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

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

\textsuperscript{1}NILU – Norwegian Institute for Air Research, Kjeller, Norway
\textsuperscript{2}Sorbonne Universités, UPMC Univ. Paris 06, Université Versailles St-Quentin, CNRS/INSU; LATMOS-IPSL, UMR 8190, Paris, France
\textsuperscript{3}Norwegian Meteorological Institute, Oslo, Norway
Arctic black carbon and sulfate in model and observations

S. Eckhardt et al.

Center for International Climate and Environmental Research – Oslo (CICERO), Oslo, Norway
Institute for Meteorology, Universität Leipzig, Leipzig, Germany
ENVS Department of Environmental Science, Aarhus University, Roskilde, Denmark
Met Office Hadley Centre, Exeter, UK
University of Reading, Reading, UK
Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA
NOAA Earth System Research Laboratory Physical Sciences Division/Polar Observations & Processes, Boulder, Colorado, USA
Environmental Chemical Processes Laboratory, Department of Chemistry, University of Crete, Heraklion, Crete, and ICE-HT/FORTH, Patras, Greece
Department of Atmospheric, Oceanic, and Space Sciences, University of Michigan, Ann Arbor, MI, USA
Alfred Wegener Institut, Helmholtz Centre for Polar- and Marine Research, Bremerhaven, Germany
International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
Climate Research Division, Atmospheric Sci. & Tech. Directorate, S & T, Environment Canada Toronto, Ontario, Canada
Swedish Meteorological and Hydrological Institute (SMHI), 60176 Norrköping, Sweden
National Oceanic and Atmospheric Administration Pacific Marine Environmental Laboratory, Seattle, WA, USA
Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia, Canada
School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada
Department of Meteorology, COMSATS Institute of Information Technology, Islamabad, Pakistan
now at: Univ. Grenoble Alpes/CNRS, Laboratoire de Glaciologie et Géophysique de l’Environnement (LGGE), 38041 Grenoble, France
Received: 4 March 2015 – Accepted: 19 March 2015 – Published: 9 April 2015
Correspondence to: S. Eckhardt (sabine.eckhardt@nilu.no)
Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

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

1 Introduction

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

The Arctic aerosol consists of a varying mixture of sulfate and organic carbon (OC), as well as ammonium, nitrate, BC and mineral dust (Quinn et al., 2007; Brock et al., 2011). Aerosols in the Arctic feature a strong annual cycle with a late winter/spring peak (the so-called Arctic Haze) and a summer minimum. Increased transport during the cold season (Stohl, 2006) and increased removal by wet deposition during the warm season can explain this annual variation (Shaw, 1995; Law and Stohl, 2007) and also shape the aerosol size distribution (Tunved et al., 2013).
Models have for a long time struggled to capture the distribution of aerosols in the Arctic (Shindell et al., 2008; Koch et al., 2009). The concentrations of BC during the Arctic Haze season in particular were underestimated, in some cases by more than an order of magnitude (Shindell et al., 2008), whereas summer concentrations were sometimes overestimated. The simulated aerosol seasonality is strongly dependent on the model treatment of aerosol removal processes. For instance, changes in the calculation of aerosol microphysical properties, size distribution and removal can change simulated concentrations by more than an order of magnitude in remote regions such as the Arctic (Vignati et al., 2010) and the calculated Arctic BC mass concentrations are very sensitive to parameterizations of BC aging (conversion from hydrophobic to hydrophilic properties) and wet scavenging (Liu et al., 2011; Huang et al., 2010).

The seasonal decrease of aerosol concentrations from winter to summer in the Arctic is likely also due to the different efficiency of scavenging by different types of clouds. There is a transition from inefficient ice-phase cloud scavenging in winter to more efficient warm cloud scavenging in summer, and there is also the appearance of warm drizzling cloud in the late spring and summer boundary layer. Including these processes in one model clearly improved its performance both in terms of absolute concentrations as well as seasonality for sulfate and BC (Browse et al., 2012). This result is in agreement with the observation-based findings that scavenging efficiencies are increased in summer both for light-scattering (of which sulfate is an important component) as well as for light-absorbing (of which BC is an important component) aerosols (Garrett et al., 2010, 2011). Another modeling problem may be excessive convective transport and underestimation of the associated wet scavenging in convective clouds, which can lead to model overestimates of BC in the upper troposphere and lower stratosphere (Allen and Landuyt, 2014; Wang et al., 2014). Despite remaining difficulties, simulations of Arctic aerosols with many models have improved considerably in the last few years by updating the model treatment of some or all of the above mentioned processes (Fisher et al., 2011; Breider et al., 2014; Sharma et al., 2013; Lund and Berntsen, 2012; Allen and Landuyt, 2014).
Remaining problems may also be due to missing emission sources or incorrect spatial or temporal distribution of emissions in the inventories used for the modeling. The main sources of BC are biomass burning and incomplete combustion of fossil fuels and biofuels (Bond et al., 2004). Sulfate originates from natural sources such as oxidation of dimethyl sulfide (DMS) or sea salt over the oceans or volcanoes. It is also produced from oxidation of SO$_2$ emitted when sulfur-containing fossil fuels are burned or by metal smelting. Studies based on observed surface concentrations repeatedly suggest that the main source regions for Arctic BC and sulfate are located in high-latitude Eurasia (e.g., Sharma et al., 2006; Eleftheriadis, 2009; Hirdman et al., 2010). Stohl et al. (2013) suggested that gas flaring in high-latitude Russia is an important source of BC which is missing from most inventories. In their simulations, BC emissions from gas flaring accounted for 42% of the annual mean BC surface concentrations in the Arctic. However, they also noted the large uncertainty of the gas flaring emissions.

