

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

5
6 **S. Eckhardt¹, B. Quennehen^{2*}, D.J.L. Olivié³, T.K. Berntsen⁴, R. Cherian⁵, J. H.**
7 **Christensen⁶, W. Collins^{7,8}, S. Crepinsek^{9,10}, N. Daskalakis¹¹, M. Flanner¹², A.**
8 **Herber¹³, C. Heyes¹⁴, Ø. Hodnebrog⁴, L. Huang¹⁵, M. Kanakidou¹¹, Z. Klimont¹⁴, J.**
9 **Langner¹⁶, K.S. Law², M. T. Lund⁴, R. Mahmood^{19,20}, A. Massling⁶, S.**
10 **Myriokefalitakis¹¹, I. E. Nielsen⁶, J. K. Nøjgaard⁶, J. Quaas⁵, P. K. Quinn¹⁷, J.-C.**
11 **Raut², S. T. Rumbold^{7,21}, M. Schulz³, S. Sharma¹⁵, R.B. Skeie⁴, H. Skov⁶, T. Uttal¹⁰,**
12 **K. von Salzen¹⁸ and A. Stohl¹**

13 [1]{NILU - Norwegian Institute for Air Research, Kjeller, Norway}

14 [2]{Sorbonne Universités, UPMC Univ. Paris 06 ; Université Versailles St-Quentin ; CNRS/INSU ; LATMOS-
15 IPSL, UMR 8190, Paris, France}

16 [3]{Norwegian Meteorological Institute, Oslo, Norway}

17 [4]{Center for International Climate and Environmental Research – Oslo (CICERO), Oslo, Norway}

18 [5]{Institute for Meteorology, Universität Leipzig, Germany}

19 [6]{ENVS Department of Environmental Science, Aarhus University, Roskilde, Denmark }

20 [7]{Met Office Hadley Centre, Exeter, UK}

21 [8]{ Department of Meteorology, University of Reading, Reading, UK}

22 [9] {Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado,
23 USA}

24 [10]. NOAA Earth System Research Laboratory Physical Sciences Division / Polar Observations & Processes.
25 Boulder, Colorado, USA

26 [11]{Environmental Chemical Processes Laboratory, Department of Chemistry, University of
27 Crete, Heraklion, Crete, and ICE-HT/FORTH, Patras, Greece }

28 [12]{Department of Atmospheric, Oceanic, and Space Sciences, University of Michigan, Ann Arbor, MI, USA}

29 [13]{Alfred Wegener Institut, Helmholtz Centre for Polar- and Marine Research, Bremerhaven, Germany}

1 [14]{International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria}
2 [15] {Climate Research Division, Atmospheric Sci. & Tech. Directorate,S & T, Environment Canada
3 Toronto, Ontario, Canada}
4 [16]{Swedish Meteorological and Hydrological Institute (SMHI), SE-60176 Norrköping, Sweden}
5 [17] {National Oceanic and Atmospheric Administration Pacific Marine Environmental Laboratory, Seattle, WA,
6 USA}
7 [18]{Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia,
8 Canada}
9 [19]{School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada}
10 [20]{Department of Meteorology, COMSATS Institute of Information Technology,
11 Islamabad, Pakistan}
12 [21]{ National Centre for Atmospheric Science, University of Reading, Reading, UK}
13
14 * - now at Univ. Grenoble Alpes/CNRS, Laboratoire de Glaciologie et Géophysique de l'Environnement (LGGE),
15 38041 Grenoble, France
16
17 Correspondence to: S. Eckhardt (sabine.eckhardt@nilu.no)
18

1 **Abstract**

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

1

2 1 Introduction

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

13 The Arctic aerosol consists of a varying mixture of sulfate and organic carbon (OC), as well as
14 ammonium, nitrate, BC and mineral dust (Quinn et al., 2007; Brock et al., 2011). Aerosols in
15 the Arctic feature a strong annual cycle with a late winter/spring peak (the so-called Arctic
16 Haze) and a summer minimum. Increased transport during the cold season (Stohl, 2006) and
17 increased removal by wet deposition during the warm season can explain this annual variation
18 (Shaw, 1995; Law and Stohl, 2007) and also shape the aerosol size distribution (Tunved et al.,
19 2013).

20 Models have for a long time struggled to capture the distribution of aerosols in the Arctic
21 (Shindell et al., 2008; Koch et al., 2009). The concentrations of BC during the Arctic Haze
22 season in particular were underestimated, in some cases by more than an order of magnitude
23 (Shindell et al., 2008), whereas summer concentrations were sometimes overestimated. The
24 simulated aerosol seasonality is strongly dependent on the model treatment of aerosol removal
25 processes. For instance, changes in the calculation of aerosol microphysical properties, size
26 distribution and removal can change simulated concentrations by more than an order of
27 magnitude in remote regions such as the Arctic (Vignati et al., 2010) and the calculated Arctic
28 BC mass concentrations are very sensitive to parameterizations of BC aging (conversion from
29 hydrophobic to hydrophilic properties) and wet scavenging (Liu et al., 2011; Huang et al.,
30 2010).

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

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

16 Remaining problems may also be due to missing emission sources or incorrect spatial or
17 temporal distribution of emissions in the inventories used for the modeling. The main sources
18 of BC are biomass burning and incomplete combustion of fossil fuels and biofuels (Bond et al.,
19 2004). Sulfate aerosols are formed by sea spray or originate from natural sources such as
20 oxidation of dimethyl sulfide (DMS) or volcanoes. It is also produced from oxidation of SO_2
21 emitted when sulfur-containing fossil fuels are burned or by metal smelting. Studies based on
22 observed surface concentrations repeatedly suggest that the main source regions for Arctic BC
23 and sulfate are located in high-latitude Eurasia (e.g., Sharma et al., 2006, Eleftheriadis 2009,
24 Hirdman et al., 2010). Stohl et al. (2013) suggested that gas flaring in high-latitude Russia is an
25 important source of BC which is missing from most inventories. In their simulations, BC
26 emissions from gas flaring accounted for 42% of the annual mean BC surface concentrations
27 in the Arctic. However, they also noted the large uncertainty of the gas flaring emissions.

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

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

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

12 **2 Methods**

13 **2.1 Measurement data**

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

20 Aerosol light absorption data were obtained from five sites in different parts of the Arctic: Alert,
21 Canada (62.3°W, 82.5°N; 210 m above sea level (asl)), Zeppelin/Ny Ålesund, Spitsbergen,
22 Norway (11.9°E, 78.9°N; 478 m asl), Tiksi, Russia (128.9° E, 71.6°N; 1 m asl), Barrow, Alaska
23 (156.6°W, 71.3°N; 11 m asl) and Pallas, Finland (24.12°E, 67.97°N; 565 m asl). The locations
24 of these measurement stations are shown in Fig. 1. Different types of particle soot absorption
25 photometers (PSAPs) were used for the measurements at Barrow, Alert, and Zeppelin, a multi-
26 angle absorption photometer was used at Pallas (Hyvärinen et al., 2011), and an aethalometer
27 was used at Tiksi. All these instruments measure the particle light absorption coefficient σ_{ap} ,
28 each at its own specific wavelength (typically at around 530–550 nm), and for different size
29 fractions of the aerosol (typically particles smaller than 1, 2.5 or 10 μm are sampled at different
30 humidities). Conversion of σ_{ap} to eBC mass concentrations is not straightforward and requires
31 certain assumptions (Petzold et al., 2013). The mass absorption efficiency used for conversion

1 can be specific to a site, the instrument and the wavelength used and is uncertain by at least a
2 factor of two. For Tiksi, the conversion is done internally by the aethalometer. For the other
3 sites, a mass absorption efficiency of $10 \text{ m}^2 \text{ g}^{-1}$, typical of aged BC aerosol (Bond and
4 Bergstrom, 2006), was used. Concentrations of eBC can be particularly uncertain and biased
5 high when substantial amounts of organic carbon are present (Cappa et al., 2008; Lack et al.,
6 2008).

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

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

25

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

1 Aircraft data were obtained from several campaigns. In the framework of POLARCAT (Polar
2 Study using Aircraft, Remote Sensing, Surface Measurements, and Models of Climate
3 Chemistry, Aerosols, and Transport; Law et al., 2014), two ARCTAS (Arctic Research of the
4 Composition of the Troposphere from Aircraft and Satellites) campaigns in April and June/July
5 2008 with a DC-8 aircraft covered mainly the North American Arctic (Jacob et al., 2010). The
6 ARCPAC (Aerosol, Radiation, and Cloud Processes affecting Arctic Climate; Brock et al.,
7 2011) campaign was conducted from Alaska together with ARCTAS in April 2008. The
8 PAMARCMiP (Polar Airborne Measurements and Arctic Regional Climate Model Simulation
9 Project) campaign covered the entire western Arctic in April 2009 (Stone et al., 2010). Two
10 HIPPO (High-Performance Instrumented Airborne Platform for Environmental Research Pole-
11 to-Pole Observations; Schwarz et al., 2010; Schwarz et al., 2013; Wofsy et al., 2011) campaigns
12 during January 2009 and October 2009 explored the North American Arctic. Flight legs north
13 of 70°N for all of these campaigns are shown in Fig. 1. Refractory BC (rBC) was measured
14 during these campaigns with single particle soot photometer (SP2) instruments (Kondo et al.,
15 2001; Schwarz et al., 2006). Observations of submicrometer aerosol sulfate mass during
16 ARCTAS were made with a particle-into liquid-Sampler (PILS) (Sullivan et al., 2006) coupled
17 to an ion chromatograph. Sulfate measurements during ARCPAC were made with a compact
18 time-of-flight aerosol mass spectrometer (Bahreini et al., 2008).

19
20 During April 2008 agricultural and boreal biomass burning influence was widespread
21 throughout the Arctic (Warneke et al., 2010; Brock et al., 2011) and ARCTAS and ARCPAC
22 often targeted these fire plumes. Anthropogenic pollution from Asia was also sampled by these
23 campaigns in the western Arctic, particularly in the mid-upper troposphere (see Law et al., 2014
24 and references therein). Pollution from Europe also made a significant contribution in the lower
25 troposphere. In contrast, PAMARCMiP and HIPPO sampled the Arctic atmosphere at times
26 with little influence from biomass burning and also did not target pollution plumes. Thus, the
27 higher mean rBC concentrations found during ARCTAS and ARCPAC than during
28 PAMARCMiP a year later are caused both by the sampling strategy of these campaigns as well
29 as the early start of the biomass burning season in 2008. Even though all available rBC and
30 sulfate data from several campaigns were used for model evaluation, the data coverage and
31 representativity for the Arctic as a whole must still be considered as rather poor. The Eastern
32 Arctic, in particular, was not sampled by any campaign.

1 ARCTAS-B was the only summertime POLARCAT campaign to make detailed measurements
2 of BC and sulfate (Jacob et al., 2010). These flights focused mainly on boreal fires over Canada
3 in July 2008 but several flights into the high Arctic sampled, for example Asian pollution close
4 to the North Pole (Sodemann et al., 2010). Plumes of Asian origin were also sampled in the
5 upper troposphere over Canada (Singh et al., 2010).

6 **2.2 Emissions**

7 All models made use of an identical emission dataset, the ECLIPSE (Evaluating the Climate
8 and Air Quality Impacts of Short-Lived Pollutants) emission inventory version V4a (Klimont
9 et al., 2015a, 2015b). The ECLIPSE inventory was created using the GAINS (Greenhouse gas
10 – Air pollution Interactions and Synergies) model (Amann et al., 2011), which provides
11 emissions of long-lived greenhouse gases and shorter-lived species in a consistent framework.
12 The proxies used in GAINS are consistent with those applied within the RCP (Representative
13 Concentration Pathways) projections as described in Lamarque et al. (2010) and as further
14 developed within the Global Energy Assessment project (GEA, 2012). They were, however,
15 modified to accommodate more recent information where available, e.g., on population
16 distribution and open biomass burning, effectively making them year specific (Riahi et al.,
17 2012; Klimont et al., 2013). Emissions for the years 2008 and 2009 were lumped into the
18 following source categories: industrial combustion, residential combustion, energy production,
19 transport, agriculture, waste treatment, shipping, agricultural waste burning and gas flaring. All
20 emission data were gridded consistently to a resolution of $0.5^{\circ}\times0.5^{\circ}$. Monthly disaggregation
21 factors were provided for the domestic heating emissions, based on ambient air temperatures.
22 For a more detailed description of the ECLIPSE emission data set, see Klimont et al. (2015a,
23 2015b). A detailed description of the high-latitude emissions in the ECLIPSE inventory and
24 comparisons with other emission inventories can be found in AMAP (2015).

