, D., Schulz, M., Kinne, S., McNaughton, C., Spackman, J. R., Balkanski, Y., Bauer, S., Berntsen, T., Bond, T. C., Boucher, O., Chin, M., Clarke, A., De Luca, N., Dentener, F., Diehl, T., Dubovik, O., Easter, R., Fahey, D. W., Feichter, J., Fillmore, D., Freitag, S., Ghan, S., Ginoux, P., Gong, S., Horowitz, L., Iversen, T., Kirkevag, A., Klimont, Z., Kondo, Y., Krol, M., Liu, X., Miller, R., Montanaro, V., Moteki, N., Myhre, G., Penner, J. E., Perlwitz, J., Pitari, G., Reddy, S., Sahu, L., Sakamoto, H., Schuster, G., Schwarz, J. P., Seland, O., Stier, P., Takegawa, N., Takemura, T., Textor, C., van Aardenne, J. A., and Zhao, Y.: Evaluation of black carbon estimations in global aerosol models, Atmospheric Chemistry and Physics, 9, 9001-9026, 2009.

Kristiansen, N. I., Stohl, A., and Wotawa, G.: Atmospheric removal times of the aerosol-bound radionuclides ^{137}Cs and ^{131}I measured after the Fukushima Dai-ichi nuclear accident – a constraint for air quality and climate models, Atmos. Chem. Phys., 12, 10759-10769, doi:10.5194/acp-12-10759-2012, 2012.

Lamarque, J. F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850-2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmospheric Chemistry and Physics, 10, 7017-7039, 10.5194/acp-10-7017-2010, 2010.

Law, K. S., and Stohl, A.: Arctic air pollution: Origins and impacts, Science, 315, 1537-1540, 10.1126/science.1137695, 2007.

Law, K. S., Stohl, A., Quinn, P. K., Brock, C. A., Burkhart, J. F., Paris, J.-D., Ancellet, G., Singh, H. B., Roiger, A., Schlager, H., Dibb, J., Jacob, D. J., Arnold, S. R., Pelon, J., and Thomas, J. L.: Arctic air

pollution: New insights from POLARCAT-IPY. *Bull. Am. Met. Soc.* **95**, 1873-1895, doi: http://dx.doi.org/10.1175/BAMS-D-13-00017.1, 2014.

Liu, J. F., Fan, S. M., Horowitz, L. W., and Levy, H.: Evaluation of factors controlling long-range transport of black carbon to the Arctic, Journal of Geophysical Research-Atmospheres, 116, 15, 10.1029/2010jd015145, 2011.

Lund, M. T., and Berntsen, T.: Parameterization of black carbon aging in the OsloCTM2 and implications for regional transport to the Arctic, Atmospheric Chemistry and Physics, 12, 6999-7014, 10.5194/acp-12-6999-2012, 2012.

Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, Geosci. Model Dev., 3, 519-551, doi:10.5194/gmd-3-519-2010, 2010.

Massling, A., Nielsen, I. E., Kristensen, D., Christensen, J. H., Sørensen, L. L., Jensen, B., Nguyen, Q. T., Nøjgaard, J. K., Glasius, M., and Skov, H.: Atmospheric black carbon and sulfate concentrations in Northeast Greenland, Atmos. Chem. Phys. Discuss., 15, 11465-11493, doi:10.5194/acpd-15-11465-2015, 2015.Myhre, G., Berglen, T. F., Johnsrud, M., Hoyle, C. R., Berntsen, T. K., Christopher, S. A., Fahey, D. W., Isaksen, I. S. A., Jones, T. A., Kahn, R. A., Loeb, N., Quinn, P., Remer, L., Schwarz, J. P., and Yttri, K. E.: Modelled radiative forcing of the direct aerosol effect with multi-observation evaluation, Atmospheric Chemistry and Physics, 9, 1365-1392, 10.5194/acp-9-1365-2009, 2009.

Myhre, G., D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura and H. Zhang: Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Myriokefalitakis, S., Tsigaridis, K., Mihalopoulos, N., Sciare, J., Nenes, A., Kawamura, K., Segers, A., and Kanakidou, M.: In-cloud oxalate formation in the global troposphere: a 3-D modeling study, Atmospheric Chemistry and Physics, 11, 5761-5782, 10.5194/acp-11-5761-2011, 2011.

Petzold, A., Ogren, J. A., Fiebig, M., Laj, P., Li, S.-M., Baltensperger, U., Holzer-Popp, T., Kinne, S., Pappalardo, G., Sugimoto, N., Wehrli, C., Wiedensohler, A., and Zhang, X.-Y.: Recommendations for reporting "black carbon" measurements, Atmos. Chem. Phys., 13, 8365-8379, doi:10.5194/acp-13-8365-2013, 2013.

Prank, M., Sofiev, M., Denier van der Gon, H. A. C., Kaasik, M., Ruuskanen, T. M., and Kukkonen, J.: A refinement of the emission data for Kola Peninsula based on inverse dispersion modelling, Atmos. Chem. Phys., 10, 10849-10865, doi:10.5194/acp-10-10849-2010, 2010.

Quinn, P. K., Miller, T. L., Bates, T. S., Ogren, J. A., Andrews, E., and Shaw, G. E.: A 3-year record of simultaneously measured aerosol chemical and optical properties at Barrow, Alaska, Journal of Geophysical Research-Atmospheres, 107, 10.1029/2001jd001248, 2002.

Quinn, P. K., Shaw, G., Andrews, E., Dutton, E. G., Ruoho-Airola, T., and Gong, S. L.: Arctic haze: current trends and knowledge gaps, Tellus Series B-Chemical and Physical Meteorology, 59, 99-114, 10.1111/j.1600-0889.2006.00238.x, 2007.

Quinn, P. K., Bates, T. S., Baum, E., Doubleday, N., Fiore, A. M., Flanner, M., Fridlind, A., Garrett, T. J., Koch, D., Menon, S., Shindell, D., Stohl, A., and Warren, S. G.: Short-lived pollutants in the Arctic: their climate impact and possible mitigation strategies, Atmospheric Chemistry and Physics, 8, 1723-1735, 10.5194/acp-8-1723-2008, 2008.

Ramanathan, V., and Carmichael, G.: Global and regional climate changes due to black carbon, Nature Geoscience, 1, 221-227, 10.1038/ngeo156, 2008.

Riahi, K. et al. (2012), Chapter 17 - Energy Pathways for Sustainable Development, in Global Energy Assessment - Toward a Sustainable Future, pp. 1203–1306, Cambridge University Press, Cambridge, UK and New York, NY, USA and the International Institute for Applied Systems Analysis, Laxenburg, Austria. [online] Available from: www.globalenergyassessment.org

Robertson, L., Langner, J., and Engardt, M.: An Eulerian limited-area atmospheric transport model, Journal of Applied Meteorology, 38, 190-210, 10.1175/1520-0450(1999)038<0190:aelaat>2.0.co;2, 1999.

Samset, B. H., Myhre, G., Herber, A., Kondo, Y., Li, S.-M., Moteki, N., Koike, M., Oshima, N., Schwarz, J. P., Balkanski, Y., Bauer, S. E., Bellouin, N., Berntsen, T. K., Bian, H., Chin, M., Diehl, T., Easter, R. C., Ghan, S. J., Iversen, T., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X., Penner, J. E., Schulz, M., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., and

Zhang, K.: Modelled black carbon radiative forcing and atmospheric lifetime in AeroCom Phase II constrained by aircraft observations, Atmos. Chem. Phys., 14, 12465-12477, doi:10.5194/acp-14-12465-2014, 2014.

Schwarz, J. P., Spackman, J. R., Gao, R. S., Watts, L. A., Stier, P., Schulz, M., Davis, S. M., Wofsy, S. C., and Fahey, D. W.: Global-scale black carbon profiles observed in the remote atmosphere and compared to models, Geophys. Res. Lett., 37, L18812, doi:10.1029/2010GL044372, 2010.

Schwarz, J. P., Samset, B. H., Perring, A. E., Spackman, J. R., Gao, R. S., Stier, P., Schulz, M., Moore, F. L., Ray, E. A., and Fahey, D. W.: Global-scale seasonally resolved black carbon vertical profiles over the Pacific, Geophys. Res. Lett., 40, 5542–5547, doi:10.1002/2013GL057775, 2013.

Sharma, S., Andrews, E., Barrie, L. A., Ogren, J. A., and Lavoue, D.: Variations and sources of the equivalent black carbon in the high Arctic revealed by long-term observations at Alert and Barrow: 1989-2003, Journal of Geophysical Research-Atmospheres, 111, 10.1029/2005jd006581, 2006.