The radiative effects of aerosols are not so much determined by the surface concentrations but by the column loadings as well as the altitude distribution of the aerosol (Samset et al., 2013; Samset and Myhre, 2011). Nevertheless, in the past, model results for the Arctic were evaluated mainly against surface measurements due to their availability over long time periods. However, surface concentrations are not representative of concentrations aloft, which are controlled, at least in part, by different source regions and different processes. It is therefore important to evaluate models not only against surface measurements but also using vertical profile information.

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

2 Methods

2.1 Measurement data

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

Aerosol light absorption data were obtained from five sites in different parts of the Arctic: Alert, Canada (62.3° W, 82.5° N; 210 m above sea level (a.s.l.)), Zeppelin/Ny Ålesund, Spitsbergen, Norway (11.9° E, 78.9° N; 478 m a.s.l.), Tiksi, Russia (128.9° E, 71.6° N; 1 m a.s.l.), Barrow, Alaska (156.6° W, 71.3° N; 11 m a.s.l.) and Pallas, Finland (24.12° E, 67.97° N; 565 m a.s.l.). The locations of these measurement stations are shown in Fig. 1. Different types of particle soot absorption photometers (PSAPs) were used for the measurements at Barrow, Alert, and Zeppelin, a multi-angle absorption photometer was used at Pallas (Hyvärinen et al., 2011), and an aethalometer was used at Tiksi. All these instruments measure the particle light absorption coefficient $\sigma_{ap}$, each at its own specific wavelength (typically at around 530–550 nm), and for different size fractions of the aerosol (typically particles smaller than 1, 2.5 or 10 µm are sampled at different humidities). Conversion of $\sigma_{ap}$ to eBC mass concentrations is not straightforward and requires certain assumptions (Petzold et al., 2013). The mass absorption efficiency used for conversion can be specific to a site, the instrument and the wavelength used and is uncertain by at least a factor of two. For Tiksi, the conversion is
done internally by the aethalometer. For the other sites, a mass absorption efficiency of $10 \, \text{m}^2 \, \text{g}^{-1}$, typical of aged BC aerosol (Bond and Bergstrom, 2006), was used. Sharma et al. (2013) used the even higher value of $19 \, \text{m}^2 \, \text{g}^{-1}$ for Barrow and Alert data.

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

Barrow and Alert data are routinely subject to data cleaning, which shall remove the influence from local sources. The Tiksi data has been quality controlled as well and episodes of local pollution have been removed. Zeppelin generally is not strongly influenced by local emissions; however, summer values are enhanced by some 11 % due to local cruise ship emissions (Eckhardt et al., 2013). Thermo-optical measurements of EC were available from Station Nord, Greenland (16.67°W, 81.6°N; 30 m a.s.l.) and from Alert. At both stations, weekly aerosol samples were collected during 2008–2009. For Station Nord a Digitel DHA 80 high volume sampler (HVS, Digitel/Riemer Messtechnik, Germany) was used for PM$_{10}$. Both stations’ samples were analyzed with a thermal-optical Lab OC/EC instrument from Sunset Laboratory Inc (Tigard, OR, USA). Punches of 2.5 cm$^2$ were cut from the filters sampled at Station Nord and analyzed according to the EUSAAR-2 protocol (Cavalli et al., 2010). The method used for the Alert carbon isotope measurements in EC/OC with baseline-separation approach was originally developed Huang et al., 2006 and further optimized Chan et al., 2010.

Sulfate measurement data were available from the stations Pallas, Zeppelin, Barrow, Nord and Alert. The sulfate data were obtained on open face filters and cations and anions were subsequently quantified by ion chromatography. Non-sea salt (nss) sulfate concentrations were obtained by subtracting the sea salt contribution via analysis of Na$^+$ and Cl$^-$ data, thus making the sulfate data directly comparable to the modeled nss-sulfate values. For Station Nord, the contribution from sea salt is only minor (Hei-
Aircraft data were obtained from several campaigns. In the framework of POLARCAT (Polar Study using Aircraft, Remote Sensing, Surface Measurements, and Models of Climate Chemistry, Aerosols, and Transport; Law et al., 2014), two ARCTAS (Arctic Research of the Composition of the Troposphere from Aircraft and Satellites) campaigns in April and June/July 2008 with a DC-8 aircraft covered mainly the North American Arctic (Jacob et al., 2010). The ARCPAC (Aerosol, Radiation, and Cloud Processes affecting Arctic Climate; Brock et al., 2011) campaign was conducted from Alaska together with ARCTAS in April 2008. The PAMARCMiP (Polar Airborne Measurements and Arctic Regional Climate Model Simulation Project) campaign covered the entire western Arctic in April 2009 (Stone et al., 2010). Two HIPPO (High-Performance Instrumented Airborne Platform for Environmental Research Pole-to-Pole Observations; Schwarz et al., 2010, 2013; Wofsy et al., 2011) campaigns during January 2009 and October 2009 explored the North American Arctic. Flight legs north of 70° N for all of these campaigns are shown in Fig. 1. Refractory BC (rBC) was measured during these campaigns with single particle soot photometer (SP2) instruments (Kondo et al., 2001; Schwarz et al., 2006). Observations of submicrometer aerosol sulfate mass during ARCTAS were made with the a particle-into liquid-Sampler (PILS) (Sullivan et al., 2006) coupled to an ion chromatograh. Sulfate measurements during ARCPAC were made with a compact time-of-flight aerosol mass spectrometer (Bahreini et al., 2008).