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

1 **2.3 Models**

2 We show results of 11 different models, whose main characteristics and references are
3 summarized in Table 1. In principle we are using two types of atmospheric models: off-line
4 models and on-line models. Both model types have certain advantages and disadvantages. Off-
5 line models based on meteorological re-analysis data can capture actual meteorological
6 situations, thus facilitating a direct comparison of measured and modeled aerosol quantities.
7 Often, they also have higher resolution than the on-line global models. However, off-line
8 models cannot be used for predictions and the off-line coupling can also cause inaccuracies in
9 the treatment of transport, chemistry and removal processes. The global on-line models in our
10 study are free-running and thus produce their own model climate, which means that they cannot
11 reproduce a given meteorological situation. Nevertheless, their modeled climate for the present
12 time should correspond to the current climatic conditions and, thus, seasonally averaged
13 quantities (i.e., averages over many different meteorological situations) should be comparable
14 to measured quantities. The main advantage of the on-line models is that they can also be used
15 for predictions.

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

27 The horizontal resolution of the individual models ranges from about $0.6^\circ \times 0.8^\circ$ to $2.8^\circ \times 2.8^\circ$.
28 We use one Lagrangian particle transport model, FLEXPART (Flexible Particle Dispersion
29 Model), which is run in backward mode for 30 days (thus, older source contributions are not
30 accounted for). The simulation is driven by $1^\circ \times 1^\circ$ operational analyses from the European
31 Centre for Medium Range Weather Forecast (ECMWF). The OsloCTM2, TM4-ECPL (Tracer
32 Model version 4 - Environmental Chemical Processes Laboratory) and SMHI MATCH

1 (Swedish Meteorological and Hydrological Institute Multi-scale Atmospheric Transport and
2 Chemistry Model) are CTMs and also use meteorological data from ECMWF (for details see
3 table 1). The DEHM (Danish Eulerian Hemispheric Model) CTM is driven by NCEP (National
4 Centers for Environmental Prediction) meteorological data. WRF-Chem (Weather Research
5 and Forecasting Model coupled with Chemistry) is an on-line atmospheric chemistry-weather
6 forecast model which was nudged to NCEP FNL (final analysis) data for this study. The
7 aerosol-climate model (ACM) ECHAM6-HAM2 (for brevity, referred to as ECHAM6 in
8 figures) is the European Centre for Medium-Range Weather Forecasts Hamburg model version
9 6 (Stevens et al., 2013) extended with the Hamburg aerosol module version 2 (HAM2) (Zhang
10 et al., 2012). ECHAM6-HAM2 and the CCMs including HadGEM3 (Met Office Hadley Centre
11 Climate Model, version 3) and CanAM4.2 (Canadian Atmospheric model, version 4.2) were
12 nudged to ECMWF data. CESM1-CAM5.2 (Community Earth System Model version 1 –
13 Community Atmosphere model version 5.2) and NorESM1-M (Norwegian Earth System
14 Model version 1 with intermediate resolution and used here in a version where aerosols are
15 fully coupled with a tropospheric gas-phase chemistry scheme, hereafter referred to as
16 NorESM) are also CCMs but were running freely, thus producing their own meteorological
17 data. These latter models cannot be compared point-to-point with the measurement data because
18 they produced meteorological conditions which were different from the actual ones; however,
19 longer-term (e.g., seasonal) medians should still be comparable with the measurements,
20 especially since sea surface temperatures (SST) and sea-ice extent were prescribed and specific
21 to the years 2008-2009. All models were sampled exactly at the locations of the measurement
22 stations and along the flight tracks at the highest possible (mostly hourly) temporal resolution.
23 Notice that not all models simulated the full 2008-2009 period and FLEXPART only simulated
24 BC.

25

26 **3. Simulated BC and sulfate concentrations**

27 Figure 2 shows the simulated BC and sulfate column mass loadings as a function of latitude for
28 the time periods of the Arctic Haze (March) and the much cleaner summer (July) in the Arctic,
29 for the models for which this information was available. For BC in March, most models show
30 a maximum near 20°N, with some models extending this maximum to 40°N. This
31 approximately covers the latitude range with the highest global emissions where the models
32 agree at least within a factor of two in their simulated column loadings. In contrast, larger

1 differences between the models are found in the Arctic, where column mass loadings vary by
2 more than an order of magnitude. Similar results are also found for sulfate in March, for which
3 most models also show a maximum around 20-40°N; however, compared to BC, the models
4 show a less pronounced decrease towards higher latitudes and two models even simulate
5 increasing sulfate burdens with latitude. The relatively good agreement between the models in
6 the BC and sulfate source region latitudes is not surprising, given that they all use the same
7 emission data set. In contrast, the differences between the atmospheric column loadings in the
8 Arctic must mainly be due to differences in the aerosol processing and removal and hence
9 aerosol lifetimes, and probably differences in atmospheric transport. Most models with
10 relatively low BC column loadings in the Arctic also have low sulfate loadings there, indicating
11 similarities in the simulated removal of these two types of aerosols. A notable exception,
12 however, is HadGEM3, which has moderately low BC but the highest sulfate loadings in the
13 Arctic.

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

26 In the simulated surface BC and sulfate mass mixing ratios the same basic patterns are found
27 as in the column loadings, but with enhanced gradients between source areas and remote regions
28 (Fig. 3). When looking at individual models, there are, however, notable differences for sulfate.
29 ECHAM6-HAM2 has the highest sulfate surface mass mixing ratios of all models, especially
30 in the northern hemisphere subtropics and mid-latitudes. Combined with the rather “normal”
31 column sulfate loadings of this model, this indicates that ECHAM6-HAM2 does not transport
32 sulfate away from the surface as quickly as the other models. On the other hand, HadGEM3,

1 which has by far the largest sulfate column loadings, has the smallest surface concentrations.
2 This deficiency was due to the implementation of the Global Model of Aerosol Processes
3 (GLOMAP; Mann et al., 2010), which in this HadGEM3 version resulted in too little removal
4 of the sulfate precursor SO_2 during the venting from the boundary layer to the free troposphere.
5 The longer sulfate lifetime there explains the high column loadings.

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

13

14 **4 Observed and simulated BC and sulfate seasonality at Arctic surface 15 measurement stations**

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

1 The eBC mass concentrations at the three other sites in the western Arctic (Alert, Barrow,
2 Pallas) are quite comparable to those at Zeppelin station, with monthly median values of about
3 20-80 ng/m³ in late winter/early spring and of less than 10 ng/m³ in summer/early fall (see Fig.
4 5). One exception is EC measured at Station Nord, which in summer is higher than eBC
5 measured at the other sites. At Alert, where both eBC and EC data are available, EC values in
6 summer are also somewhat higher than eBC values (although lower than the Station Nord EC
7 values), probably due to systematic differences in measurement techniques.

8 At the Tiksi station, which is closer to the main source regions of Arctic BC in high-latitude
9 Eurasia (Hirdman et al., 2010), higher monthly median eBC values were measured (more than
10 100 ng /m³ in winter/spring, about 20-40 ng /m³ in summer) and the annual mean (81 ng/m³) is
11 2.5 times higher than the average for the other stations (31 ng/m³). The seasonality of measured
12 eBC is strongest at Alert where the summer concentrations are very low, but the winter/spring
13 concentrations are similar to the other sites in the western Arctic. This result points to a
14 deepening of the seasonal minimum with latitude. While the aerosol concentrations in the Arctic
15 during late winter/early spring are comparable to remote regions further south, the
16 concentrations in summer/early fall are lower because of the effective cleansing of the
17 atmosphere (Garrett et al., 2010, 2011; Browse et al., 2012; Tunved et al., 2013) and less
18 efficient transport from source regions (Stohl, 2006). The highest eBC concentrations were
19 observed in January (Alert), February (Barrow), March (Pallas, Tiksi) or April (Zeppelin), with
20 no clear dependence of the time of the maximum on latitude; however, the maximum occurred
21 earlier at the two North American sites than at the other sites.

22 The models capture the Arctic BC concentrations with variable success (Fig. 5). Most models
23 capture the much higher concentrations in winter/spring than summer/fall, and some models
24 can approximately reproduce the concentrations reached during the Arctic Haze season (see
25 also Breider et al., 2014). However, as already seen for the Zeppelin station (Fig. 4) and the
26 annual mean surface mass mixing ratios (Fig. 3), there is a large variability between individual
27 models, with seasonal median values varying by about an order of magnitude both in spring
28 and summer even when excluding the most extreme models (see also Table 2). Seasonal mean
29 concentrations during January to March are underestimated by up to a factor of 27 for individual
30 models and by more than a factor 2 for the mean over all models, and only one model slightly
31 overestimates the measured concentrations (Table 2). Nevertheless, this indicates clear progress
32 since earlier studies (e.g. Shindell et al., 2008; Koch et al., 2009; AMAP, 2011), where it was

1 reported that most models had a completely wrong seasonality and systematically
2 underpredicted the Arctic Haze concentrations. For instance, in Shindell et al. (2008), none of
3 their models came close to the measured concentrations at Barrow and Alert during winter and
4 spring, with a model-mean underestimate of about one order of magnitude (their Fig. 7). It is
5 also important to keep in mind that the eBC measurements are uncertain and could be biased
6 high. However, EC and eBC values at Alert are very similar and we find a similar model
7 underestimate of measured EC at Station Nord as well.

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

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

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

1 and found that the differences between the different inventories grow with latitude and are
2 largest north of 70°N (i.e., high-latitude Eurasian emissions).

3 The seasonal cycle of sulfate at the monitoring stations is similar to that of eBC, with a clear
4 maximum during the Arctic Haze season and a minimum in summer/early fall (Fig. 6).
5 However, the seasonal cycle at the northernmost stations is less strong than for eBC, with about
6 a factor of 5 difference between spring and summer, compared to a factor of 15 for eBC (Table
7 2). This is probably due to the influence of biogenic sources of sulfate in summer (Quinn et al.,
8 2002) and/or a weaker seasonality in the emissions (e.g., smelter emissions of SO_2 are probably
9 relatively constant throughout the year).

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

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

29 At Pallas, the lowest-latitude station in this comparison, most models severely underestimate
30 sulfate throughout the year (Fig. 6), although they tend to overestimate BC in spring there. One
31 likely reason for the sulfate underestimation is the proximity of the Pallas station to the Kola
32 peninsula, where metal smelters are a strong source of sulfur. According to AMAP (2006), SO_2

1 emissions in Nikel, Zapolyarnyy and Monchegorsk together were about 170 kt/yr in the year
2 2002. In the ECLIPSE version 4a inventory used for this study the SO₂ emissions in these areas
3 are only about 33 kt/yr in total for the year 2005. Similar deficiencies were in fact reported also
4 for other emission inventories for this region (Prank et al., 2010). Strong underestimation of the
5 SO₂ emissions from metal smelting in the Kola peninsula is therefore a likely explanation for
6 why almost all models underestimate sulfate at Pallas so strongly. Similar discrepancies were
7 in fact found for SO₂ emissions in Norilsk, prompting a regridding of the ECLIPSE emissions
8 (now available version 5a) using better location information for the metal smelting industry.

9 **5 Vertical Profiles**

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

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

30 When plotted as a function of altitude (two bottom panel rows in Fig. 7), the measured values
31 peak in the 4-5 km altitude bin, both for sub-Arctic and Arctic latitudes. The models, except for
32 CanAM4.2, underestimate the measured median values throughout the entire depth of the

1 profile. Some of the models, mainly those driven by observed meteorology, capture the rBC
2 maximum in the mid-troposphere in the Arctic. However, the lower-latitude 4-5 km maximum
3 is hardly reproduced by any of the models. One likely reason for the modeling problems is the
4 strong biomass burning activity during spring 2008, which influenced a substantial fraction of
5 the measurement data (Warneke et al., 2010; Brock et al., 2011). Even though this should be
6 reflected in the GFED emission data for 2008, it seems possible that the GFED emissions are
7 underestimated. Furthermore, as some of the flights targeted biomass burning plumes
8 specifically, the influence of the biomass burning may be enhanced in the measurement data
9 compared to the models, especially if the models did not capture the plume transport well
10 enough and thus potentially simulated the biomass burning plumes at other locations than
11 observed. This sampling bias is particularly strong for the CCMs which are not driven by
12 observed meteorological fields.