Sharma, S., Ishizawa, M., Chan, D., Lavoue, D., Andrews, E., Eleftheriadis, K., and Maksyutov, S.: 16-year simulation of Arctic black carbon: Transport, source contribution, and sensitivity analysis on deposition, Journal of Geophysical Research-Atmospheres, 118, 943-964, 10.1029/2012jd017774, 2013.

Shaw, G. E.: The Arctic Haze Phenomenon, Bulletin of the American Meteorological Society, 76, 2403 – 2412, 1995.

Shindell, D., and Faluvegi, G.: Climate response to regional radiative forcing during the twentieth century, Nature Geoscience, 2, 294-300, 10.1038/ngeo473, 2009.

Shindell, D. T., Chin, M., Dentener, F., Doherty, R. M., Faluvegi, G., Fiore, A. M., Hess, P., Koch, D. M., MacKenzie, I. A., Sanderson, M. G., Schultz, M. G., Schulz, M., Stevenson, D. S., Teich, H., Textor, C., Wild, O., Bergmann, D. J., Bey, I., Bian, H., Cuvelier, C., Duncan, B. N., Folberth, G., Horowitz, L. W., Jonson, J., Kaminski, J. W., Marmer, E., Park, R., Pringle, K. J., Schroeder, S., Szopa, S., Takemura, T., Zeng, G., Keating, T. J., and Zuber, A.: A multi-model assessment of pollution transport to the Arctic, Atmospheric Chemistry and Physics, 8, 5353-5372, 2008.

Skeie, R. B., Berntsen, T. K., Myhre, G., Tanaka, K., Kvalevag, M. M., and Hoyle, C. R.: Anthropogenic radiative forcing time series from pre-industrial times until 2010, Atmospheric Chemistry and Physics, 11, 11827-11857, 10.5194/acp-11-11827-2011, 2011a.

Skeie, R. B., Berntsen, T., Myhre, G., Pedersen, C. A., Ström, J., Gerland, S., and Ogren, J. A.: Black carbon in the atmosphere and snow, from pre-industrial times until present, Atmos. Chem. Phys., 11, 6809-6836, doi:10.5194/acp-11-6809-2011, 2011b.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth System Model: ECHAM6-HAM2, Journal of Advances in Modeling Earth Systems, 5, 146-172, 10.1002/jame.20015, 2013.

Stohl, A.: Characteristics of atmospheric transport into the Arctic troposphere. *J. Geophys. Res.* **111**, D11306, doi:10.1029/2005JD006888, 2006.

Stohl, A., Hittenberger, M., and Wotawa, G.: Validation of the Lagrangian particle dispersion model FLEXPART against large scale tracer experiments. Atmospheric Environment **32**, 4245-4264, 1998.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmospheric Chemistry and Physics, 5, 2461-2474, 2005.

Stohl, A., Klimont, Z., Eckhardt, S., Kupiainen, K., Shevchenko, V. P., Kopeikin, V. M., and Novigatsky, A. N.: Black carbon in the Arctic: the underestimated role of gas flaring and residential combustion emissions, Atmospheric Chemistry and Physics, 13, 8833-8855, 10.5194/acp-13-8833-2013, 2013.

Stone, R. S., Herber, A., Vitale, V., Mazzola, M., Lupi, A., Schnell, R. C., Dutton, E. G., Liu, P. S. K., Li, S. M., Dethloff, K., Lampert, A., Ritter, C., Stock, M., Neuber, R., and Maturilli, M.: A three-dimensional characterization of Arctic aerosols from airborne Sun photometer observations: PAM-ARCMIP, April 2009, Journal of Geophysical Research-Atmospheres, 115, 10.1029/2009jd013605, 2010.

Sullivan, A. P., Peltier, R. E., Brock, C. A., de Gouw, J. A., Holloway, J. S., Warneke, C., Wollny, A. G., and Weber, R. J.: Airborne measurements of carbonaceous aerosol soluble in water over northeastern United States: Method development and an investigation into water-

soluble organic carbon sources, Journal of Geophysical Research-Atmospheres, 111, 14, 10.1029/2006jd007072, 2006.

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

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmos. Chem. Phys., 10, 11707-11735, doi:10.5194/acp-10-11707-2010, 2010.

Vignati, E., M. Karl, M. Krol, J. C. Wilson, P. Stier and F. Cavalli, Sources of uncertainties in modelling black carbon at the global scale, Atmospheric Chemistry and Physics 10, 2595 – 2611, 2010.

von Salzen, K.: Piecewise log-normal approximation of size distributions for aerosol modelling, Atmospheric Chemistry and Physics, 6, 1351-1372, 2006.

von Salzen, K., Scinocca, J. F., McFarlane, N. A., Li, J. N., Cole, J. N. S., Plummer, D., Verseghy, D., Reader, M. C., Ma, X. Y., Lazare, M., and Solheim, L.: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of Physical Processes, Atmosphere-Ocean, 51, 104-125, 10.1080/07055900.2012.755610, 2013.

Wang, H., Easter, R. C., Rasch, P. J., Wang, M., Liu, X., Ghan, S. J., Qian, Y., Yoon, J. H., Ma, P. L., and Vinoj, V.: Sensitivity of remote aerosol distributions to representation of cloud-aerosol interactions in a global climate model, Geoscientific Model Development, 6, 765-782, 10.5194/gmd-6-765-2013, 2013.

Warneke, C., Froyd, K. D., Brioude, J., Bahreini, R., Brock, C. A., Cozic, J., de Gouw, J. A., Fahey, D. W., Ferrare, R., Holloway, J. S., Middlebrook, A. M., Miller, L., Montzka, S., Schwarz, J. P., Sodemann, H., Spackman, J. R., and Stohl, A.: An important contribution to springtime Arctic aerosol from biomass burning in Russia, Geophysical Research Letters, 37, 10.1029/2009gl041816, 2010.

Wang, Q., D. J. Jacob, J. R. Spackman, A. E. Perring, J. P. Schwarz, N. Moteki, E. A. Marais, C. Ge, J. Wang, and S. R. H. Barrett, Global budget and radiative forcing of black carbon

aerosol: Constraints from pole-to-pole (HIPPO) observations across the Pacific, J. Geophys. Res. Atmos., 119, 195–206, doi:10.1002/2013JD020824, 2014.

Wofsy, S. C., Team, H. S., Cooperating Modellers, T., and Satellite, T.: HIAPER Pole-to-Pole Observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences, 369, 2073-2086, 10.1098/rsta.2010.0313, 2011.

Zaveri, R. A., and Peters, L. K.: A new lumped structure photochemical mechanism for large-scale applications, Journal of Geophysical Research-Atmospheres, 104, 30387-30415, 10.1029/1999jd900876, 1999.

Zaveri, R. A., Easter, R. C., Fast, J. D., and Peters, L. K.: Model for Simulating Aerosol Interactions and Chemistry (MOSAIC), Journal of Geophysical Research-Atmospheres, 113, 29, 10.1029/2007jd008782, 2008.

Zhang, K., O'Donnell, D., Kazil, J., Stier, P., Kinne, S., Lohmann, U., Ferrachat, S., Croft, B., Quaas, J., Wan, H., Rast, S., and Feichter, J.: The global aerosol-climate model ECHAM-HAM, version 2: sensitivity to improvements in process representations, Atmospheric Chemistry and Physics, 12, 8911-8949, 10.5194/acp-12-8911-2012, 2012.

1 Table 1. Model overview

2

Model Name	Model Type[1]	Horizontal/vertical resolution Model domain	Meteorological fields; treatment of aerosol mixtures	Periods simulated / output temporal resolution	References
FLEXPART	LPDM	Met. Input data: 1° x1°92L global	ECMWF Operational Analyses; none	2008-2009 3h	Stohl et al. (1998, 2005)
OsloCTM2	CTM	2.8°x2.8°, 60L global	ECMWF IFS Forecasts ; aerosol externally mixed	2008-2009 3h	Myhre et al. (2009), Skeie et al. (2011a, 2011b)
NorESM	CCM	1.9°x2.5°, 26L global	Internal, observed SST prescribed; BC internally mixeed	2008-2009 3h	Kirkevåg et al. (2013), Bentsen et al. (2013)
TM4-ECPL	CTM	2°x3°, 34L global	ECMWF ERA-interim; aerosols externally mixed	2008-2009 24h	Myriokefalitakis et al. (2011); Kanakidou et al. (2012); Daskalakis et al. (2014)
ECHAM6-HAM2	ACM	1.8°x1.8°, 31L global	ECMWFReanalysis; aerosols internally mixed	March-August, 2008, 1h	Stevens et al. (2013), Zhang et al. (2012)
SMHI-MATCH	CTM	0.57°x0.75°, 38L 20-90°N	ECMW – ERA-Interim; BC internally mixed	2008, 2009 1h	Andersson et al. (2007), Robertson et al. (1999)
CanAM4.2	ACM	2.8°x2.8°, 49L, global	Nudged to ECMWF temp.and winds; aged BC internally, near emission externally	2008-2009 3h	Von Salzen et al. (2013), von Salzen (2006)
DEHM	CTM	150km <60° 50km >60°N, 29L 0-90°N	NCEP; internally mixed aerosols	2008-2009 3h	Christensen (1997), Brandt et al. (2012)
CESM1/CAM5.2	CCM	1.9°x2.5°, 30L global	Internal, observed SST prescribed; internally mixed aerosols	2008-2009 1h	Liu et al. (2012), Wang et al. (2013)
WRF-Chem	RCCM	100kmx100km 38L 27-90° N	Nudged every 6h to FNL to all levels above the PBL; internally mixed aerosols	March-July 2008 3h	Grell et al. (2005), Zaveri et al. (1999), Zaveri et al. (2008)
HadGEM3	CCM	1.9°x1.3°, 63L global	ECMWF ERA-interim; internally mixed aerosol	March-June, November 2008, January, May and November 2009 2h	Hewitt et al. (2011), Mann et al. (2010)