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

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

2.2 Emissions

All models made use of an identical emission dataset, the ECLIPSE (Evaluating the Climate and Air Quality Impacts of Short-Lived Pollutants) emission inventory version V4a (Klimont et al., 2015a, b). The ECLIPSE inventory was created using the GAINS (Greenhouse gas – Air pollution Interactions and Synergies) model (Amann et al., 2011), which provides emissions of long-lived greenhouse gases and shorter-lived species in a consistent framework. The proxies used in GAINS are consistent with those applied within the RCP (Representative Concentration Pathways) projections as described in Lamarque et al. (2010) and as further developed within the Global Energy Assessment project (GEA, 2012). They were, however, modified to accommodate more recent information where available, e.g., on population distribution and open biomass burning, effectively making them year specific (Riahi et al., 2012; Klimont et al., 2013).

For this study, emissions were provided for the years 2008 and 2009 and were lumped into the following source categories: industrial combustion, residential combustion, energy production, transport, agriculture, waste treatment, shipping, agricultural waste burning and gas flaring. All emission data were gridded consistently to a resolution...
of 0.5° × 0.5°. Monthly disaggregation factors were provided for the domestic heating emissions, based on ambient air temperatures. For a more detailed description of the ECLIPSE emission data set, see Klimont et al. (2015a, b). Non-agricultural biomass burning emissions were not available through GAINS and were therefore taken from the Global Fire Emission Database (GFED), version 3.1 (van der Werf et al., 2010). No attempt was made to harmonize sulfur emissions from volcanic sources or the ocean, which could explain some differences in simulated sulfate concentrations.

2.3 Models

We show results of 11 different models, whose main characteristics and references are summarized in Table 1. The horizontal resolution of the individual models ranges from about 0.6° × 0.8° to 2.8° × 2.8°. We use one Lagrangian particle transport model, FLEXPART (Flexible Particle Dispersion Model), which is run in backward mode for 30 days (thus, older source contributions are not accounted for). The simulation is driven by 1° × 1° operational analyses from the European Centre for Medium Range Weather Forecast (ECMWF). The OsloCTM2, TM4-ECPL (Tracer Model version 4 – Environmental Chemical Processes Laboratory) and SMHI MATCH (Swedish Meteorological and Hydrological Institute Multi-scale Atmospheric Transport and Chemistry Model) are CTMs and also use meteorological data from ECMWF (for details see Table 1). The DEHM (Danish Eulerian Hemispheric Model) CTM is driven by NCEP (National Centers for Environmental Prediction) meteorological data. WRF-Chem (Weather Research and Forecasting Model coupled with Chemistry) is an on-line atmospheric chemistry–weather forecast model which was nudged to NCEP FNL (final analysis) data for this study. The aerosol–climate model (ACM) ECHAM6-HAM2 (for brevity, referred to as ECHAM6 in figures) is the European Centre for Medium-Range Weather Forecasts Hamburg model version 6 (Stevens et al., 2013) extended with the Hamburg aerosol module version 2 (HAM2) (Zhang et al., 2012). ECHAM6-HAM2 and the CCMs HadGEM3 (Met Office Hadley Centre Climate Model, version 3) and CanAM4.2 (Canadian Atmospheric model, version 4.2) were nudged to ECMWF data. CESP1-
CAM5.2 (Community Earth System Model version 1 – Community Atmosphere model version 5.2) and NorESM1-M (Norwegian Earth System Model version 1 with intermediate resolution and used here in a version where aerosols are fully coupled with a tropospheric gas-phase chemistry scheme, hereafter referred to as NorESM) are also CCMs but were running freely, thus producing their own meteorological data. These latter models cannot be compared point-to-point with the measurement data because they produced meteorological conditions which were different from the actual ones; however, longer-term (e.g., seasonal) medians should still be comparable with the measurements, especially since sea surface temperatures (SST) and sea-ice extent were prescribed and specific to the years 2008–2009. All models were sampled exactly at the locations of the measurement stations and along the flight tracks at the highest possible (mostly hourly) temporal resolution. Notice that not all models simulated the full 2008–2009 period and FLEXPART only simulated BC.

3 Simulated BC and sulfate concentrations

Figure 2 shows the simulated BC and sulfate column mass loadings as a function of latitude for the time periods of the Arctic Haze (March) and the much cleaner summer (July) in the Arctic, for the models for which this information was available. For BC in March, most models show a maximum near 20° N, with some models extending this maximum to 40° N. This approximately covers the latitude range with the highest global emissions where the models agree fairly well in their simulated column loadings. In contrast, large differences between the models are found in the Arctic, where column mass loadings vary by more than an order of magnitude. Similar results are also found for sulfate in March, for which most models also show a maximum around 20–40° N; however, compared to BC, the models show a less pronounced decrease towards higher latitudes and two models even simulate increasing sulfate burdens with latitude. The relatively good agreement between the models in the BC and sulfate source region latitudes is not surprising, given that they all use the same emission data set. In contrast,
the differences between the atmospheric column loadings in the Arctic must mainly be
due to differences in the aerosol processing and removal and hence aerosol lifetimes,
and probably differences in atmospheric transport. Most models with relatively low BC
column loadings in the Arctic also have low sulfate loadings there, indicating similarities
in the simulated removal of these two types of aerosols. A notable exception, however,
is HadGEM3, which has moderately low BC but the highest sulfate loadings in the
Arctic.