13 Comparisons like those shown in Fig. 7 were also performed for the other aircraft campaigns.
14 For the sake of brevity, we further aggregate the data and only show results for latitudes north
15 of 70°N and for median values below and above 3 km altitude (Fig. 8). For spring 2008, the
16 aggregate plots for BC (Fig. 8e-f) show even more clearly than Fig. 7 that all models except
17 CanAM4.2 underestimate the measured rBC concentrations both at low and high altitudes. The
18 spring 2009 PAMARCMiP campaign, however, shows a different picture (Fig. 8c-d). This
19 campaign was influenced very little by biomass burning. The measured median rBC mass
20 concentrations at low (high) altitudes were about a factor two (three) lower than for the spring
21 2008 campaigns. Most models also simulated lower median BC concentrations than a year
22 earlier but the modeled reductions were less pronounced than the measured ones and, thus,
23 about half of the models under- and the other half overestimated the measured median values.
24 The vertical gradient of measured BC was also different in 2008 and in 2009. While in spring
25 2008, the concentrations above 3 km were higher than those below, the opposite was true in
26 spring 2009, likely because of the weaker biomass burning influence in 2009. This feature can
27 be seen very clearly in the vertical profiles shown in Fig. 9 and it is not well captured by the
28 models, most of which showed a relatively flat vertical BC distribution.

29 The concentrations measured by the ARCTAS summer campaign in 2008 are much lower than
30 those measured in spring 2008 and 2009, both at low and high altitudes (Fig. 8g-h), which is in
31 agreement with the seasonality seen at the surface stations. Some of the models under- and
32 others overestimate the measured concentrations, with the majority of the models

1 overestimating, especially below 3 km. The mean values, averaged over all models, are about
2 two (three) times as high as the measurements for altitudes above (below) 3 km. Some of the
3 models reproduce the measured rBC maximum at 6 km (Fig. 9).

4 The HIPPO campaign in fall 2009 (Fig. 8i-j) was conducted about one month after the seasonal
5 minimum at most surface sites and measured very low rBC mass concentrations, which is
6 consistent with the surface observations. Most of the models overestimate the measured
7 concentrations throughout the entire vertical profile (Fig. 9).

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

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

26 For sulfate, measured median concentrations in the Arctic during spring 2008 were lower above
27 3 km than below 3 km (Fig. 10a-b). All models, except CanAM4.2, strongly underestimate the
28 measured sulfate concentrations, some models by more than an order of magnitude. This is
29 consistent with the findings from the surface station comparisons (Fig. 6, Table 2). The models
30 also do not give a consistent picture of the vertical distribution of sulfate, with some models
31 correctly simulating lower concentrations above 3 km than below but others giving the opposite
32 result. The model underestimates for sulfate are likely not related to a sampling bias towards

1 frequent encounters of biomass burning plumes, as biomass burning plumes are relatively poor
2 in sulfate (e.g. Brock et al., 2011). Instead, the underestimation suggests other missing sulfur
3 sources or a too quick removal of sulfate from the atmosphere. Indeed, the latter would be
4 consistent with the suggestion of Kristiansen et al. (2012) that sulfate lifetimes in models are
5 too short in spring.

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

12

13 **6 Station vs. low-altitude aircraft measurements**

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

21 To investigate how consistent a picture the aircraft campaigns give vis a vis the station
22 measurements, we compare all aircraft data from the lowest 3 km and lowest 1 km to the values
23 obtained from the surface stations for the same months (Fig. 11). Selecting data only for even
24 lower altitudes is problematic as the data coverage becomes very poor. In Fig. 11, we also show
25 the station measurements obtained for the years 2008 and 2009 separately. For eBC, the
26 measurements obtained for the same month at the different stations and during different years
27 are (with a few exceptions such as Barrow in January 2008) quite comparable with each other.
28 In particular, April 2008 did not show higher eBC values than April 2009. This is consistent
29 with the finding that the biomass burning layers in 2008 did not extend to the surface (Brock et
30 al., 2011). At Alert, the EC values are similar to the eBC values, whereas the Station Nord EC
31 values are in summer and fall higher than eBC values at other stations. The aircraft rBC
32 measurements for all campaigns show consistently lower values than the eBC or EC

1 measurements at the ground, except for the HIPPO campaign in January 2009 where, however,
2 the data coverage particularly below 1 km is poor. It is possible that the BC concentrations
3 show a strong gradient in the lowest 1 km and that surface concentrations are indeed
4 systematically higher than concentrations just aloft. However, an alternative explanation could
5 be that the rBC measurements are biased low against the eBC or EC measurements, given the
6 different measurement techniques used. A direct comparison of all three measurement
7 techniques at the Alert station also suggests a low bias of rBC against eBC and EC
8 concentrations (S. Sharma, personal communication). For sulfate (Fig. 12) the measurements
9 show a much larger variability than for BC, both between stations and between the two different
10 years. For instance, the 25th percentile of the sulfate concentrations at Alert in January 2009 is
11 higher than the 75th percentile of the other stations and also of Alert in January 2008. On the
12 other hand, the sulfate concentrations measured during the two available flight campaigns in
13 spring and summer 2008 are not systematically different from those measured at the stations,
14 although the median concentration in summer 2008 is somewhat lower than at the stations. This
15 is consistent with the eBC or rBC differences.

16

17 **7 Sulfate/BC correlations**

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

1 Figure 13 shows correlation plots between monthly mean sulfate and eBC for the measurements
2 and the models sampled at the different stations. In the observations, sulfate and eBC
3 correlations for Alert, Pallas and Zeppelin are statistically significant at the 99.9% level (Table
4 3). The slopes of the regression lines shown in Fig. 13 are reported in Table 3. For the
5 observations, they are very similar: 10.1, 8.4 and 8.9 $\text{ng}[\text{SO}_4]/\text{m}^3 / \text{ng}[\text{eBC}]/\text{m}^3$ for Alert, Pallas
6 and Zeppelin, respectively. For Barrow, where the correlation is not significant because of two
7 eBC-rich outlier data points, the slope is smaller ($6.4 \text{ ng}[\text{SO}_4]/\text{m}^3 / \text{ng}[\text{eBC}]/\text{m}^3$). The strong
8 correlation between sulfate and eBC and the similarity of the slopes suggests that the sources
9 contributing to the measurements at the different stations are similar and that the removal of
10 sulfate and eBC is highly correlated, which would be expected for internally mixed aged aerosol
11 as is typical for the Arctic.

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

24 Based on the ECLIPSE inventory which is available for BC and for SO_2 , we estimated ratios
25 between those two substances under the assumption that all SO_2 is converted to sulfate. The
26 SO_2 to BC emission ratio of anthropogenic emissions in the ECLIPSE inventory is 25 globally
27 and 40 north of 50°N . For the GFED biomass burning emissions the emission ratio is only 1.7
28 globally and 2.5 north of 50°N , and for the sum of anthropogenic and biomass burning
29 emissions, we obtain ratios of 19 globally and 25 north of 50°N . The mean observed slopes of
30 the observations ($9.1 \text{ ng}[\text{SO}_4]/\text{m}^3 / \text{ng}[\text{eBC}]/\text{m}^3$) and the slopes modeled by DEHM (5.4
31 $\text{ng}[\text{SO}_4]/\text{m}^3 / \text{ng}[\text{BC}]/\text{m}^3$), CESM1-CAM5.2 ($9.9 \text{ ng}[\text{SO}_4]/\text{m}^3 / \text{ng}[\text{BC}]/\text{m}^3$) and WRF-Chem
32 ($8.5 \text{ ng}[\text{SO}_4]/\text{m}^3 / \text{ng}[\text{BC}]/\text{m}^3$) are much lower than the emission ratio of anthropogenic

1 emissions in the ECLIPSE inventory and they are also lower than the emission ratio for mixed
2 anthropogenic and biomass burning emissions. This suggests that biomass burning emissions
3 are relatively more important in the Arctic than elsewhere, that there are missing BC sources,
4 that sulfur emissions are overestimated (although this is not so likely, given the too low SO₂
5 emissions in high latitude Russia in the ECLIPSE version 4a inventory used here), and/or that
6 there exists a mechanism that enriches aerosols in BC relative to sulfate in the Arctic
7 atmosphere. The latter could be related to the hydrophobic nature of freshly emitted BC.

8 **Conclusions**

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

- 11 • The simulation of BC concentrations in the Arctic has improved compared to earlier
12 studies (e.g. Shindell et al., 2008; Koch et al., 2009; AMAP, 2011). For instance, our
13 model-mean underestimate of Arctic eBC at Barrow and Alert is about a factor of 2,
14 compared to one order of magnitude reported in Shindell et al. (2008). Nevertheless, the
15 aerosol seasonality at the surface is still too weak in most models. Concentrations of
16 eBC and sulfate averaged over three surface sites in the western Arctic are
17 underestimated in winter/spring in all but one model (model means for January-March
18 underestimated by 59% and 37% for BC and sulfate), whereas concentrations in summer
19 are overestimated in the model mean (by 88% and 44% for July-September), but with
20 over- as well as underestimates present in individual models.
- 21 • For the aircraft campaigns, the models overestimated measured rBC during all seasons
22 except for spring and throughout the depth of the troposphere. In spring 2009, no
23 overestimate was found, and in spring 2008 the models underestimated both rBC and
24 sulfate strongly. For rBC, this could have been due to underestimation of the strong
25 influence of biomass burning emissions observed during that campaign. The largest
26 eBC underestimates are found for the station Tiksi, which is closest to potential Russian
27 source regions and where the annual mean eBC concentration is three times higher than
28 the average annual mean for all other stations. This suggests an underestimate of BC
29 sources in Russia in the emission inventory used, even though this inventory contains
30 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.

1 **References**

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

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

9 AMAP: AMAP Assessment 2006: Acidifying Pollutants, Arctic Haze, and Acidification in the
10 Arctic. Arctic Monitoring and Assessment Programme (AMAP), Oslo, Norway. xii+112 pp,
11 2006.

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

14 Andersson, C., Langner, J., and Bergstrom, R.: Interannual variation and trends in air pollution
15 over Europe due to climate variability during 1958-2001 simulated with a regional CTM
16 coupled to the ERA40 reanalysis, *Tellus Series B-Chemical and Physical Meteorology*, 59, 77-
17 98, 10.1111/j.1600-0889.2006.00196.x, 2007.

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

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

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

1 Bond, T. C., and Bergstrom, R. W.: Light absorption by carbonaceous particles: An
2 investigative review, *Aerosol Science and Technology*, 40, 27-67,
3 10.1080/02786820500421521, 2006.

4 Brandt, J., Silver, J. D., Frohn, L. M., Geels, C., Gross, A., Hansen, A. B., Hansen, K. M.,
5 Hedegaard, G. B., Skjoth, C. A., Villadsen, H., Zare, A., and Christensen, J. H.: An integrated
6 model study for Europe and North America using the Danish Eulerian Hemispheric Model with
7 focus on intercontinental transport of air pollution, *Atmospheric Environment*, 53, 156-176,
8 10.1016/j.atmosenv.2012.01.011, 2012.

9 Breider, T. J., Mickley, L. J., Jacob, D. J., Wang, Q., Fisher, J. A., Chang, R. Y. W., and
10 Alexander, B.: Annual distributions and sources of Arctic aerosol components, aerosol optical
11 depth, and aerosol absorption, *Journal of Geophysical Research-Atmospheres*, 119, 4107-4124,
12 2014.

13 Brock, C. A., Cozic, J., Bahreini, R., Froyd, K. D., Middlebrook, A. M., McComiskey, A.,
14 Brioude, J., Cooper, O. R., Stohl, A., Aikin, K. C., de Gouw, J. A., Fahey, D. W., Ferrare, R.
15 A., Gao, R. S., Gore, W., Holloway, J. S., Hubler, G., Jefferson, A., Lack, D. A., Lance, S.,
16 Moore, R. H., Murphy, D. M., Nenes, A., Novelli, P. C., Nowak, J. B., Ogren, J. A., Peischl, J.,
17 Pierce, R. B., Pilewskie, P., Quinn, P. K., Ryerson, T. B., Schmidt, K. S., Schwarz, J. P.,
18 Sodemann, H., Spackman, J. R., Stark, H., Thomson, D. S., Thornberry, T., Veres, P., Watts,
19 L. A., Warneke, C., and Wollny, A. G.: Characteristics, sources, and transport of aerosols
20 measured in spring 2008 during the aerosol, radiation, and cloud processes affecting Arctic
21 Climate (ARCPAC) Project, *Atmospheric Chemistry and Physics*, 11, 2423-2453, 10.5194/acp-
22 11-2423-2011, 2011.

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

26 Cappa, C., Lack, D., Burkholder, J., and Ravishankara, A.: Bias
27 in filter based aerosol light absorption measurements due to organic aerosol loading: evidence
28 from laboratory measurements, *Aerosol Sci. Technol.*, 42, 1022–1032,
29 doi:[10.1080/02786820802389285](https://doi.org/10.1080/02786820802389285), 2008.

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

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

7 Christensen, J. H.: The Danish Eulerian hemispheric model - A three-dimensional air pollution
8 model used for the Arctic, *Atmospheric Environment*, 31, 4169-4191, 10.1016/s1352-
9 2310(97)00264-1, 1997.