3 [1]Chemistry transport model (CTM), Lagrangian particle dispersion model (LPDM), chemistry climate model (CCM), aerosol climate model
4 (ACM), regional climate model coupled with a chemistry module (RCM)
5

Table 2. Median observed eBC and modeled BC mass surface concentrations in ng/m^3 as well as measured and modeled sulfate (SO_4) concentrations in the Arctic during winter/spring (January to March) and summer (July to September). The data used are from the years 2008 and 2009 and were averaged for the three stations Alert, Barrow and Zeppelin. Notice that some models do not cover the whole periods completely (see Table 1).

Model/obs	Winter/Spring BC [ng/m^3]	Summer BC [ng/m^3]	Winter/Spring SO_4 [ng/m^3]	Summer SO_4 [ng/m^3]
Measured	49.4	3.3	561.0	103.2
Model mean:	20.1	6.2	353.6	148.6
FLEXPART	40.2	7.7		
OsloCTM2	8.4	1.3	90.2	109.7
NorESM	13.0	4.4	394.2	70.8
TM4-ECPL	5.4	1.3	71.3	149.7
ECHAM6-HAM2	1.9	2.1	488.7	388.9
SMHI-MATCH	38.6	1.1	603.3	151.1
CanAM4.2	38.8	1.6	791.3	270.9
DEHM	57.1	11.6	434.6	61.1
CESM1-CAM5	21.3	5.1	210.5	21.9
WRF-Chem	14.9	32.3	408.8	246.6
HadGEM3	1.8	0.7	43.2	15.9

Table 3: Slopes of regression lines between monthly mean concentrations of sulfate and (e)BC for the different stations. Slopes are calculated both for the observations and the model values. Values that are statistically significant at the 99.9% level are marked with an asterisk. For the mean over all sites/models, only the statistically significant values were averaged.

	Alert	Barrow	Pallas	Zeppelin	Mean
Observations	**10.1**	6.4	**8.4**	**8.9**	9.1
Model mean	17.3	16.6	6.7	9.7	12.6
OsloCTM2	-8.6	2.4	-2.0	-5.5	-
NorESM	**35.3**	27.8	0.4	12.1	35.3
TM4-ECPL	9.5	**33.2**	**5.8**	8.1	19.5
ECHAM6-HAM2	30.0	90.4	1.0	-746.4	-
SMHI-MATCH	**25.6**	**25.9**	0.4	10.9	25.7
CanAM4.2	**18.2**	2.5	7.1	**12.4**	15.3
DEHM	**7.5**	**5.7**	**1.6**	**6.7**	5.4
CESM1-CAM5.2	**11.1**	**8.9**	**9.6**	**9.9**	9.9
WRF-Chem	**6.4**	**9.3**	**9.8**	2.4	8.5
HadGEM3	10.7	-8.7	-0.81	3.2	-

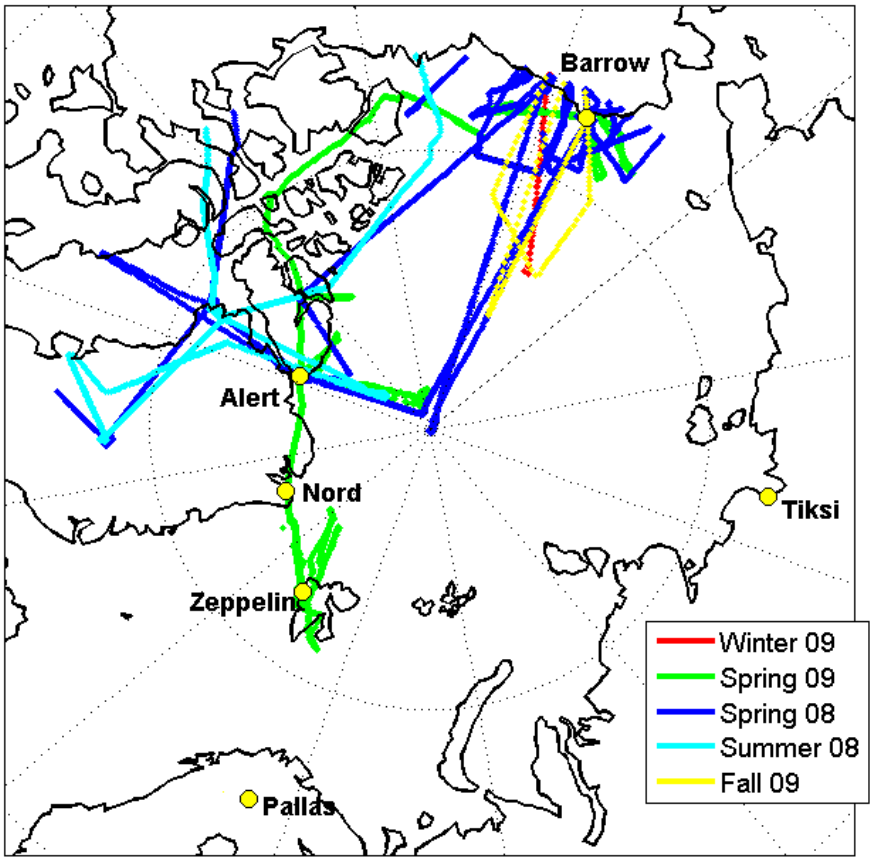


Figure 1. Map showing the locations of the measurement stations (yellow circles) and the flight tracks north of 70°N of all aircraft campaigns used in this study. Aircraft data were from the HIPPO (winter 2009 and fall 2009), ARCTAS (spring and summer 2008), ARCPAC (spring 2008) and PAMARCMiP (spring 2009) campaigns.

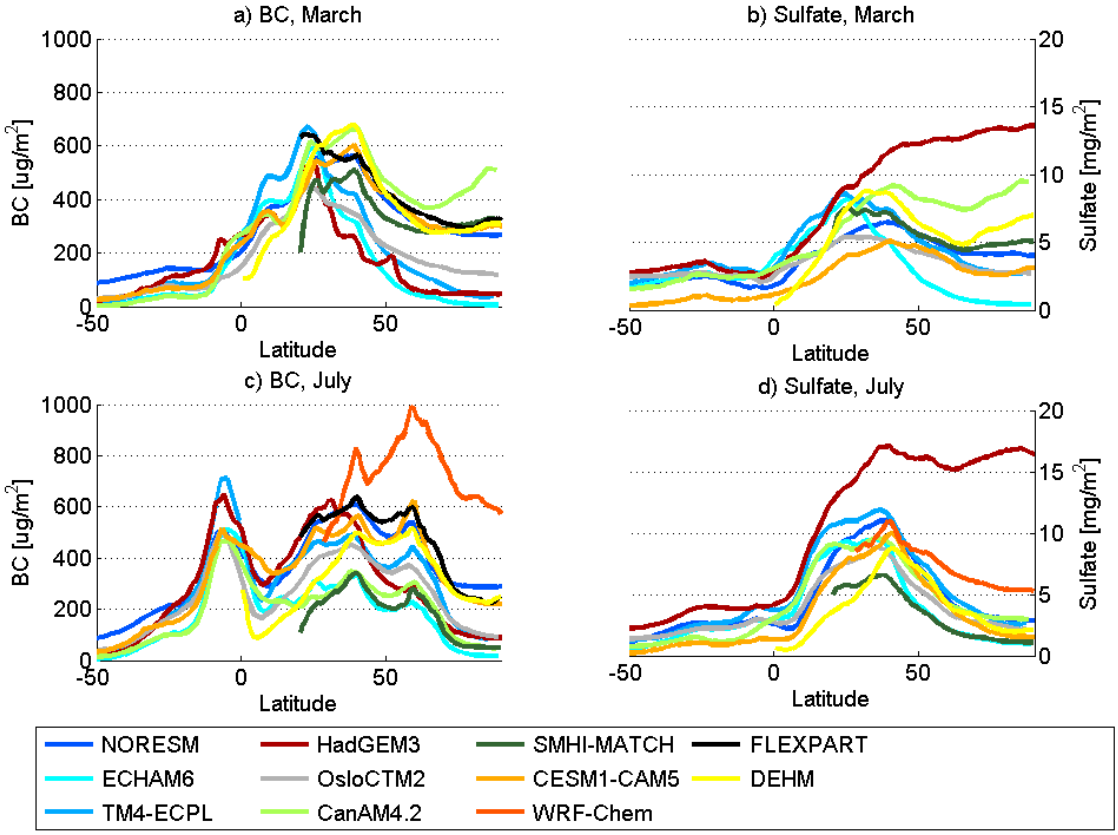


Figure 2: BC (a, c) and sulfate (b, d) column mass loadings for the year 2008 averaged over all longitudes as a function of latitude (for the range 50°S to 90°N) for March (a-b) and July (c-d).