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

In the simulated surface BC and sulfate mass mixing ratios the same basic patterns
are found as in the column loadings, but with enhanced gradients between source ar-
eas and remote regions (Fig. 3). When looking at individual models, there are, however,
notable differences for sulfate. ECHAM6-HAM2 has the highest sulfate surface mass
mixing ratios of all models, especially in the Northern Hemisphere subtropics and mid-
latitudes. Combined with the rather “normal” column sulfate loadings of this model,
this indicates that ECHAM6-HAM2 does not transport sulfate away from the surface as
quickly as the other models. On the other hand, HadGEM3, which has by far the largest
sulfate column loadings, has the smallest surface concentrations. This deficiency was
due to the implementation of the Global Model of Aerosol Processes (GLOMAP; Mann et al., 2010), which in this HadGEM3 version resulted in too little removal of the sulfate precursor SO$_2$ during the venting from the boundary layer to the free troposphere. The longer sulfate lifetime there explains the high column loadings.

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

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

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

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

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

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

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

For Tiksi, data the comparison is less direct as measurement data from July 2009–June 2010 were used. Nevertheless, it is clear that except for CanAM4.2 (which produces the highest modeled values at most sites) the models strongly underpredict for this site, especially in winter/spring. The most likely explanation for this is that the BC emissions in high-latitude Russia are underestimated in the ECLIPSE inventory. It is difficult to know where exactly the missing sources are located. However, we find that
in the ECLIPSE inventory the BC emissions in Norilsk (88.2° E, 69.3° N; population 170 000) are zero. We do not suggest that Norilsk emissions are responsible for the strong underestimation of BC concentrations at Tiksi, but these discrepancies (and others for sulfur emissions discussed later) suggest that the high-latitude Russian pollutant emissions are underestimated and/or wrongly placed in the ECLIPSE inventory (and in most other global emission inventories).

The seasonal cycle of sulfate at the monitoring stations is similar to that of eBC, with a clear maximum during the Arctic Haze season and a minimum in summer/early fall (Fig. 6). However, the seasonal cycle at the northernmost stations is less strong than for eBC, with about a factor of 5 difference between spring and summer, compared to a factor of 15 for eBC (Table 2). This is probably due to the influence of biogenic sources of sulfate in summer (Quinn et al., 2002) and/or a weaker seasonality in the emissions (e.g., smelter emissions of SO$_2$ are probably relatively constant throughout the year).

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

The models generally underpredict sulfate most strongly at the northernmost station (Alert), which is consistent with the BC results (compare Figs. 5 and 6). The CanAM4.2
model, which had some of the highest BC concentrations, also gives the highest sulfate values (Table 2). It is the only model that matches the high measured sulfate values at Alert and Station Nord in spring.

At Pallas, the lowest-latitude station in this comparison, most models severely underestimate sulfate throughout the year (Fig. 6), although they tend to overestimate BC in spring there. One likely reason for the sulfate underestimation is the proximity of the Pallas station to the Kola peninsula, where metal smelters are a strong source of sulfur. According to AMAP (2006), SO$_2$ emissions in Nikel, Zаполярный and Monchegorsk together were about 170 ktyr$^{-1}$ in the year 2002. In the ECLIPSE version 4a inventory used for this study the SO$_2$ emissions in these areas are only about 33 ktyr$^{-1}$ in total for the year 2005. Similar deficiencies were in fact reported also for other emission inventories for this region (Prank et al., 2010). Strong underestimation of the SO$_2$ emissions from metal smelting in the Kola peninsula is therefore a likely explanation for why almost all models underestimate sulfate at Pallas so strongly. Similar discrepancies were in fact found for SO$_2$ emissions in Norilsk, prompting a regridding of the ECLIPSE emissions (now available version 5a) using better location information for the metal smelting industry.

5 Vertical profiles

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

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

When plotted as a function of altitude (two bottom panel rows in Fig. 7), the measured values peak in the 4–5 km altitude bin, both for sub-Arctic and Arctic latitudes. The models, except for CanAM4.2, underestimate the measured median values throughout the entire depth of the profile. Some of the models, mainly those driven by observed meteorology, capture the rBC maximum in the mid-troposphere in the Arctic. However, the lower-latitude 4–5 km maximum is hardly reproduced by any of the models. One likely reason for the modeling problems is the strong biomass burning activity during spring 2008, which influenced a substantial fraction of the measurement data (Warneke et al., 2010; Brock et al., 2011). Even though this should be reflected in the GFED emission data for 2008, it seems possible that the GFED emissions are underestimated. Furthermore, as some of the flights targeted biomass burning plumes specifically, the influence of the biomass burning may be enhanced in the measurement data compared to the models, especially if the models did not capture the plume transport well enough and thus potentially simulated the biomass burning plumes at other locations.
than observed. This sampling bias is particularly strong for the CCMs which are not driven by observed meteorological fields.