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

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

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

21 Gadhavi, H. S., Renuka, K., Kiran, V. R., Jayaraman, A., Stohl, A., Klimont, Z., and Beig, G.:
22 Evaluation of black carbon emission inventories using a Lagrangian dispersion model - a case
23 study over southern India. *Atmos. Chem. Phys.* 15, 1447-1461, doi:10.5194/acp-15-1447-2015,
24 2015.

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

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

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

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

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

10 Hewitt, H. T., Copsey, D., Culverwell, I. D., Harris, C. M., Hill, R. S. R., Keen, A. B., McLaren,
11 A. J., and Hunke, E. C.: Design and implementation of the infrastructure of HadGEM3: the
12 next-generation Met Office climate modelling system, *Geoscientific Model Development*, 4,
13 223-253, 10.5194/gmd-4-223-2011, 2011.

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

18 Huang, Lin., J.R. Brook, W. Zhang, S-M. Li, L. Graham, D. Ernst, A. Chivulescu and G. Lu.
19 Stable isotope measurements of carbon fractions (OC/EC) in airborne particulate: A new
20 dimension for source characterization and apportionment. *Atmospheric Environment* 40: 2690–
21 2705, 2006

22 Huang, L., Gong, S. L., Jia, C. Q., and Lavoue, D.: Relative contributions of anthropogenic
23 emissions to black carbon aerosol in the Arctic, *Journal of Geophysical Research-Atmospheres*,
24 115, 11, 10.1029/2009jd013592, 2010.

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

1 Jacobson, M. Z.: Climate response of fossil fuel and biofuel soot, accounting for soot's feedback
2 to snow and sea ice albedo and emissivity, *Journal of Geophysical Research-Atmospheres*, 109,
3 10.1029/2004jd004945, 2004.

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

8 Klimont, Z., Smith, S. J., and Cofala, J.: The last decade of global anthropogenic sulfur dioxide:
9 2000-2011 emissions, *Environmental Research Letters*, 8, 10.1088/1748-9326/8/1/014003,
10 2013.

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

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

16 Koch, D., Schulz, M., Kinne, S., McNaughton, C., Spackman, J. R., Balkanski, Y., Bauer, S.,
17 Berntsen, T., Bond, T. C., Boucher, O., Chin, M., Clarke, A., De Luca, N., Dentener, F., Diehl,
18 T., Dubovik, O., Easter, R., Fahey, D. W., Feichter, J., Fillmore, D., Freitag, S., Ghan, S.,
19 Ginoux, P., Gong, S., Horowitz, L., Iversen, T., Kirkevag, A., Klimont, Z., Kondo, Y., Krol,
20 M., Liu, X., Miller, R., Montanaro, V., Moteki, N., Myhre, G., Penner, J. E., Perlitz, J., Pitari,
21 G., Reddy, S., Sahu, L., Sakamoto, H., Schuster, G., Schwarz, J. P., Seland, O., Stier, P.,
22 Takegawa, N., Takemura, T., Textor, C., van Aardenne, J. A., and Zhao, Y.: Evaluation of black
23 carbon estimations in global aerosol models, *Atmospheric Chemistry and Physics*, 9, 9001-
24 9026, 2009.

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

29 Lack, D. A., Cappa, C. D., Covert, D. S., Baynard, T., Massoli, P., Sierau, B., Bates, T. S.,
30 Quinn, P. K., Lovejoy, E. R., and Ravishankara, A. R.: Bias in filter based aerosol light ab-
31 sorption measurements due to organic aerosol loading: evidence from ambient measurements,

1 Aerosol Sci. Tech., 42, 1033–1041, doi:10.1080/02786820802389277, 2008. Lamarque, J. F.,
2 Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A.,
3 Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O.
4 R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.:
5 Historical (1850-2000) gridded anthropogenic and biomass burning emissions of reactive gases
6 and aerosols: methodology and application, Atmospheric Chemistry and Physics, 10, 7017-
7 7039, 10.5194/acp-10-7017-2010, 2010.

8 Law, K. S., and Stohl, A.: Arctic air pollution: Origins and impacts, Science, 315, 1537-1540,
9 10.1126/science.1137695, 2007.

10 Law, K. S., Stohl, A., Quinn, P. K., Brock, C. A., Burkhardt, J. F., Paris, J.-D., Aucellet, G., Singh, H.
11 B., Roiger, A., Schlager, H., Dibb, J., Jacob, D. J., Arnold, S. R., Pelon, J., and Thomas, J. L.: Arctic air
12 pollution: New insights from POLARCAT-IPY. *Bull. Am. Met. Soc.* **95**, 1873-1895, doi:
13 <http://dx.doi.org/10.1175/BAMS-D-13-00017.1>, 2014.

14 Liu, J. F., Fan, S. M., Horowitz, L. W., and Levy, H.: Evaluation of factors controlling long-
15 range transport of black carbon to the Arctic, *Journal of Geophysical Research-Atmospheres*,
16 116, 15, 10.1029/2010jd015145, 2011.

17 Lund, M. T., and Berntsen, T.: Parameterization of black carbon aging in the OsloCTM2 and
18 implications for regional transport to the Arctic, *Atmospheric Chemistry and Physics*, 12, 6999-
19 7014, 10.5194/acp-12-6999-2012, 2012.

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

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

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

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

12 Petzold, A., Ogren, J. A., Fiebig, M., Laj, P., Li, S.-M., Baltensperger, U., Holzer-Popp, T.,
13 Kinne, S., Pappalardo, G., Sugimoto, N., Wehrli, C., Wiedensohler, A., and Zhang, X.-Y.:
14 Recommendations for reporting "black carbon" measurements, *Atmos. Chem. Phys.*, 13, 8365-
15 8379, doi:10.5194/acp-13-8365-2013, 2013.

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

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

22 Quinn, P. K., Shaw, G., Andrews, E., Dutton, E. G., Ruoho-Airola, T., and Gong, S. L.: Arctic
23 haze: current trends and knowledge gaps, *Tellus Series B-Chemical and Physical Meteorology*,
24 59, 99-114, 10.1111/j.1600-0889.2006.00238.x, 2007.

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

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

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

6 Robertson, L., Langner, J., and Engardt, M.: An Eulerian limited-area atmospheric transport
7 model, *Journal of Applied Meteorology*, 38, 190-210, 10.1175/1520-
8 0450(1999)038<0190:aelaat>2.0.co;2, 1999.

9 Samset, B. H., Myhre, G., Herber, A., Kondo, Y., Li, S.-M., Moteki, N., Koike, M., Oshima, N.,
10 Schwarz, J. P., Balkanski, Y., Bauer, S. E., Bellouin, N., Berntsen, T. K., Bian, H., Chin, M.,
11 Diehl, T., Easter, R. C., Ghan, S. J., Iversen, T., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X.,
12 Penner, J. E., Schulz, M., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., and
13 Zhang, K.: Modelled black carbon radiative forcing and atmospheric lifetime in AeroCom
14 Phase II constrained by aircraft observations, *Atmos. Chem. Phys.*, 14, 12465-12477,
15 doi:10.5194/acp-14-12465-2014, 2014.

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

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

24 Sharma, S., Andrews, E., Barrie, L. A., Ogren, J. A., and Lavoue, D.: Variations and sources of
25 the equivalent black carbon in the high Arctic revealed by long-term observations at Alert and
26 Barrow: 1989-2003, *Journal of Geophysical Research-Atmospheres*, 111,
27 10.1029/2005jd006581, 2006.

28 Sharma, S., Ishizawa, M., Chan, D., Lavoue, D., Andrews, E., Eleftheriadis, K., and Maksyutov,
29 S.: 16-year simulation of Arctic black carbon: Transport, source contribution, and sensitivity
30 analysis on deposition, *Journal of Geophysical Research-Atmospheres*, 118, 943-964,
31 10.1029/2012jd017774, 2013.

1 Shaw, G. E.: The Arctic Haze Phenomenon, Bulletin of the American Meteorological Society,
2 76, 2403 – 2412, 1995.

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

5 Shindell, D. T., Chin, M., Dentener, F., Doherty, R. M., Faluvegi, G., Fiore, A. M., Hess, P.,
6 Koch, D. M., MacKenzie, I. A., Sanderson, M. G., Schultz, M. G., Schulz, M., Stevenson, D.
7 S., Teich, H., Textor, C., Wild, O., Bergmann, D. J., Bey, I., Bian, H., Cuvelier, C., Duncan, B.
8 N., Folberth, G., Horowitz, L. W., Jonson, J., Kaminski, J. W., Marmer, E., Park, R., Pringle,
9 K. J., Schroeder, S., Szopa, S., Takemura, T., Zeng, G., Keating, T. J., and Zuber, A.: A multi-
10 model assessment of pollution transport to the Arctic, *Atmospheric Chemistry and Physics*, 8,
11 5353-5372, 2008.

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

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

18 Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M.,
19 Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U.,
20 Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth System
21 Model: ECHAM6-HAM2, *Journal of Advances in Modeling Earth Systems*, 5, 146-172,
22 10.1002/jame.20015, 2013.

23 Stohl, A.: Characteristics of atmospheric transport into the Arctic troposphere. *J. Geophys. Res.*
24 **111**, D11306, doi:10.1029/2005JD006888, 2006.

25 Stohl, A., Hittenberger, M., and Wotawa, G.: Validation of the Lagrangian particle dispersion
26 model FLEXPART against large scale tracer experiments. *Atmospheric Environment* **32**, 4245-
27 4264, 1998.

28 Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian
29 particle dispersion model FLEXPART version 6.2, *Atmospheric Chemistry and Physics*, 5,
30 2461-2474, 2005.

1 Stohl, A., Klimont, Z., Eckhardt, S., Kupiainen, K., Shevchenko, V. P., Kopeikin, V. M., and
2 Novigatsky, A. N.: Black carbon in the Arctic: the underestimated role of gas flaring and
3 residential combustion emissions, *Atmospheric Chemistry and Physics*, 13, 8833-8855,
4 10.5194/acp-13-8833-2013, 2013.

5 Stone, R. S., Herber, A., Vitale, V., Mazzola, M., Lupi, A., Schnell, R. C., Dutton, E. G., Liu,
6 P. S. K., Li, S. M., Dethloff, K., Lampert, A., Ritter, C., Stock, M., Neuber, R., and Maturilli,
7 M.: A three-dimensional characterization of Arctic aerosols from airborne Sun photometer
8 observations: PAM-ARCMIP, April 2009, *Journal of Geophysical Research-Atmospheres*,
9 115, 10.1029/2009jd013605, 2010.

10 Sullivan, A. P., Peltier, R. E., Brock, C. A., de Gouw, J. A., Holloway, J. S., Warneke, C.,
11 Wollny, A. G., and Weber, R. J.: Airborne measurements of carbonaceous aerosol soluble in
12 water over northeastern United States: Method development and an investigation into water-
13 soluble organic carbon sources, *Journal of Geophysical Research-Atmospheres*, 111, 14,
14 10.1029/2006jd007072, 2006.

15 Tunved, P., Ström, J., and Krejci, R.: Arctic aerosol life cycle: linking aerosol size distributions
16 observed between 2000 and 2010 with air mass transport and precipitation at Zeppelin station,
17 Ny-Ålesund, Svalbard, *Atmos. Chem. Phys.*, 13, 3643-3660, doi:10.5194/acp-13-3643-2013,
18 2013.

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

23 Vignati, E., M. Karl, M. Krol, J. C. Wilson, P. Stier and F. Cavalli, Sources of uncertainties in
24 modelling black carbon at the global scale, *Atmospheric Chemistry and Physics* 10, 2595 –
25 2611, 2010.

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

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

1 Wang, H., Easter, R. C., Rasch, P. J., Wang, M., Liu, X., Ghan, S. J., Qian, Y., Yoon, J. H., Ma,
2 P. L., and Vinoj, V.: Sensitivity of remote aerosol distributions to representation of cloud-
3 aerosol interactions in a global climate model, *Geoscientific Model Development*, 6, 765-782,
4 10.5194/gmd-6-765-2013, 2013.

5 Warneke, C., Froyd, K. D., Brioude, J., Bahreini, R., Brock, C. A., Cozic, J., de Gouw, J. A.,
6 Fahey, D. W., Ferrare, R., Holloway, J. S., Middlebrook, A. M., Miller, L., Montzka, S.,
7 Schwarz, J. P., Sodemann, H., Spackman, J. R., and Stohl, A.: An important contribution to
8 springtime Arctic aerosol from biomass burning in Russia, *Geophysical Research Letters*, 37,
9 10.1029/2009gl041816, 2010.