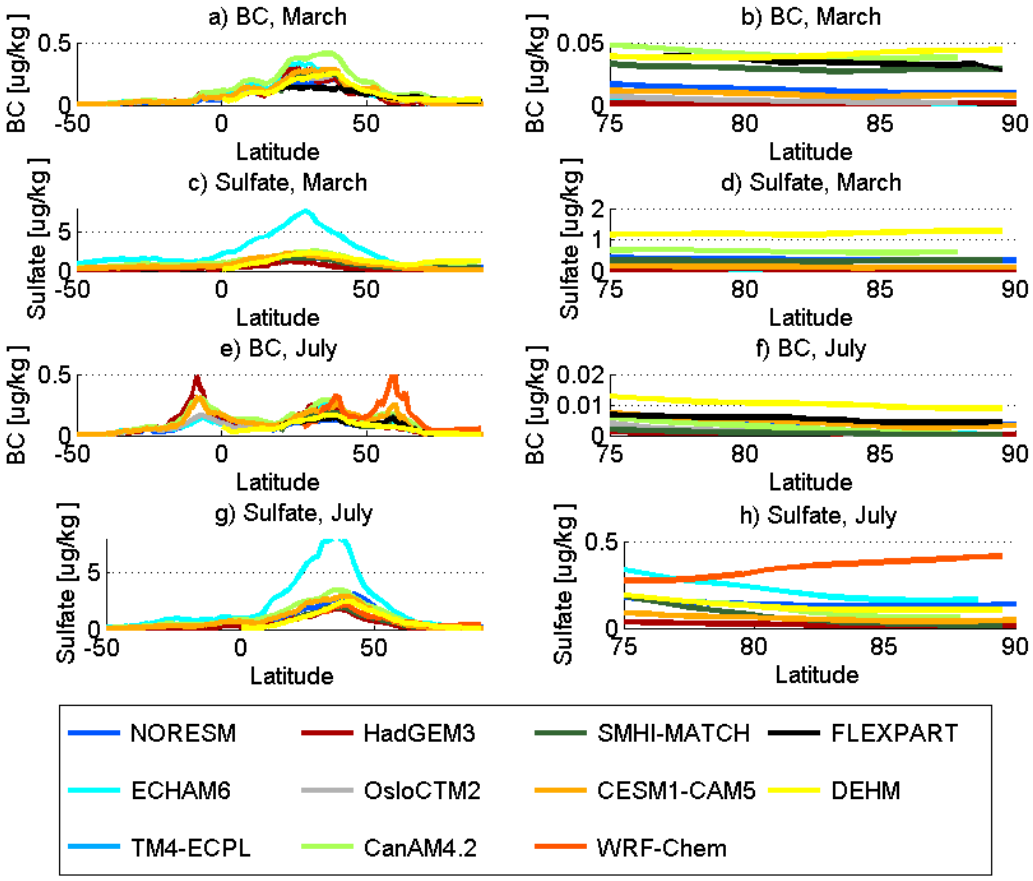


Figure 3: BC (a-b, e-f) and sulfate (c-d, g-h) mass mixing ratios for the year 2008 at the surface averaged over all longitudes as a function of latitude (for the range 50°S to 90°N) for March (a-d) and July (e-h). The right panels show the same data as the left panels, but only for 70-90°N and with an adjusted ordinate scale.

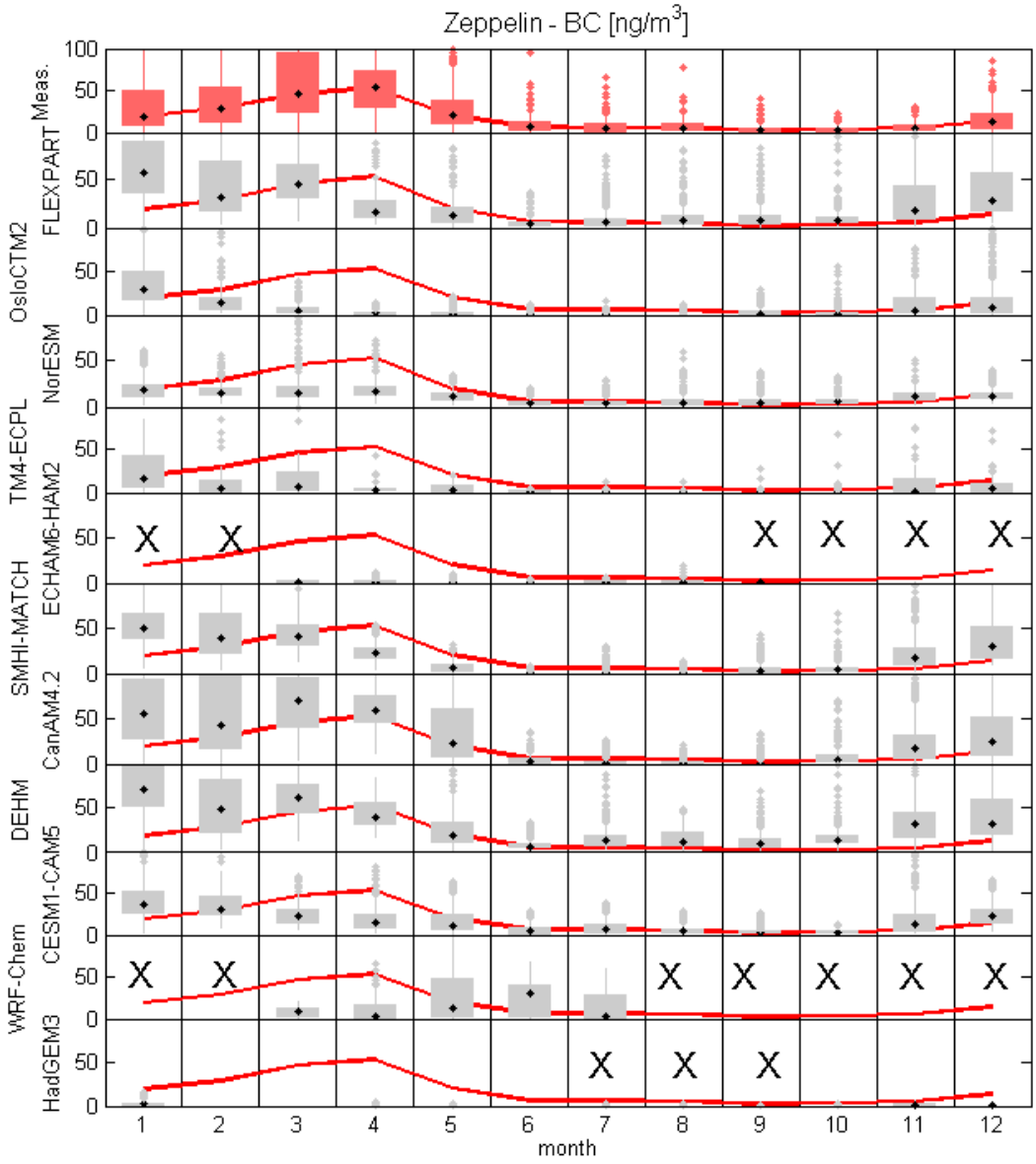


Figure 4. Observed and simulated mean annual cycle of (equivalent) BC mass concentrations [ng/m^3] at the Zeppelin station. Shown is the monthly frequency distributions using data from the years 2008 and 2009. The uppermost panel (red boxes) shows monthly frequency distributions of the observed eBC concentrations. The other panels below (grey boxes) show monthly frequency distributions of the modeled BC concentrations. Black dots depict the monthly median value, the grey boxes span the range between the 25th and 75th percentiles, red and grey dots represent values which are outside the 1.5 fold of this interquartile range (grey lines). The red line connects the monthly medians of the observed eBC concentrations in the uppermost panel and is repeated in all other panels for the convenience of comparing modeled

and measured values. Missing model data are denoted with “X”. Notice that some models have very low BC mass concentrations, which are difficult to see on the scale used.