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

The concentrations measured by the ARCTAS summer campaign in 2008 are much lower than those measured in spring 2008 and 2009, both at low and high altitudes (Fig. 8g and h), which is in agreement with the seasonality seen at the surface stations. Some of the models under- and others overestimate the measured concentrations, with the majority of the models overestimating, especially below 3 km. The mean values, averaged over all models, are about two (three) times as high as the measurements for altitudes above (below) 3 km. Some of the models reproduce the measured rBC maximum at 6 km (Fig. 9).
The HIPPO campaign in fall 2009 (Fig. 8i–j) was conducted about one month after the seasonal minimum at most surface sites and measured very low rBC mass concentrations, which is consistent with the surface observations. Most of the models overestimate the measured concentrations throughout the entire vertical profile (Fig. 9).

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

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

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

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

6 Station vs. low-altitude aircraft measurements

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

To investigate how consistent a picture the aircraft campaigns give vis a vis the station measurements, we compare all aircraft data from the lowest 3 km and lowest 1 km to the values obtained from the surface stations for the same months (Fig. 11). Selecting data only for even lower altitudes is problematic as the data coverage becomes
very poor. In Fig. 11, we also show the station measurements obtained for the years 2008 and 2009 separately. For eBC, the measurements obtained for the same month at the different stations and during different years are (with a few exceptions such as Barrow in January 2008) quite comparable with each other. In particular, April 2008 did not show higher eBC values than April 2009. This is consistent with the finding that the biomass burning layers in 2008 did not extend to the surface (Brock et al., 2011).

At Alert, the EC values are similar to the eBC values, whereas the Station Nord EC values are in summer and fall higher than eBC values at other stations. The aircraft rBC measurements for all campaigns show consistently lower values than the eBC/EC measurements at the ground, except for the HIPPO campaign in January 2009 where, however, the data coverage particularly below 1 km is poor. It is possible that the BC concentrations show a strong gradient in the lowest 1 km and that surface concentrations are indeed systematically higher than concentrations just aloft. However, an alternative explanation could be that the rBC measurements are biased low against the eBC/EC measurements, given the different measurement techniques used. A direct comparison of all three measurement techniques at the Alert station also suggests a low bias of rBC against eBC and EC concentrations (S. Sharma, personal communication, 2015). For sulfate (Fig. 12) the measurements show a much larger variability than for BC, both between stations and between the two different years. For instance, the 25th percentile of the sulfate concentrations at Alert in January 2009 is higher than the 75th percentile of the other stations and also of Alert in January 2008. On the other hand, the sulfate concentrations measured during the two available flight campaigns in spring and summer 2008 are not systematically different from those measured at the stations, although the median concentration in summer 2008 is somewhat lower than at the stations. This is consistent with the eBC/rBC differences.
7 Sulfate/BC correlations

Figure 13 shows correlation plots between monthly mean sulfate and eBC for the measurements and the models sampled at the different stations. In the observations, sulfate and eBC correlations for Alert, Pallas and Zeppelin are statistically significant at the 99.9% level (Table 3). The slopes of the regression lines shown in Fig. 13 are reported in Table 3. For the observations, they are very similar: 10.1, 8.4 and 8.9 ng[SO$_4$]$_3$−(ng[eBC]m$^{-3}$)$^{-1}$ for Alert, Pallas and Zeppelin, respectively. For Barrow, where the correlation is not significant because of two eBC-rich outlier data points, the slope is smaller (6.4 ng[SO$_4$]$_3$−(ng[eBC]m$^{-3}$)$^{-1}$). The strong correlation between sulfate and eBC and the similarity of the slopes suggests that the sources contributing to the measurements at the different stations are similar and that the removal of sulfate and eBC is highly correlated, which would be expected for internally mixed aged aerosol as is typical for the Arctic.

Most of the models, on the other hand, show much weaker correlation between sulfate and BC and some of the models have no significant correlation at all. Exceptions are DEHM, CESM1-CAM5.2 and WRF-Chem which show mainly significant correlations, and slopes that are comparable at the different stations and which are also quite similar to the observed slopes. This suggests that, with the given emissions, it is possible to reproduce the observed correlations. The lack of correlation between sulfate and BC in the other models – in disagreement with the observations – therefore suggests that they treat the two species differently, probably having a too large fraction of the aerosol as externally mixed. Correlations could also be degraded by a too strong influence of biogenic (dimethyl sulfide) emissions from the oceans or factors influencing SO$_2$ to sulfate conversion such as the level of oxidants in the models. This could lead to varying fractions of sulfur present as SO$_2$ and maybe these fractions are more variable in the models than in reality.

The SO$_2$ (converted to sulfate) to BC emission ratio of anthropogenic emissions in the ECLIPSE inventory is 25 globally and 40 north of 50°N. For the GFED
biomass burning emissions the emission ratio is only 1.7 globally and 2.5 north of
50°N, and for the sum of anthropogenic and biomass burning emissions, we ob-
tain ratios of 19 globally and 25 north of 50°N. The mean observed slopes of
the observations (9.1 ng[SO$_4$]/m$^3$/ng[eBC]/m$^3$) and the slopes modeled by DEHM
(5.4 ng[SO$_4$]/m$^3$/ng[BC]/m$^3$), CESM1-CAM5.2 (9.9 ng[SO$_4$]/m$^3$/ng[BC]/m$^3$) and WRF-
Chem (8.5 ng[SO$_4$]/m$^3$/ng[BC]/m$^3$) are much lower than the emission ratio of anthro-
pogenic emissions in the ECLIPSE inventory and they are also lower than the emission
ratio for mixed anthropogenic and biomass burning emissions. This suggests that
biomass burning emissions are relatively more important in the Arctic than elsewhere,
that there are missing BC sources, that sulfur emissions are overestimated (although
this is not so likely, given the too low SO$_2$ emissions in high latitude Russia in the
ECLIPSE version 4a inventory used here), and/or that there exists a mechanism that
enriches aerosols in BC relative to sulfate in the Arctic atmosphere. The latter could be
related to the hydrophobic nature of freshly emitted BC.