10 Wang, Q., D. J. Jacob, J. R. Spackman, A. E. Perring, J. P. Schwarz, N. Moteki, E. A. Marais,
11 C. Ge, J. Wang, and S. R. H. Barrett, Global budget and radiative forcing of black carbon
12 aerosol: Constraints from pole-to-pole (HIPPO) observations across the Pacific, *J. Geophys.*
13 *Res. Atmos.*, 119, 195–206, doi:10.1002/2013JD020824, [2014](#).

14 Wofsy, S. C., Team, H. S., Cooperating Modellers, T., and Satellite, T.: HIAPER Pole-to-Pole
15 Observations (HIPPO): fine-grained, global-scale measurements of climatically important
16 atmospheric gases and aerosols, *Philosophical Transactions of the Royal Society a-*
17 *Mathematical Physical and Engineering Sciences*, 369, 2073-2086, 10.1098/rsta.2010.0313,
18 2011.

19 Zaveri, R. A., and Peters, L. K.: A new lumped structure photochemical mechanism for large-
20 scale applications, *Journal of Geophysical Research-Atmospheres*, 104, 30387-30415,
21 10.1029/1999jd900876, 1999.

22 Zaveri, R. A., Easter, R. C., Fast, J. D., and Peters, L. K.: Model for Simulating Aerosol
23 Interactions and Chemistry (MOSAIC), *Journal of Geophysical Research-Atmospheres*, 113,
24 29, 10.1029/2007jd008782, 2008.

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

29

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. 1° Input data: global x1°92L	ECMWF Analyses; none	Operational 2008-2009 3h	Stohl et al. (1998, 2005)
OsloCTM2	CTM	2.8°x2.8°, global	60L ECMWF IFS Forecasts ; aerosol externally mixed	2008-2009 3h	Myhre et al. (2009), Skeie et al. (2011a, 2011b)
NorESM	CCM	1.9°x2.5°, global	26L Internal, observed SST prescribed; BC internally mixed	2008-2009 3h	Kirkevåg et al. (2013), Bentsen et al. (2013)
TM4-ECPL	CTM	2°x3°, global	34L 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°, global	31L ECMWF Reanalysis; aerosols internally mixed	March-August, 2008, 1h	Stevens et al. (2013), Zhang et al. (2012)
SMHI-MATCH	CTM	0.57°x0.75°, 20-90°N	38L ECMWF – ERA-Interim; BC internally mixed	2008, 1h	2009 Andersson et al. (2007), Robertson et al. (1999)
CanAM4.2	ACM	2.8°x2.8°, global	49L, 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, 0-90°N	29L NCEP; internally mixed aerosols	2008-2009 3h	Christensen (1997), Brandt et al. (2012)
CESM1/CAM5.2	CCM	1.9°x2.5°, global	30L Internal, observed SST prescribed; internally mixed aerosols	2008-2009 1h	Liu et al. (2012), Wang et al. (2013)
WRF-Chem	RCCM	100kmx100km 27-90° N	38L 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°, global	63L 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
4
5 ¹Chemistry transport model (CTM), Lagrangian particle dispersion model (LPDM), chemistry climate model (CCM), aerosol climate model (ACM), regional climate model coupled with a chemistry module (RCM)

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

Model/obs	Winter/Spring	Summer	Winter/Spring	Summer
	BC [ng/m^3]	BC [ng/m^3]	SO_4 [ng/m^3]	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

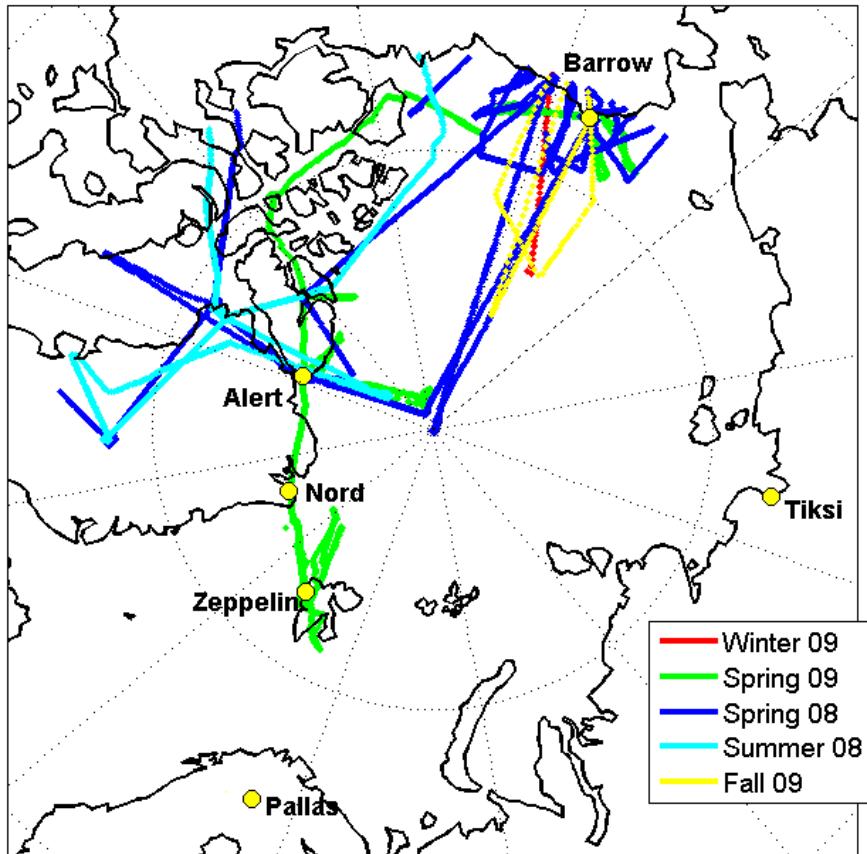
6
 7

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

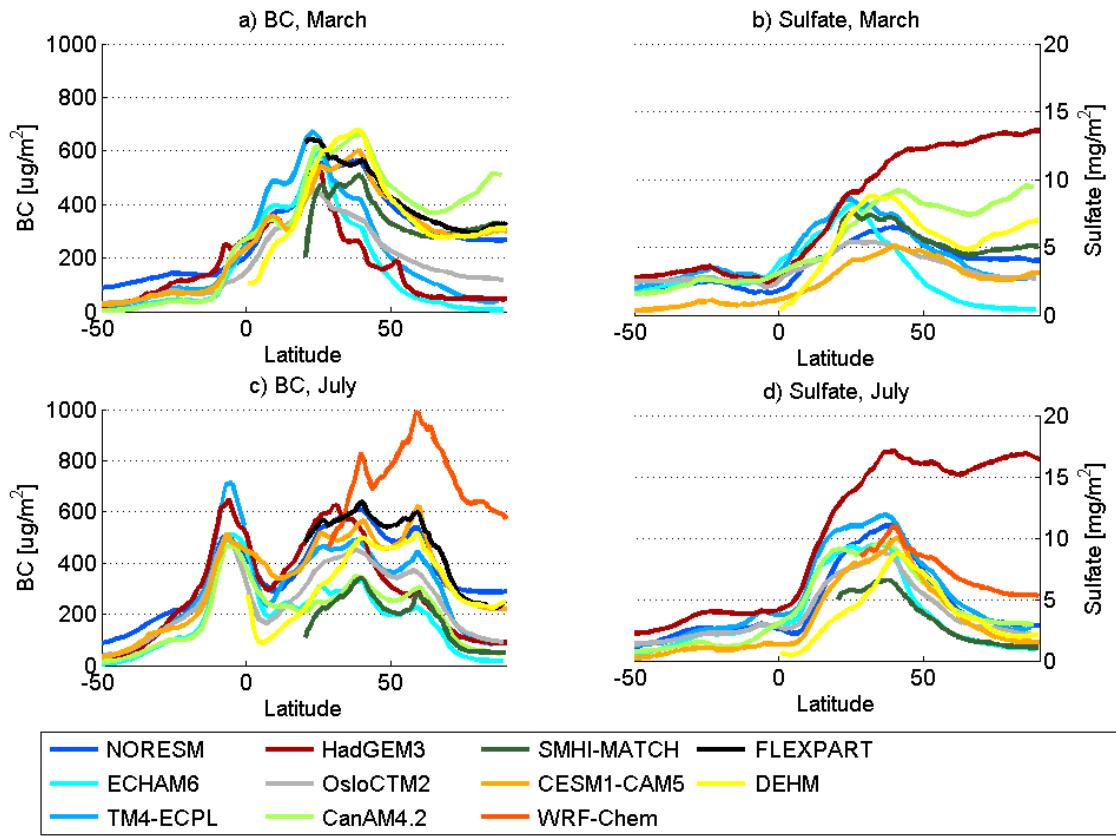
	Alert	Barrow	Pallas	Zeppelin	Mean
Observations	10.1	6.4	8.4	8.9	9.1
Model mean	17.3	16.6	6.7	9.7	12.6
OsloCTM2	-8.6	2.4	-2.0	-5.5	-
NorESM	35.3	27.8	0.4	12.1	35.3
TM4-ECPL	9.5	33.2	5.8	8.1	19.5
ECHAM6- HAM2	30.0	90.4	1.0	-746.4	-
SMHI- MATCH	25.6	25.9	0.4	10.9	25.7
CanAM4.2	18.2	2.5	7.1	12.4	15.3
DEHM	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	-