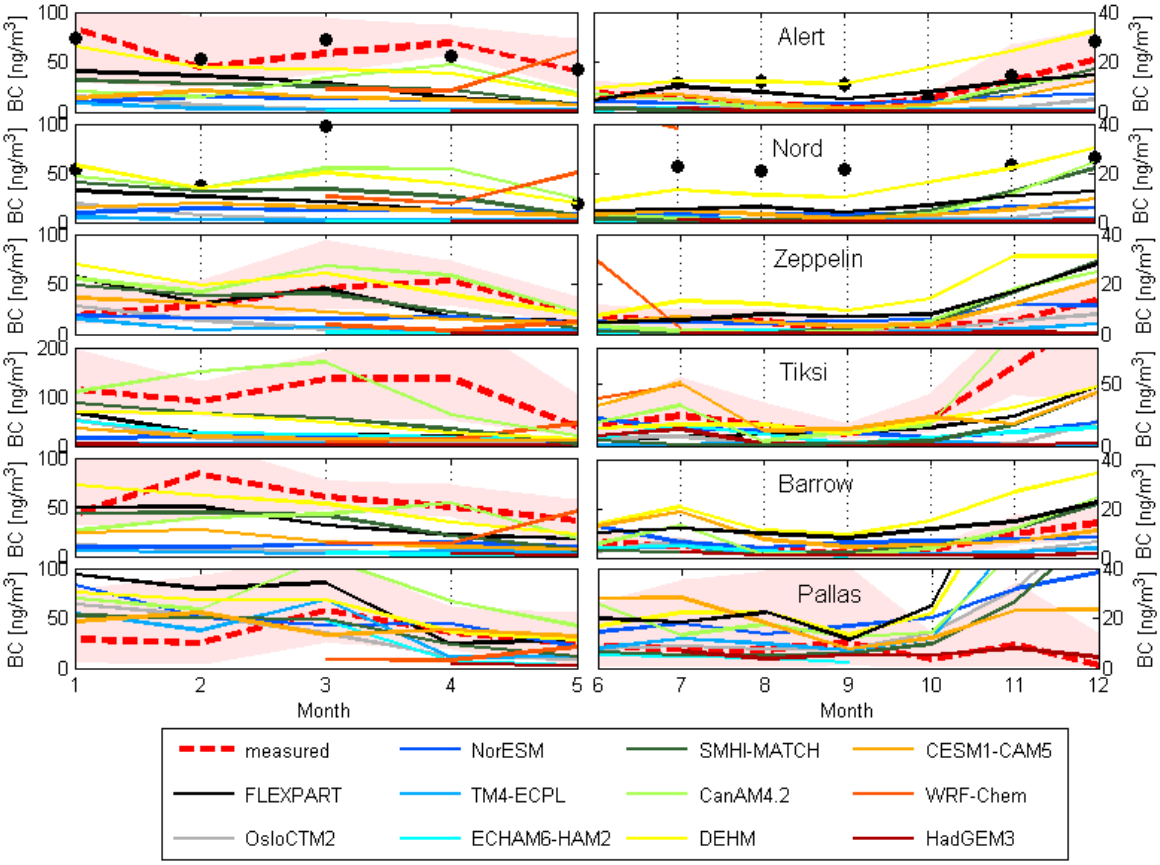


Figure 5. Surface concentrations of monthly (month is displayed on the abscissa) median observed eBC or EC and modeled BC. Each row represents one station: (from top) Alert, Nord, Zeppelin, Tiksi, Barrow and Pallas, for late winter/spring (left column) and summer/fall (right column). The red dashed lines connect the observed median eBC values, the light red shaded areas span from the 25th to the 75th percentile of the observations. The black dots are the EC concentrations, which are available for Alert and Station Nord. Modeled median values are shown with different lines according to the legend. Notice the difference in concentration scales used for the left and right panels and also for the Tiksi station.

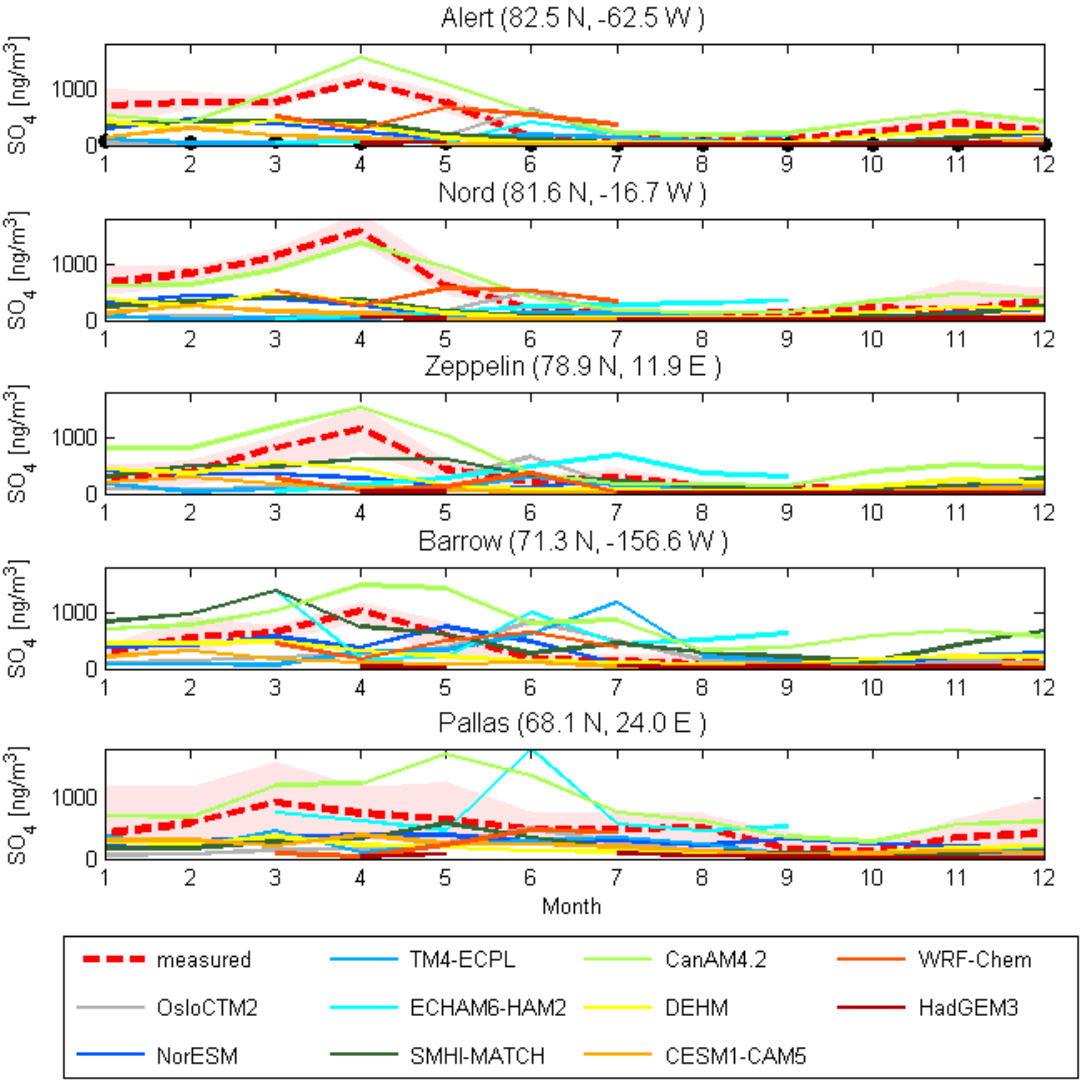


Figure 6. Monthly (month is displayed on the abscissa) median observed and modeled sulfate surface concentrations for the stations (from top) Alert, Nord, Zeppelin, Barrow and Pallas. The red dashed lines connect the observed median values. The light red shaded areas span from the 25th to the 75th percentile of the observations. Modeled median values are shown with different lines according to the legend.