8 Conclusions

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

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

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

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

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

**Acknowledgements.** The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 282688 – ECLIPSE. Some of the work was conducted for and funded by the Arctic Monitoring and Assessment Programme (AMAP). French authors also acknowledge support from the CLIMSLIP-ANR project and computer resources provided by IDRIS HPC resources under the
allocation 2014-017141 under GENCI. Contributions by SMHI were funded by the Swedish Environmental Protection Agency under contract NV-09414-12 and through the Swedish Climate and Clean Air research program, SCAC. Simulations with CanAM4.2 were supported by the Network on Climate and Aerosols: Addressing Key Uncertainties in Remote Canadian Environments (NETCARE), with partial funding from the Natural Sciences and Engineering Research Council of Canada (NSERC). This is PMEL contribution number 4276. ECMWF gave access to their meteorological data. Environment Canada (S. Sharma) provided the sulfate data and eBC data. S.-M. 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. J. Schmale is acknowledged for valuable discussion.

References


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


<table>
<thead>
<tr>
<th>Model Name</th>
<th>Model Type*</th>
<th>Horizontal/vertical resolution</th>
<th>Meteorological fields</th>
<th>Periods simulated/output temporal resolution</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>OsloCTM2</td>
<td>CTM</td>
<td>2.8° × 2.8°, 60L</td>
<td>ECMWF Operational Analyses</td>
<td>2008–2009, 3 h</td>
<td>Myhre et al. (2009), Skeie et al. (2011a, b)</td>
</tr>
<tr>
<td>NorESM</td>
<td>CCM</td>
<td>1.9° × 2.5°, 26L</td>
<td>Internal, observed SST prescribed ECMWF</td>
<td>2008–2009, 3 h, 24 h</td>
<td>Kirkevåg et al. (2013), Bentsen et al. (2013), Myroikofitalakis et al. (2011), Kanakidou et al. (2012), Daskalakis et al. (2014)</td>
</tr>
<tr>
<td>ECHAM6-HAM2</td>
<td>ACM</td>
<td>1.8° × 1.8°, 31L</td>
<td>ECMWF Reanalysis</td>
<td>March–August, 2008, 1 h</td>
<td>Liu et al. (2012), Wang et al. (2013), Grell et al. (2005), Zaveri et al. (1999, 2008), Hewitt et al. (2011), Mann et al. (2010)</td>
</tr>
<tr>
<td>SMHI-MATCH</td>
<td>CTM</td>
<td>0.57° × 0.75°, 38L</td>
<td>ECMWF – ERA-interim</td>
<td>2008 (2009), 1 h</td>
<td>Zaveri et al. (1999, 2008)</td>
</tr>
<tr>
<td>CanAM4.2</td>
<td>ACM</td>
<td>2.8° × 2.8°, 49L</td>
<td>Nudged to ECMWF temp. and winds NCEP</td>
<td>2008–2009, 3 h</td>
<td>Liu et al. (2012), Wang et al. (2013), Grell et al. (2005), Zaveri et al. (1999, 2008), Hewitt et al. (2011), Mann et al. (2010)</td>
</tr>
<tr>
<td>DEHM</td>
<td>CTM</td>
<td>150 km &lt; 60°, 50 km &gt; 60° N, 29L</td>
<td>NCEP</td>
<td>2008–2009, 3 h</td>
<td>Liu et al. (2012), Wang et al. (2013), Grell et al. (2005), Zaveri et al. (1999, 2008), Hewitt et al. (2011), Mann et al. (2010)</td>
</tr>
<tr>
<td>CESM1/CAM5.2</td>
<td>CCM</td>
<td>1.9° × 2.5°, 30L</td>
<td>Internal, observed SST prescribed</td>
<td>2008–2009, 1 h</td>
<td>Liu et al. (2012), Wang et al. (2013), Grell et al. (2005), Zaveri et al. (1999, 2008), Hewitt et al. (2011), Mann et al. (2010)</td>
</tr>
<tr>
<td>WRF-Chem</td>
<td>CTM</td>
<td>100 km × 100 km 38L</td>
<td>Nudged to FNL</td>
<td>March–August 2008, 3 h</td>
<td>Liu et al. (2012), Wang et al. (2013), Grell et al. (2005), Zaveri et al. (1999, 2008), Hewitt et al. (2011), Mann et al. (2010)</td>
</tr>
</tbody>
</table>

* Chemistry transport model (CTM), Lagrangian particle dispersion model (LPDM), chemistry climate model (CCM), aerosol chemistry model (ACM)
Table 2. Median observed eBC and modeled BC mass concentrations in ng m\(^{-3}\) as well as measured and modeled sulfate (SO\(_4\)) concentrations in the Arctic during winter/spring (January to March) and summer (July to September). The data used are from the years 2008 and 2009 and were averaged for the three stations Alert, Barrow and Zeppelin. Notice that some models do not cover the whole periods completely (see Table 1).