5

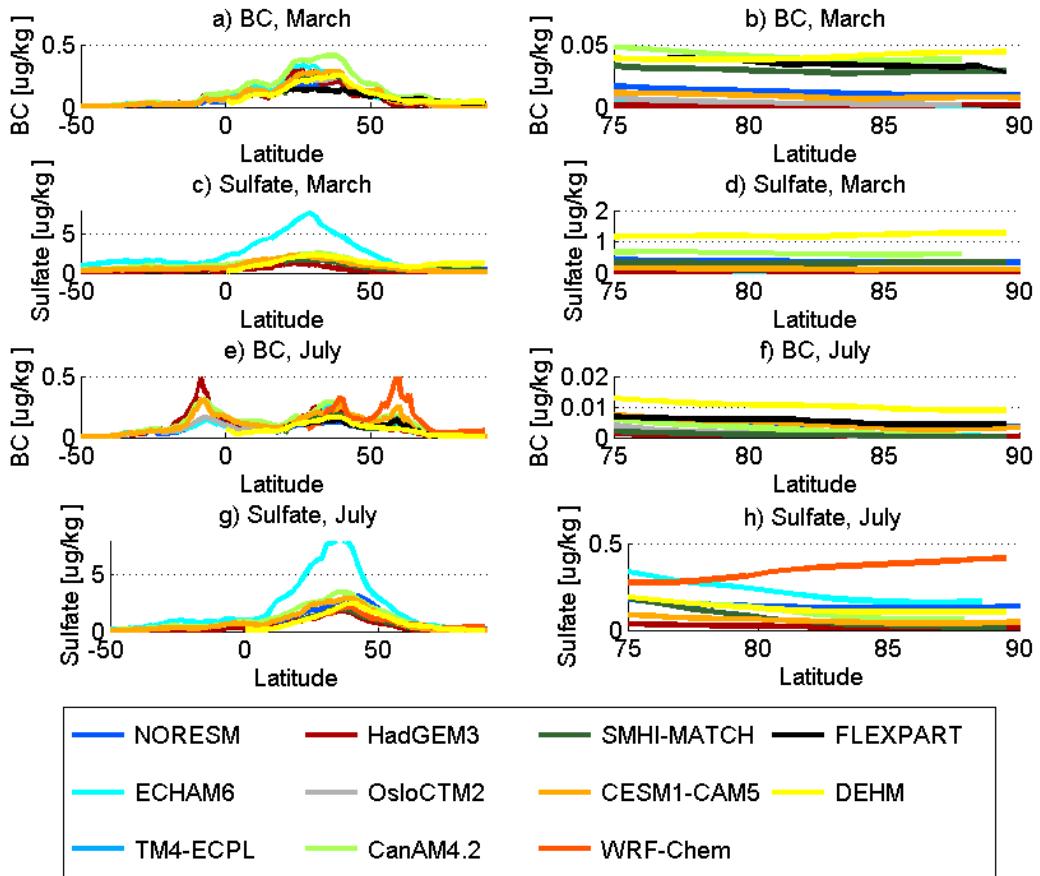
6



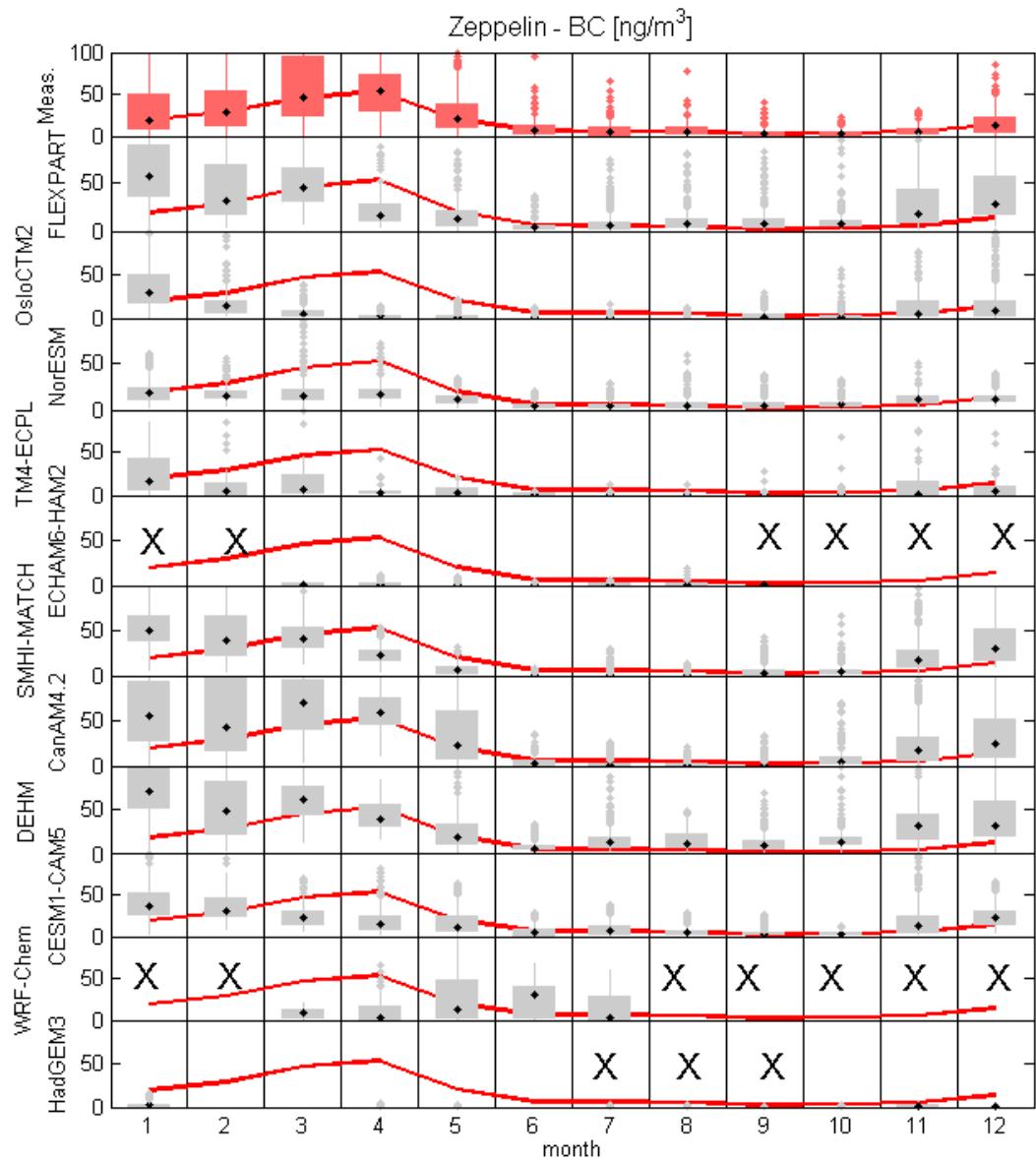
1
2 Figure 1. Map showing the locations of the measurement stations (yellow circles) and the flight
3 tracks north of 70°N of all aircraft campaigns used in this study. Aircraft data were from the
4 HIPPO (winter 2009 and fall 2009), ARCTAS (spring and summer 2008), ARCPAC (spring
5 2008) and PAMARCMiP (spring 2009) campaigns.
6



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



1
 2 Figure 3: BC (a-b, e-f) and sulfate (c-d, g-h) mass mixing ratios for the year 2008 at the
 3 surface averaged over all longitudes as a function of latitude (for the range 50°S to 90°N) for
 4 March (a-d) and July (e-h). The right panels show the same data as the left panels, but only
 5 for 70-90°N and with an adjusted ordinate scale.
 6



1
 2 Figure 4. Observed and simulated mean annual cycle of (equivalent) BC mass concentrations
 3 [ng/m^3] at the Zeppelin station. Shown is the monthly frequency distributions using data from
 4 the years 2008 and 2009. The uppermost panel (red boxes) shows monthly frequency
 5 distributions of the observed eBC concentrations. The other panels below (grey boxes) show
 6 monthly frequency distributions of the modeled BC concentrations. Black dots depict the
 7 monthly median value, the grey boxes span the range between the 25th and 75th percentiles,
 8 red and grey dots represent values which are outside the 1.5 fold of this interquartile range (grey
 9 lines). The red line connects the monthly medians of the observed eBC concentrations in the
 10 uppermost panel and is repeated in all other panels for the convenience of comparing modeled

1 and measured values. Missing model data are denoted with “X”. Notice that some models have
2 very low BC mass concentrations, which are difficult to see on the scale used.
3

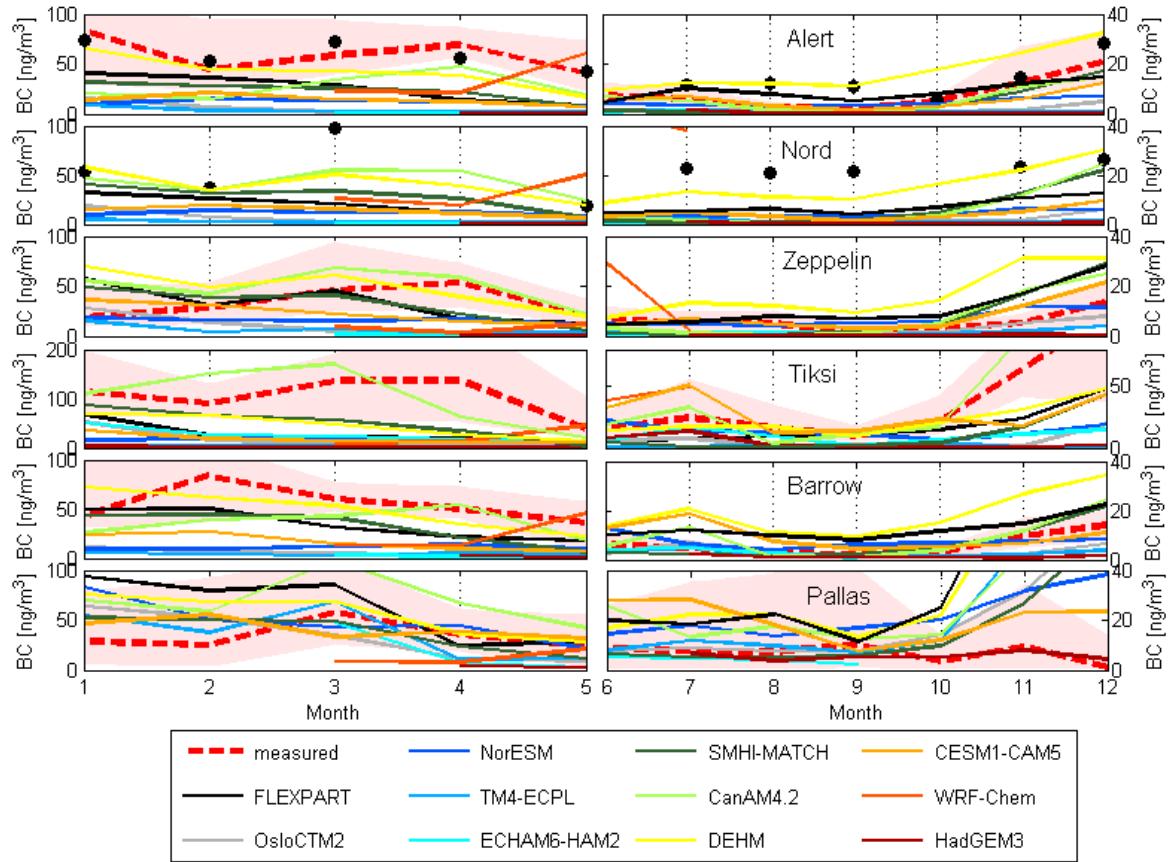
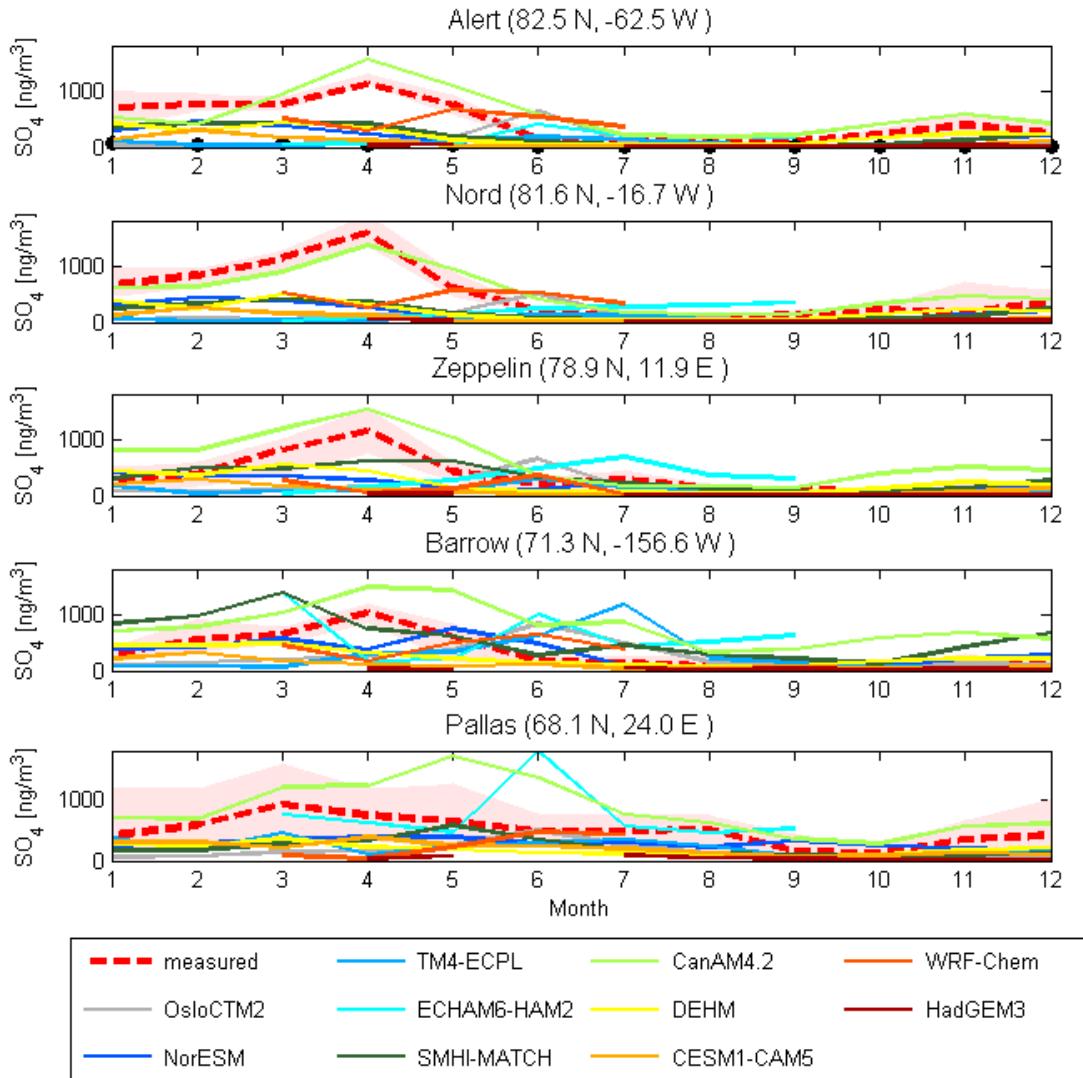


Figure 5. Surface concentrations of monthly (month is displayed on the abscissa) median observed eBC or EC and modeled BC. Each row represents one station: (from top) Alert, Nord, Zeppelin, Tiksi, Barrow and Pallas, for late winter/spring (left column) and summer/fall (right column). The red dashed lines 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.



1
2 Figure 6. Monthly (month is displayed on the abscissa) median observed and modeled sulfate
3 surface concentrations for the stations (from top) Alert, Nord, Zeppelin, Barrow and Pallas. The
4 red dashed lines connect the observed median values. The light red shaded areas span from the
5 25th to the 75th percentile of the observations. Modeled median values are shown with different
6 lines according to the legend.