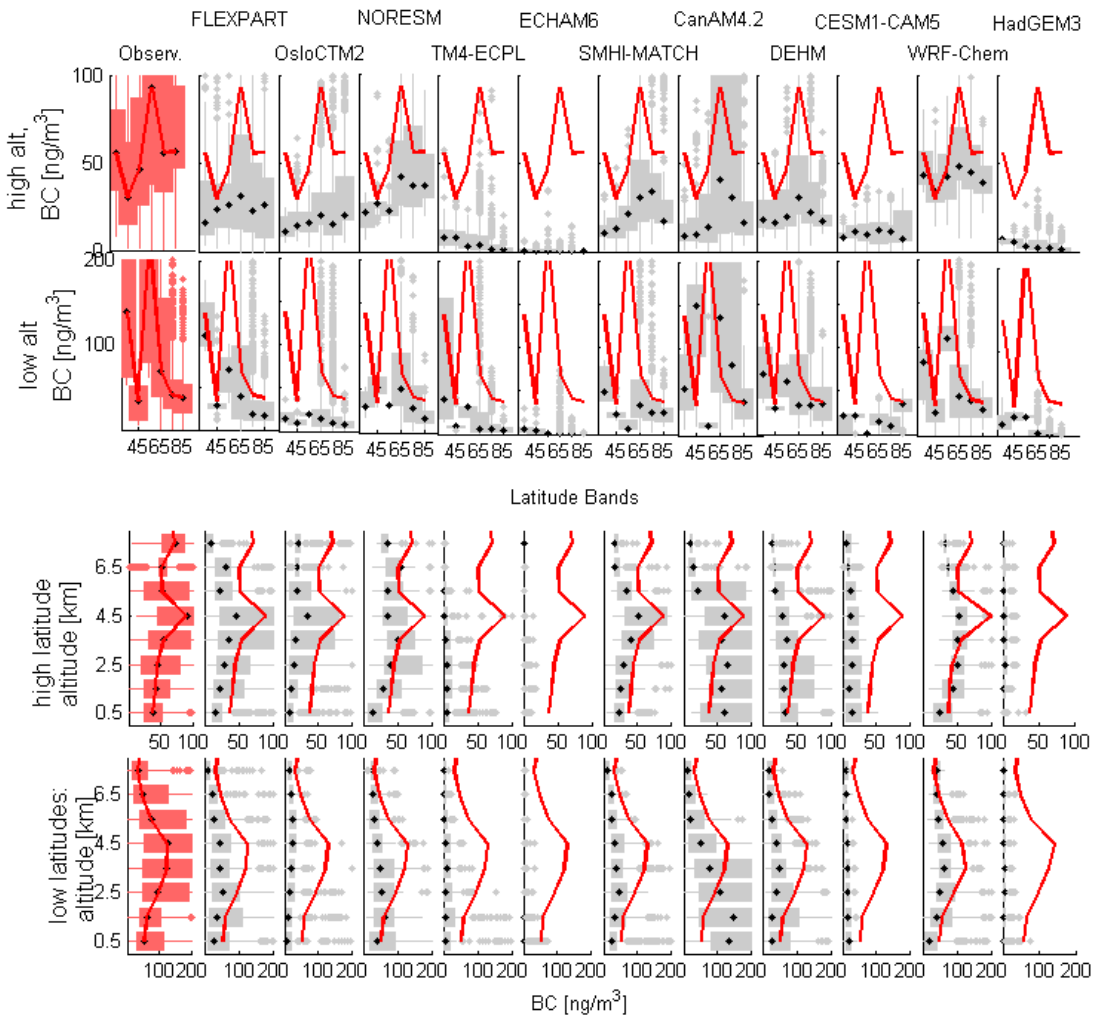


Figure 7: Comparison of modeled BC with observed rBC (red boxes and red lines) mass concentrations from the ARCTAS-spring and ARCPAC campaigns in spring 2008. The leftmost column shows box and whisker plots (like in Fig 4: boxes go from 25th to 75th percentile, whiskers span the 1.5fold interquartile range) of observed rBC concentrations in ng/m^3. The black dots as well as the red lines represent the median values. The other columns show the modeled BC concentrations for FLEXPART, OsloCTM2, NorESM, TM4-ECPL, ECHAM6-HAM2, SMHI-MATCH, CanAM4.2, DEHM, CESM1-CAM5.2, WRF-Chem and HadGEM3. The top row represents median (r)BC concentrations for altitudes below 3 km asl as a function of latitude by binning the data into 10° latitude bands. The second row represents median (r)BC concentrations for altitudes above 3 km asl. The third (bottom) row shows median (r)BC concentrations for latitudes north of (south of) 70^{0}N as a function of altitude by binning the data into 1-km height intervals.

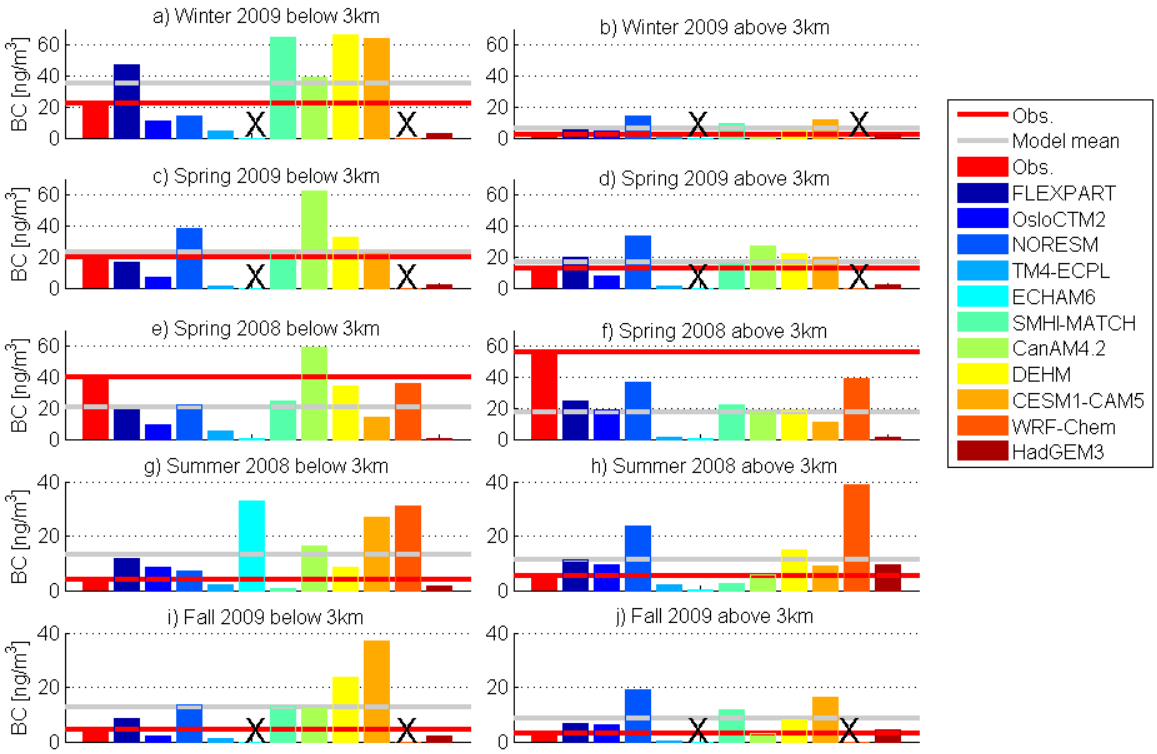


Figure 8: Median observed rBC and modeled BC mass concentrations for the winter 2009 HIPPO (a, b) spring 2009 PAMARCMiP (c-d) spring 2008 ARCTAS/ARCPAC (e-f) summer 2008 ARCTAS (g-h) and the fall 2009 HIPPO (i-j) aircraft campaigns. The red bar and the red horizontal line show the observations, the other colored bars the various models, the grey line shows the mean value of all model medians. Results are shown separately for measurements below 3 km (left panels) and above 3 km (right panels). Notice that the concentration scales on the ordinates are different for the individual panels.