<table>
<thead>
<tr>
<th>Model/obs</th>
<th>Winter/Spring BC [ng m(^{-3})]</th>
<th>Summer BC [ng m(^{-3})]</th>
<th>Winter/Spring SO(_4) [ng m(^{-3})]</th>
<th>Summer SO(_4) [ng m(^{-3})]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured</td>
<td>49.4</td>
<td>3.3</td>
<td>561.0</td>
<td>103.2</td>
</tr>
<tr>
<td>Model mean:</td>
<td>20.1</td>
<td>6.2</td>
<td>353.6</td>
<td>148.6</td>
</tr>
<tr>
<td>FLEXPART</td>
<td>40.2</td>
<td>7.7</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>OsloCTM2</td>
<td>8.4</td>
<td>1.3</td>
<td>90.2</td>
<td>109.7</td>
</tr>
<tr>
<td>NorESM</td>
<td>13.0</td>
<td>4.4</td>
<td>394.2</td>
<td>70.8</td>
</tr>
<tr>
<td>TM4-ECPL</td>
<td>5.4</td>
<td>1.3</td>
<td>71.3</td>
<td>149.7</td>
</tr>
<tr>
<td>ECHAM6-HAM2</td>
<td>1.9</td>
<td>2.1</td>
<td>488.7</td>
<td>388.9</td>
</tr>
<tr>
<td>SMHI-MATCH</td>
<td>38.6</td>
<td>1.1</td>
<td>603.3</td>
<td>151.1</td>
</tr>
<tr>
<td>CanAM4.2</td>
<td>38.8</td>
<td>1.6</td>
<td>791.3</td>
<td>270.9</td>
</tr>
<tr>
<td>DEHM</td>
<td>57.1</td>
<td>11.6</td>
<td>434.6</td>
<td>61.1</td>
</tr>
<tr>
<td>CESM1-CAM5</td>
<td>21.3</td>
<td>5.1</td>
<td>210.5</td>
<td>21.9</td>
</tr>
<tr>
<td>WRF-Chem</td>
<td>14.9</td>
<td>32.3</td>
<td>408.8</td>
<td>246.6</td>
</tr>
<tr>
<td>HadGEM3</td>
<td>1.8</td>
<td>0.7</td>
<td>43.2</td>
<td>15.9</td>
</tr>
</tbody>
</table>
Table 3. Slopes of regression lines between monthly mean concentrations of sulfate and (e)BC for the different stations. Slopes are calculated both for the observations and the model values. Values that are statistically significant at the 99.9 % level are marked with an asterisk. For the mean over all sites/models, only the statistically significant values were averaged.