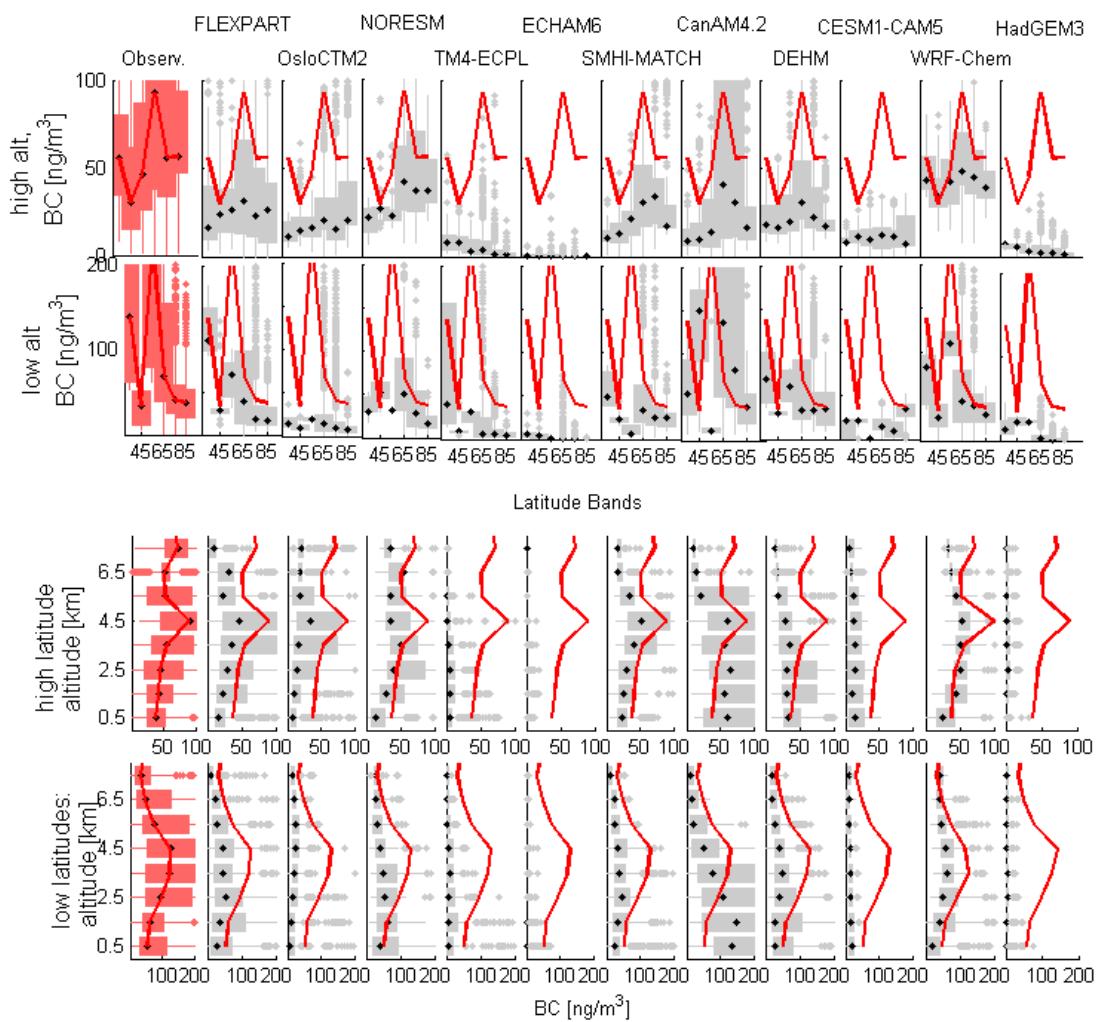
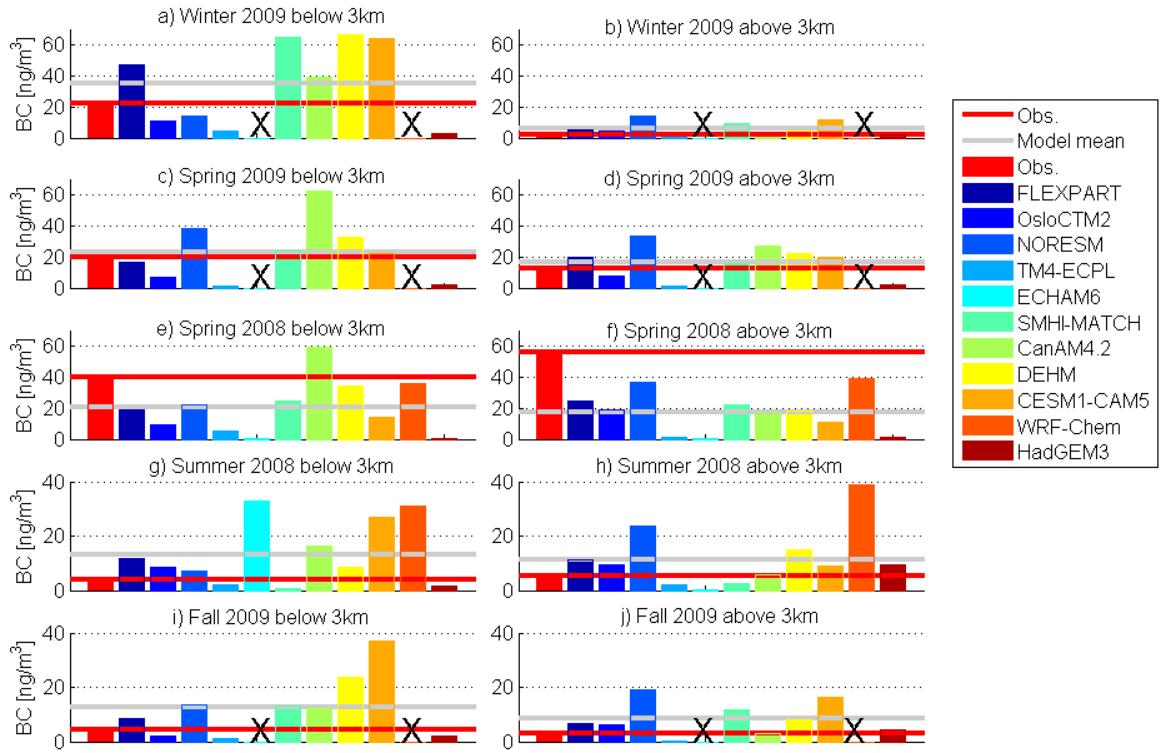


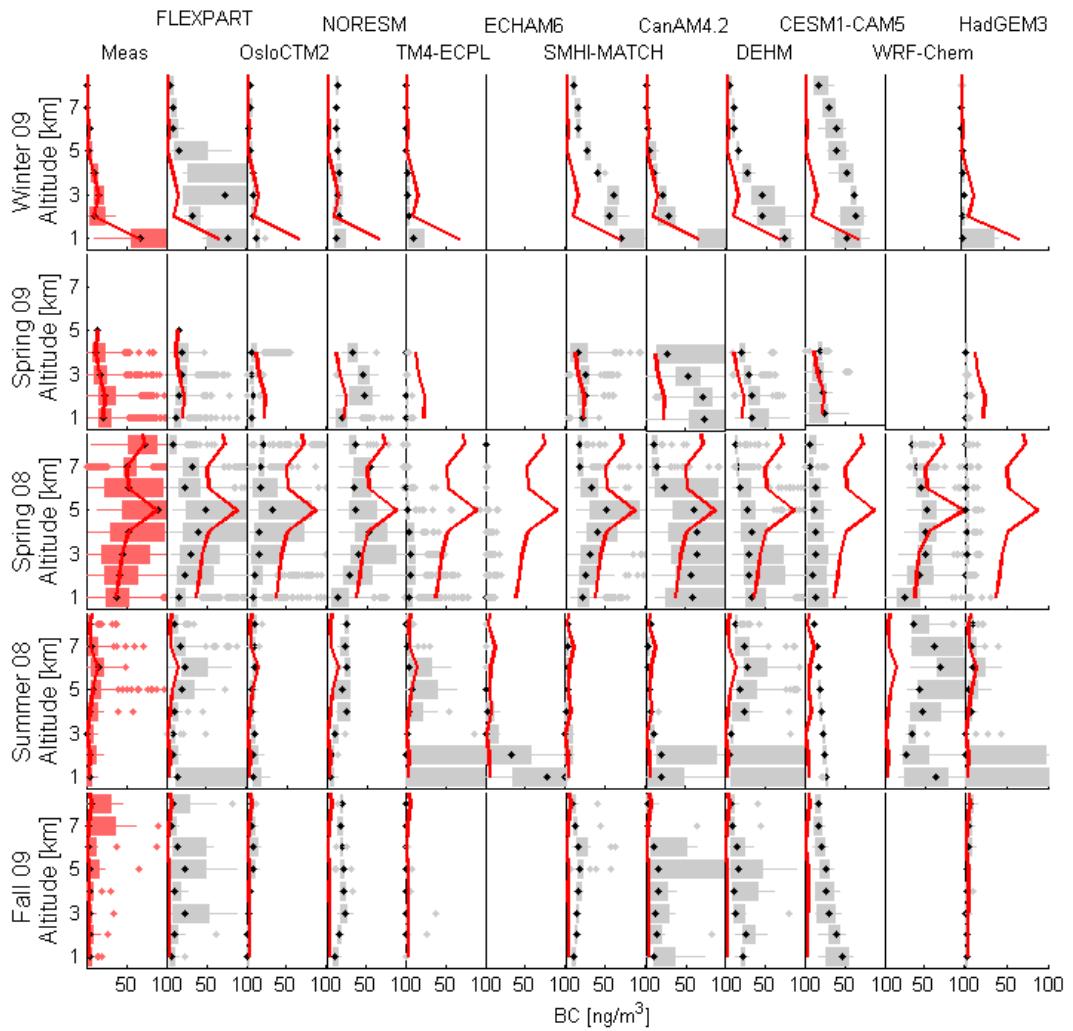
Figure 7: Comparison of modeled BC with observed rBC (red boxes and red lines) mass concentrations from the ARCTAS-spring and ARCPAC campaigns in spring 2008. The leftmost column shows box and whisker plots (like in Fig 4: boxes go from 25th to 75th percentile, whiskers span the 1.5fold interquartile 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°N as a function of altitude by binning the data into 1-km height intervals.

14

15



1
 2 Figure 8: Median observed rBC and modeled BC mass concentrations for the winter 2009
 3 HIPPO (a, b) spring 2009 PAMARCMiP (c-d) spring 2008 ARCTAS/ARCPAC (e-f) summer
 4 2008 ARCTAS (g-h) and the fall 2009 HIPPO (i-j) aircraft campaigns. The red bar and the red
 5 horizontal line show the observations, the other colored bars the various models, the grey line
 6 shows the mean value of all model medians. Results are shown separately for measurements
 7 below 3 km (left panels) and above 3 km (right panels). Notice that the concentration scales on
 8 the ordinates are different for the individual panels.