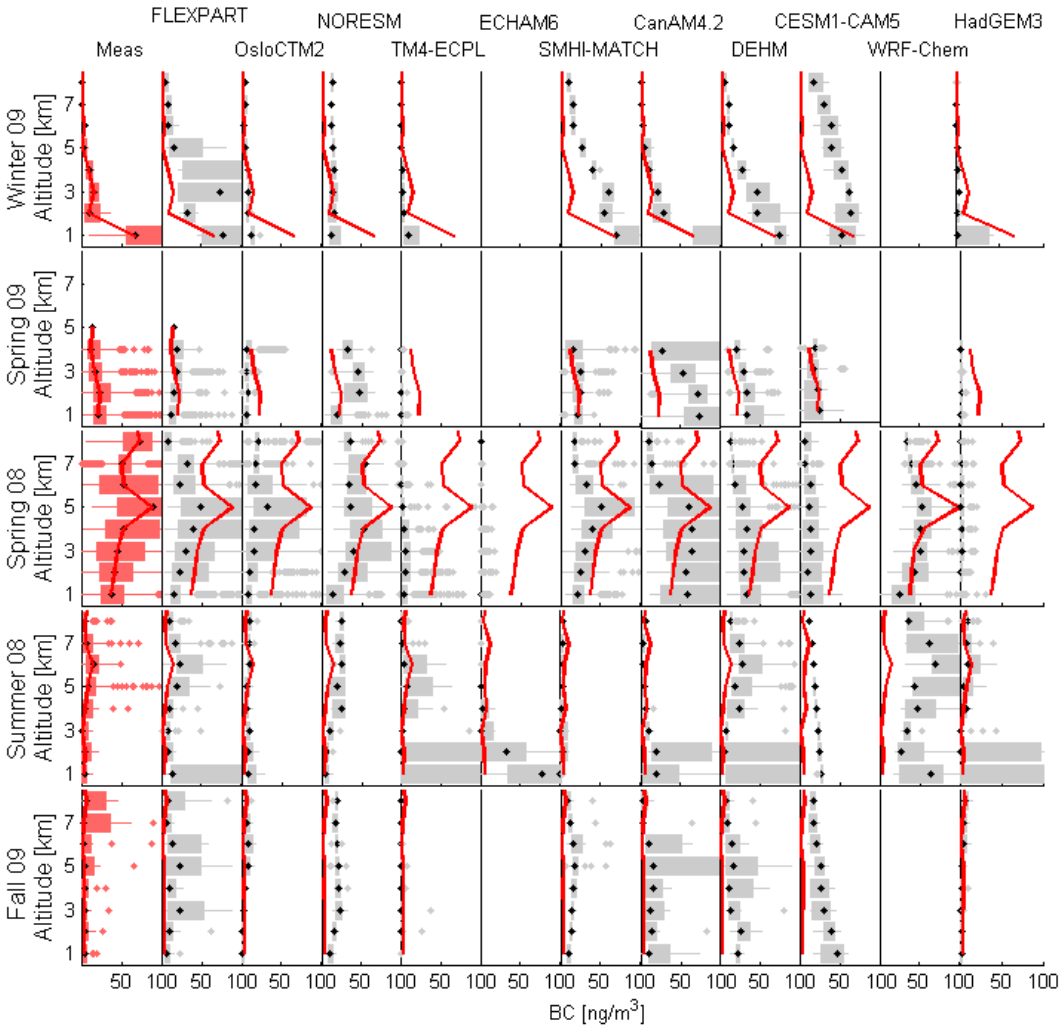


Figure 9: Comparison of modeled BC with observed rBC mass concentrations as a function of altitude for all data taken north of 70°N, for the different campaigns (same as in Fig. 8). The leftmost column shows box and whisker plots of observed rBC concentrations in ng/m^3. The black dots as well as the red lines represent the median values. The other columns show the modeled BC concentrations for FLEXPART, OsloCTM2, NorESM, TM4-ECPL, ECHAM6-HAM2, SMHI-MATCH, CanAM4.2, DEHM, CESM1-CAM5.2, WRF-Chem and HadGEM3.

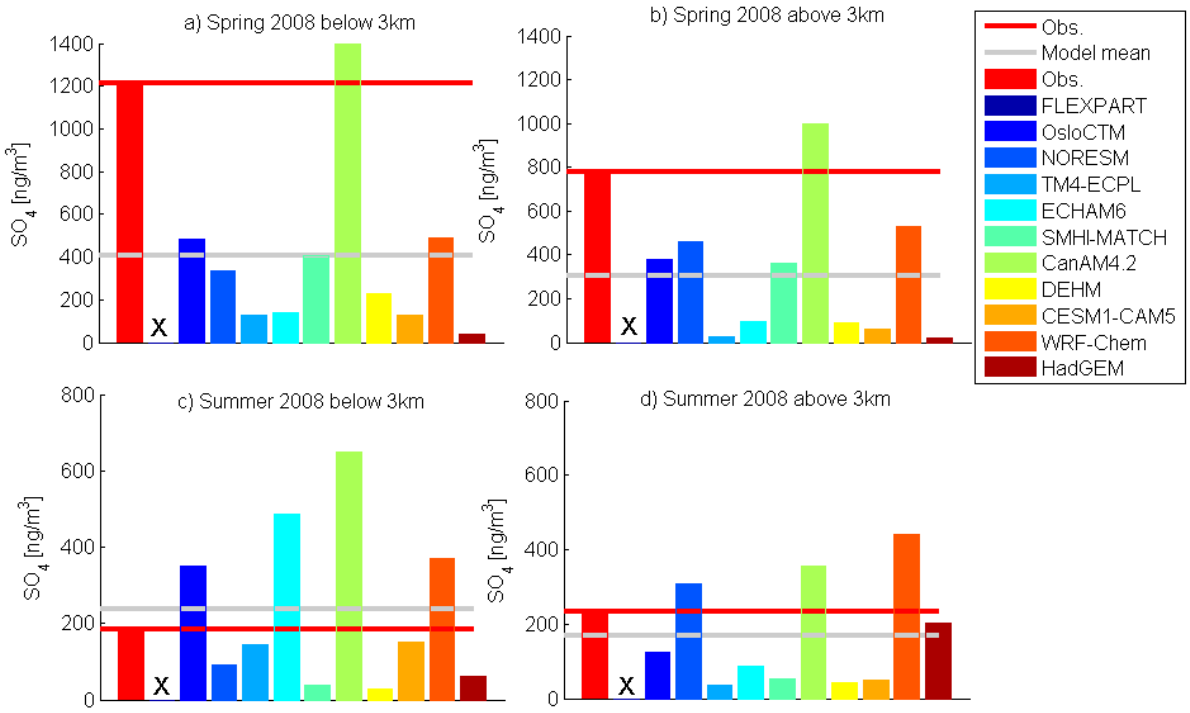


Figure 10: Median SO_4 concentrations for the ARCTAS/ARCPAC spring 2008 (a-b) and ARCTAS summer 2008 (c-d) campaigns. The red bar and the red horizontal line show the observations, the other colored bars the various models. The analysis is performed for measurements below 3 km (left panels) and above 3 km (right panels). Note: each row has a different y-axis.