<table>
<thead>
<tr>
<th></th>
<th>Alert</th>
<th>Barrow</th>
<th>Pallas</th>
<th>Zeppelin</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>10.1*</td>
<td>6.4</td>
<td>8.4*</td>
<td>8.9*</td>
<td>9.1</td>
</tr>
<tr>
<td>Model mean</td>
<td>17.3</td>
<td>16.6</td>
<td>6.7</td>
<td>9.7</td>
<td>12.6</td>
</tr>
<tr>
<td>OsloCTM2</td>
<td>−8.6</td>
<td>2.4</td>
<td>−2.0</td>
<td>−5.5</td>
<td>−</td>
</tr>
<tr>
<td>NorESM</td>
<td>35.3*</td>
<td>27.8</td>
<td>0.4</td>
<td>12.1</td>
<td>35.3</td>
</tr>
<tr>
<td>TM4-ECPL</td>
<td>9.5</td>
<td>33.2*</td>
<td>5.8*</td>
<td>8.1</td>
<td>19.5</td>
</tr>
<tr>
<td>ECHAM6-HAM2</td>
<td>30.0</td>
<td>90.4</td>
<td>1.0</td>
<td>−746.4</td>
<td>−</td>
</tr>
<tr>
<td>SMHI-MATCH</td>
<td>25.6*</td>
<td>25.9*</td>
<td>0.4</td>
<td>10.9</td>
<td>25.7</td>
</tr>
<tr>
<td>CanAM4.2</td>
<td>18.2*</td>
<td>2.5</td>
<td>7.1</td>
<td>12.4*</td>
<td>15.3</td>
</tr>
<tr>
<td>DEHM</td>
<td>7.5*</td>
<td>5.7*</td>
<td>1.6*</td>
<td>6.7*</td>
<td>5.4</td>
</tr>
<tr>
<td>CESM1-CAM5.2</td>
<td>11.1*</td>
<td>8.9*</td>
<td>9.6*</td>
<td>9.9*</td>
<td>9.9</td>
</tr>
<tr>
<td>WRF-Chem</td>
<td>6.4*</td>
<td>9.3*</td>
<td>9.8*</td>
<td>2.4</td>
<td>8.5</td>
</tr>
<tr>
<td>HadGEM3</td>
<td>10.7</td>
<td>−8.7</td>
<td>−0.81</td>
<td>3.2</td>
<td>−</td>
</tr>
</tbody>
</table>
Figure 1. Map showing the locations of the measurement stations (yellow circles) and the flight tracks north of 70° N of all aircraft campaigns used in this study. Aircraft data were from the HIPPO (winter 2009 and fall 2009), ARCTAS (spring and summer 2008), ARCPAC (spring 2008) and PAMARCMiP (spring 2009) campaigns.
Figure 2. BC (a, c) and sulfate (b, d) column mass loadings for the year 2008 averaged over all longitudes as a function of latitude (for the range 50° S to 90° N) for March (a, b) and July (c, d).
Figure 3. BC (a, b, e, f) and sulfate (c, d, g, h) mass mixing ratios for the year 2008 at the surface averaged over all longitudes as a function of latitude (for the range 50° S to 90° N) for March (a–d) and July (e–h). The right panels show the same data as the left panels, but only for 70–90° N and with an adjusted ordinate scale.
Figure 4. Observed and simulated mean annual cycle of (equivalent) BC mass concentrations [ng m\(^{-3}\)] at the Zeppelin station, shown as monthly frequency distributions using data from the years 2008 and 2009. The uppermost panel shows monthly frequency distributions of the observed eBC concentrations, the other panels show monthly frequency distributions of the modeled BC concentrations. Black dots depict the monthly median value, the grey boxes span the range between the 25th and 75th percentiles, red and grey dots represent values which are outside the 1.5 fold of this interquartile range. The red line connects the monthly medians of the observed eBC concentrations in the uppermost panel and is repeated in all other panels for the convenience of comparing modeled and measured values. Missing model data are denoted with “X”. Notice that some models have very low BC mass concentrations, which are difficult to see on the scale used.
Figure 5. Monthly (month is displayed on the abscissa) median observed eBC/EC and modeled BC concentrations for the stations (from top) Alert, Nord, Zeppelin, Tiksi, Barrow and Pallas, for late winter/spring (left column) and summer/fall (right column). The red dashed lines connect the observed median eBC values, the light red shaded areas span from the 25th to the 75th percentile of the observations. The black dots are the EC concentrations, which are available for Alert and Station Nord. Modeled median values are shown with different lines according to the legend. Notice the difference in concentration scales used for the left and right panels and also for the Tiksi station.
Figure 6. Monthly (month is displayed on the abscissa) median observed and modeled sulfate concentrations for the stations (from top) Alert, Nord, Zeppelin, Barrow and Pallas. The red dashed lines connect the observed median values, the light red shaded areas span from the 25th to the 75th percentile of the observations. Modeled median values are shown with different lines according to the legend.
Figure 7. Comparison of modeled BC with observed rBC mass concentrations from the ARCTAS-spring and ARCPAC campaigns in spring 2008. The leftmost column shows box and whisker plots of observed rBC concentrations in ng m$^{-3}$. The black dots as well as the red lines represent the median values. The other columns show the modeled BC concentrations for FLEXPART, OsloCTM2, NorESM, TM4-ECPL, ECHAM6-HAM2, SMHI-MATCH, CanAM4.2, DEHM, CESM1-CAM5.2, WRF-Chem and HadGEM3. The top (second from top) row represents median (r)BC concentrations for altitudes below (above) 3 km a.s.l. as a function of latitude by binning the data into 10° latitude bands. 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.
Figure 8. Median observed rBC and modeled BC mass concentrations for the winter 2009 HIPPO (a, b) spring 2009 PAMARCMiP (c, d) spring 2008 ARCTAS/ARCPAC (e, f) summer 2008 ARCTAS (g, h) and the fall 2009 HIPPO (i, j) aircraft campaigns. The red bar and the red horizontal line show the observations, the other colored bars the various models, the grey line shows the mean value of all model medians. Results are shown separately for measurements below 3 km (left panels) and above 3 km (right panels). Notice that the concentration scales on the ordinates are different for the individual panels.
Figure 9. Comparison of modeled BC with observed rBC mass concentrations as a function of altitude for all data taken north of 70° N, for the different campaigns (same as in Fig. 8). The leftmost column shows box and whisker plots of observed rBC concentrations in ng m$^{-3}$. The black dots as well as the red lines represent the median values. The other columns show the modeled BC concentrations for FLEXPART, OsloCTM2, NorESM, TM4-ECPL, ECHAM6-HAM2, SMHI-MATCH, CanAM4.2, DEHM, CESM1-CAM5.2, WRF-Chem and HadGEM3.
Figure 10. Median SO$_4$ concentrations for the ARCTAS/ARCPAC spring 2008 (a, b) and ARCTAS summer 2008 (c, d) campaigns. The red bar and the red horizontal line show the observations, the other colored bars the various models. The analysis is performed for measurements below 3 km (left panels) and above 3 km (right panels). Note: each row has a different y-axis.
Figure 11. Comparison of eBC [ng m\(^{-3}\)] measured at the stations Zeppelin (Zep), Alert (Alt), and Barrow (Brw) (grey bars), EC measured at Alert and Station Nord (Nord) (green dots and bars) and rBC [ng m\(^{-3}\)] measured by aircraft (Air) in the lowest 3 and 1 km, north of 70° N (blue bars) for the years 2008 and 2009 for (a) January, (b) April, (c) June and July and (d) October and November. The black dots represent the median, and the boxes the interquartile range. For the aircraft measurements, the blue boxes show the results for the lowest 3 km, the black box outlines show the results for the lowest 1 km.
Figure 12. Same as Fig. 9, but for sulfate.
Figure 13. Correlation plots of monthly mean sulfate and (e)BC concentrations for the observations (top left) and the different models sampled at the observation sites.