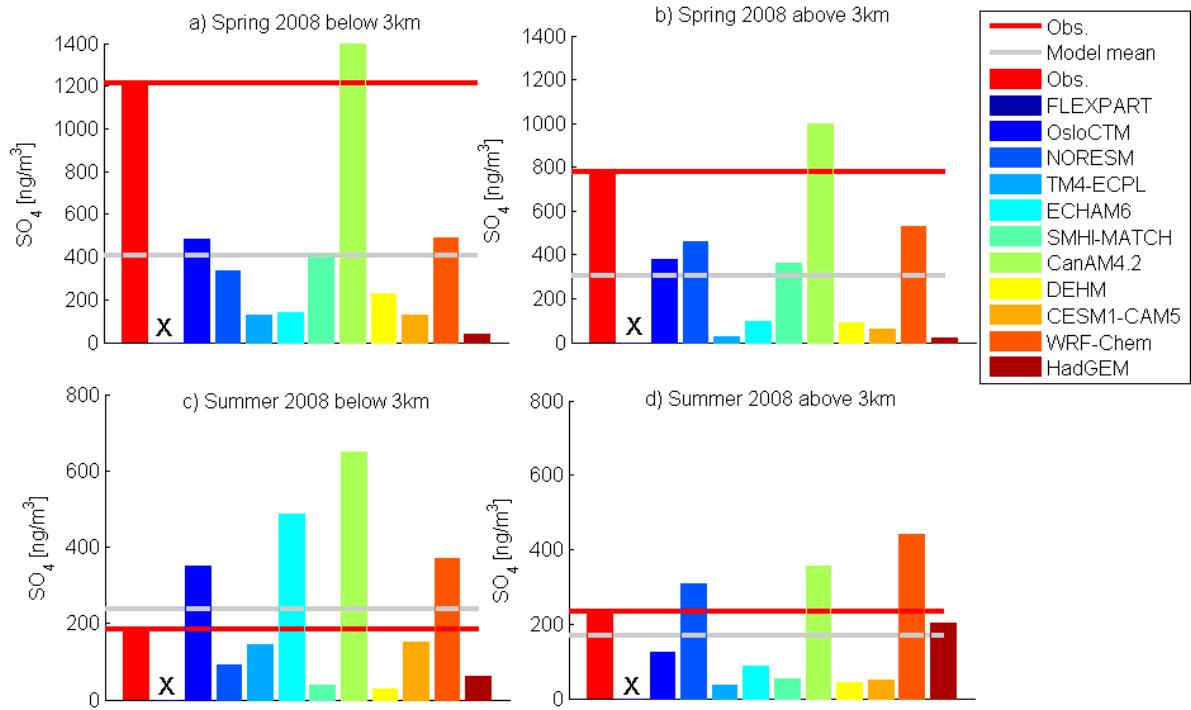


1

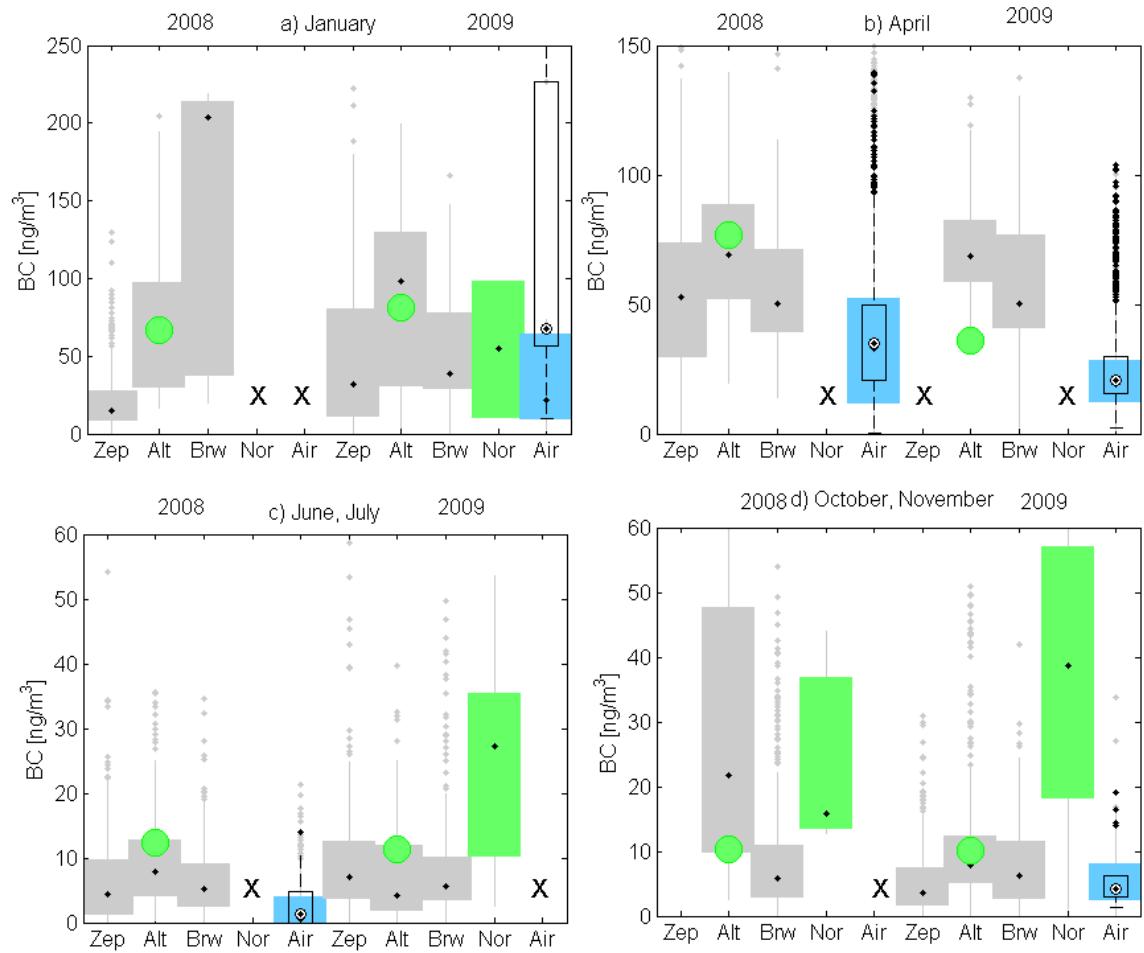
2 Figure 9: Comparison of modeled BC with observed rBC mass concentrations as a function of
 3 altitude for all data taken north of 70°N, for the different campaigns (same as in Fig. 8). The
 4 leftmost column shows box and whisker plots of observed rBC concentrations in ng/m^3 . The
 5 black dots as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

8

9

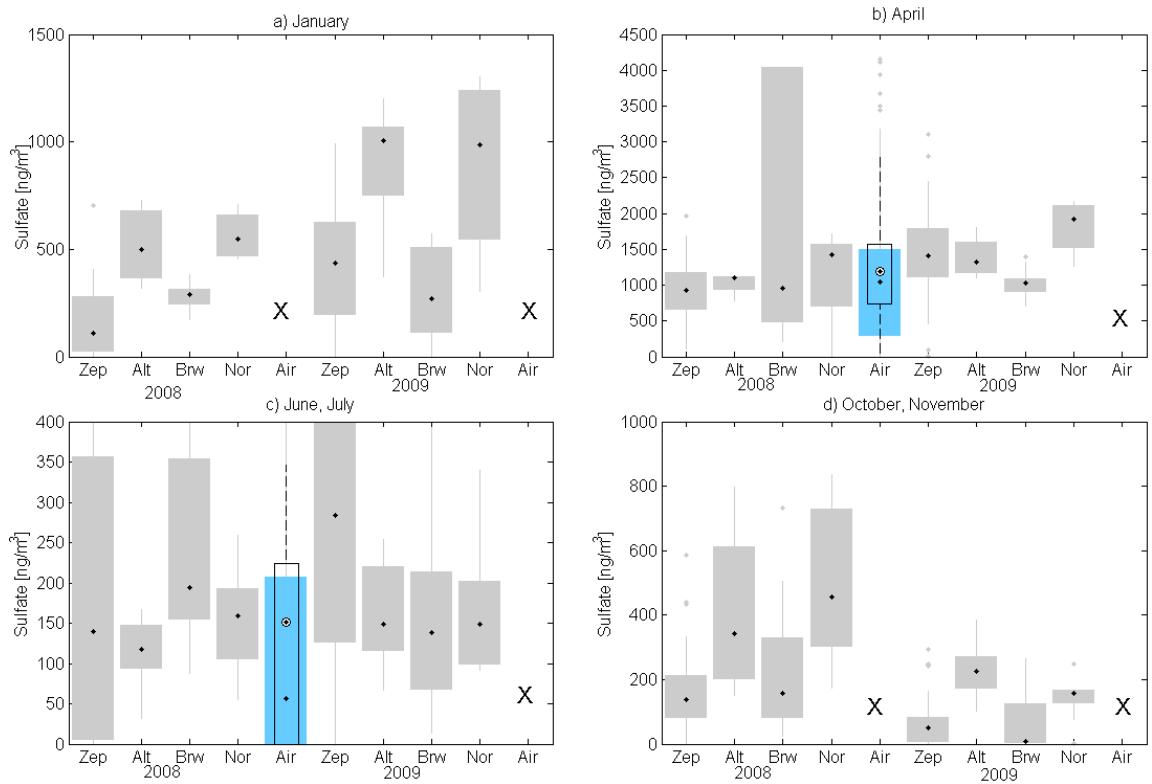


1
 2 Figure 10: Median SO₄ concentrations for the ARCTAS/ARCPAC spring 2008 (a-b) and
 3 ARCTAS summer 2008 (c-d) campaigns. The red bar and the red horizontal line show the
 4 observations, the other colored bars the various models. The analysis is performed for
 5 measurements below 3 km (left panels) and above 3 km (right panels). Note: each row has a
 6 different y-axis.
 7
 8



1
2 Figure 11: Comparison of eBC [ng/m³] measured at the stations Zeppelin (Zep), Alert (Alt),
3 and Barrow (Brw) (grey bars), EC measured at Alert and Station Nord (Nord) (green dots and
4 bars) and rBC [ng/m³] measured by aircraft (Air) in the lowest 3 km and 1 km, north of 70°N
5 (blue bars) for the years 2008 and 2009 for a) January, b) April, c) June and July and d) October
6 and November. The black dots represent the median, and the boxes the interquartile range. For
7 the aircraft measurements, the blue boxes show the results for the lowest 3 km, the black box
8 outlines show the results for the lowest 1 km.
9

1



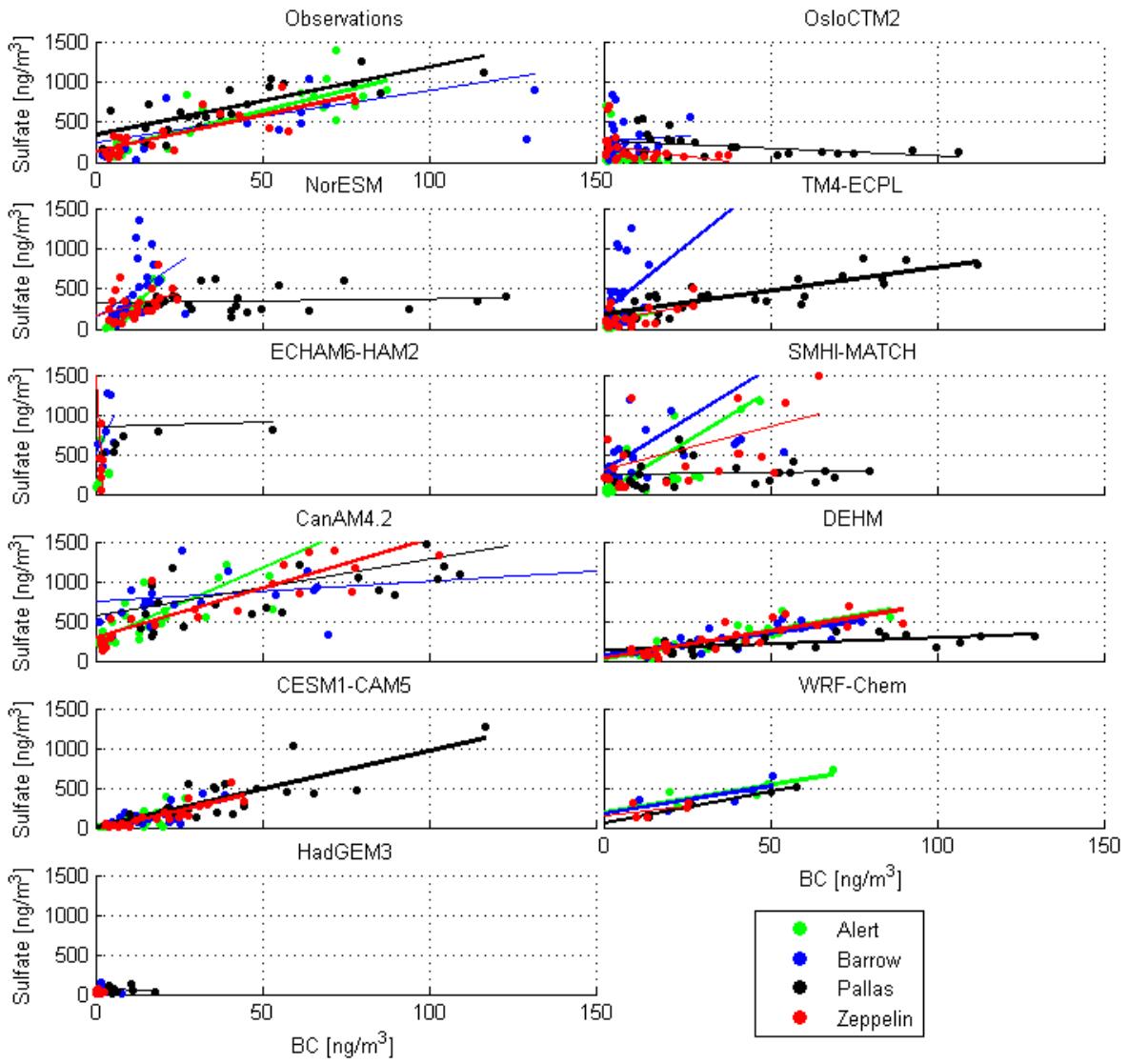
2

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

4

5

6



1
2 Figure 13: Correlation plots of monthly mean sulfate and (e)BC concentrations for the
3 observations (top left) and the different models sampled at the observation sites. Thick lines
4 denote significant correlations.
5
6