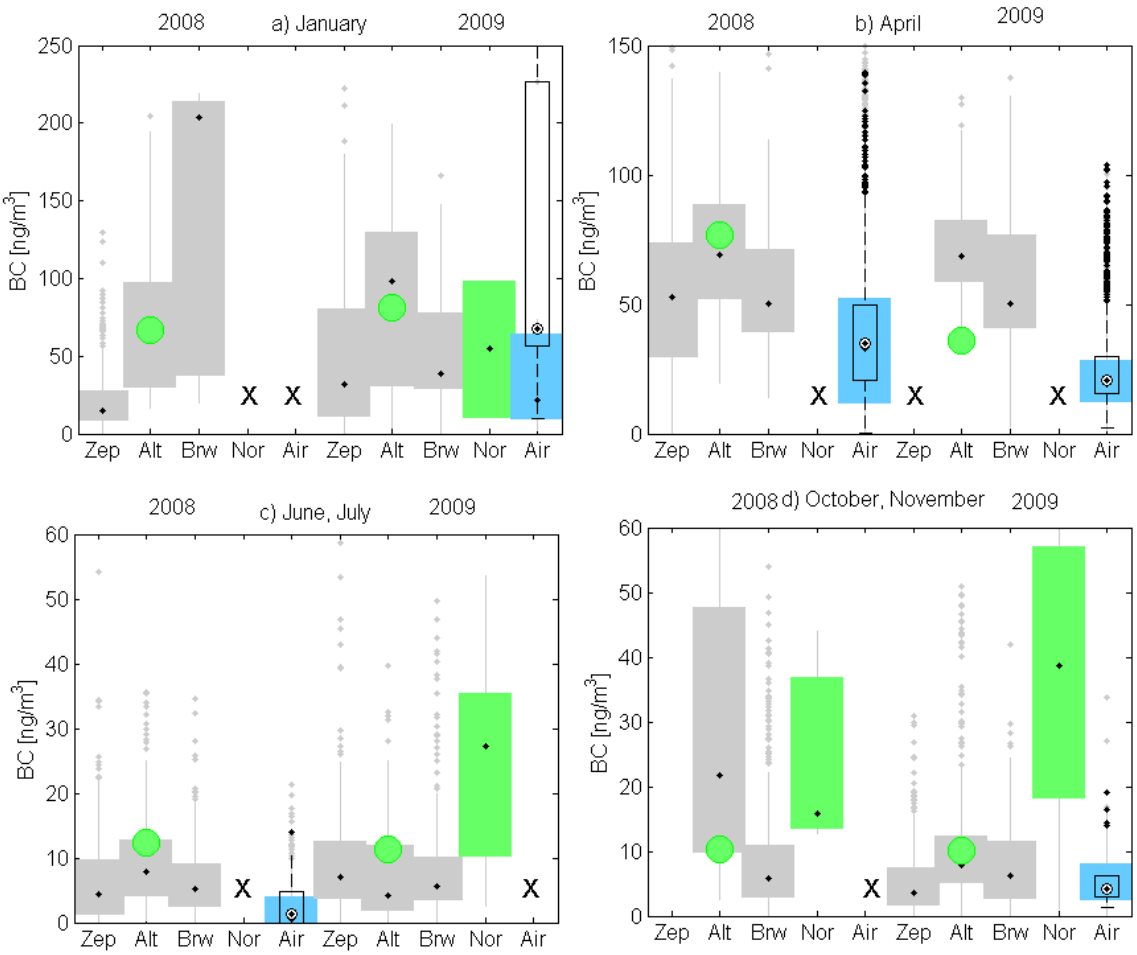
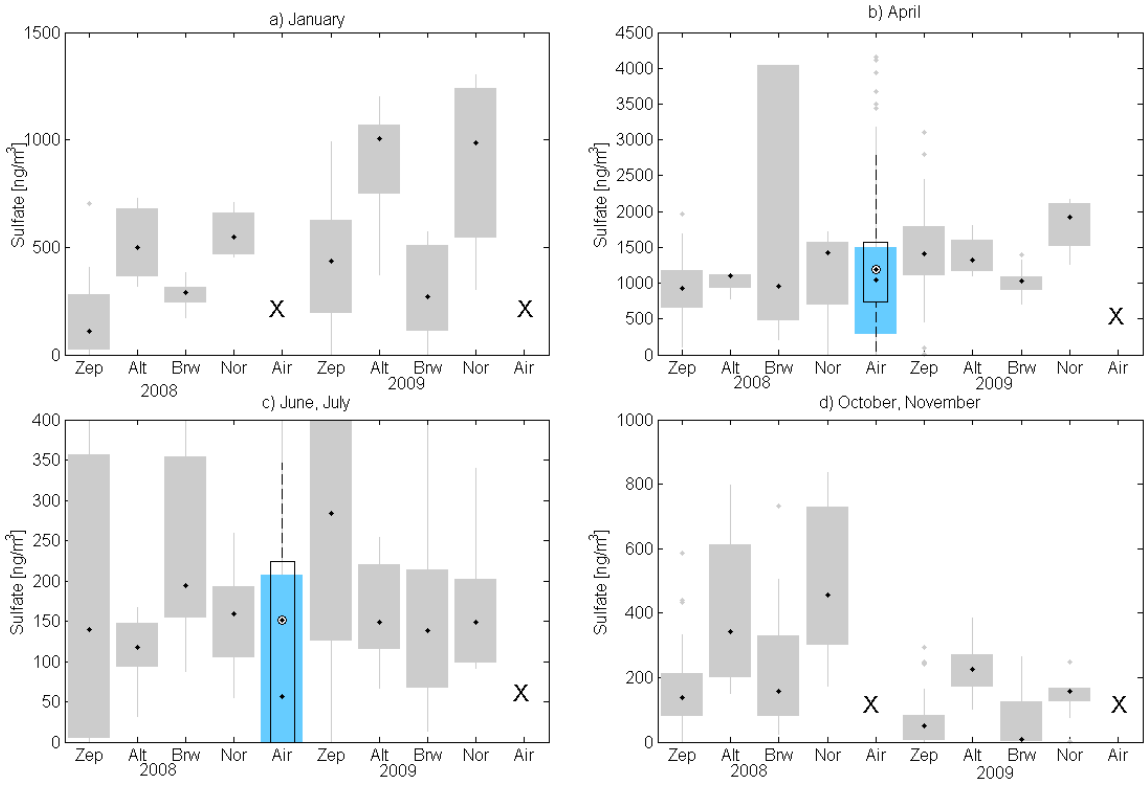


Figure 11: Comparison of eBC [ng/m^3] measured at the stations Zeppelin (Zep), Alert (Alt), and Barrow (Brw) (grey bars), EC measured at Alert and Station Nord (Nord) (green dots and bars) and rBC [ng/m^3] measured by aircraft (Air) in the lowest 3 km and 1 km, north of 70°N (blue bars) for the years 2008 and 2009 for a) January, b) April, c) June and July and d) October and November. The black dots represent the median, and the boxes the interquartile range. For the aircraft measurements, the blue boxes show the results for the lowest 3 km, the black box outlines show the results for the lowest 1 km.

1



2

3 Figure 12: Same as figure 9, but for sulfate.

4

5

6

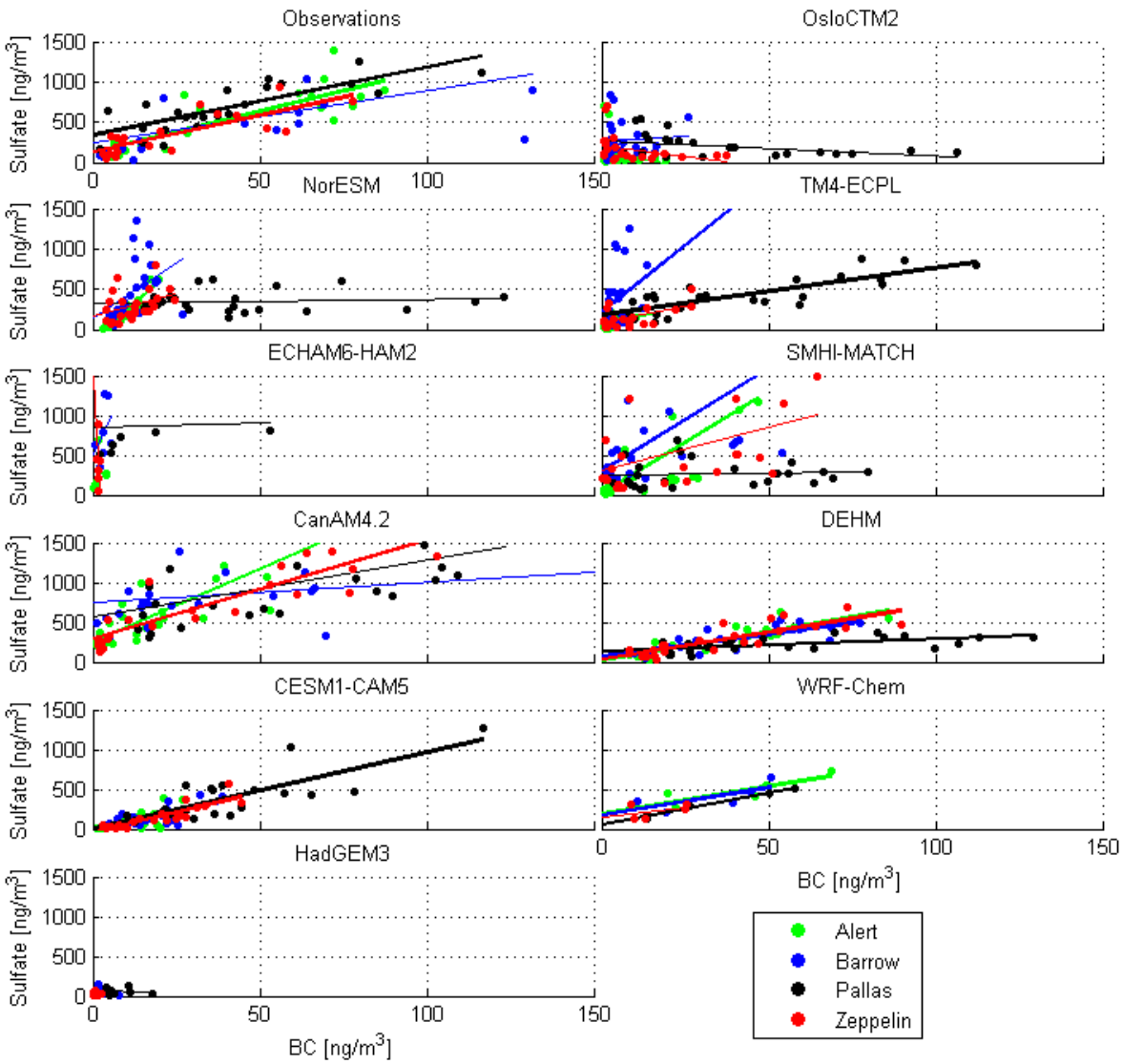


Figure 13: Correlation plots of monthly mean sulfate and (e)BC concentrations for the observations (top left) and the different models sampled at the observation sites. Thick lines denote significant correlations.