**Atmospheric
Chemistry
and Physics**

Using measurements for evaluation of black carbon modeling

S. Gilardoni, E. Vignati, and J. Wilson

European Commission, Institute for Environment and Sustainability, Climate Change Unit, Ispra, Italy

Abstract. The ever increasing use of air quality and climate model assessments to underpin economic, public health, and environmental policy decisions makes effective model evaluation critical. This paper discusses the properties of black carbon and light attenuation and absorption observations that are the key to a reliable evaluation of black carbon model and compares parametric and nonparametric statistical tools for the quantification of the agreement between models and observations. Black carbon concentrations are simulated with TM5/M7 global model from July 2002 to June 2003 at four remote sites (Alert, Jungfraujoch, Mace Head, and Trinidad Head) and two regional background sites (Bondville and Ispra). Equivalent black carbon (EBC) concentrations are calculated using light attenuation measurements from January 2000 to December 2005. Seasonal trends in the measurements are determined by fitting sinusoidal functions and the representativeness of the period simulated by the model is verified based on the scatter of the experimental values relative to the fit curves. When the resolution of the model grid is larger than $1° \times 1°$, it is recommended to verify that the measurement site is representative of the grid cell. For this purpose, equivalent black carbon measurements at Alert, Bondville and Trinidad Head are compared to light absorption and elemental carbon measurements performed at different sites inside the same model grid cells. Comparison of these equivalent black carbon and elemental carbon measurements indicates that uncertainties in black carbon optical properties can compromise the comparison between model and observations. During model evaluation it is important to examine the extent to which a model is able to simulate the variability in the observations over different integration periods as this will help to identify the most appropriate timescales. The agreement between model and obser-

vation is accurately described by the overlap of probability distribution (PD) curves. Simple monthly median comparisons, the Student's t-test, and the Mann-Whitney test are discussed as alternative statistical tools to evaluate the model performance. The agreement measured by the Student's t-test, when applied to the logarithm of EBC concentrations, overestimates the higher PD agreements and underestimates the lower PD agreements; the Mann-Whitney test can be employed to evaluate model performance on a relative scale when the shape of model and experimental distributions are similar.

1 Introduction

The term black carbon (BC) refers to light absorbing carbon (LAC) aerosols produced by combustion processes (Bond and Bergstrom, 2006; Cachier, 1998), also known as soot. Its ability to absorb visible light is due to the molecular structure, which is characterized by valence electrons in π-orbitals; the energy gaps between bonding and anti-bonding molecular orbitals is small enough to allow the absorption of visible light (380–880 nm). BC concentration is determined from light attenuation measurements at a constant wavelength or at multiple wavelengths. A detailed list of direct and indirect LAC measurement techniques is reported by Bond and Bergstrom (2006). The measurements discussed here are direct measurements of the light attenuation produced by aerosol either suspended in the air or deposited on a filter.

BC modeling is a crucial component of global climate model because BC is the most ubiquitous aerosol component which absorbs visible light and is a key contributor to the climate forcing by aerosol (Liousse et al., 1993; Cooke et al., 1997; Andreae and Merlet, 2001; Menon et al., 2002; Hansen et al., 2005; Ramanathan and Carmichael, 2008). BC aerosols can be transported far from their sources and affect

remote and pristine areas, as well as rural and urban locations (Hansen et al., 1988; Hara et al., 2008; Eleftheriadis et al., 2009). The first attempt to model BC concentration used a simple approach assuming that BC was externally mixed and had a constant size distribution (Haywood and Shine, 1995; Tegen et al., 1997). More recent models included BC size fractionation and internal mixing with soluble aerosol components (Jacobson, 2001; Wilson et al., 2001; Stier et al., 2005).

Model evaluation is crucial because it helps understanding pollutant dynamics and makes it possible to use model output to quantify the effect of environmental policies on regional and global scale (Koch et al., 2009). Several approaches have been proposed to obtain accurate assessments of local and regional air quality model performance (Chang and Hanna, 2004; Seigneur et al., 2000; Boylan and Russell, 2006; Yu et al., 2006; Zhang et al., 2006); on the contrary, evaluation of large scale models (like global models) is commonly performed in a simplistic manner by comparing monthly means (or medians) of model data and observations.

This study is part of the EUCAARI – European Integrated Project on Aerosol Cloud Climate and Air Quality Interactions (Kulmala et al., 2009) and AEROCOM – Aerosol Comparisons between Observations and Models (http://nansen.ipsl.jussieu.fr/AEROCOM) (Textor et al., 2006) efforts to improve model validation methods. The objectives of this paper are:

– the optimization of a validation procedure that includes evaluation of observation representativeness,

– the discussion of statistical tools to compare model and observations.

First, we recognize that the model/observation comparison is only meaningful when the temporal and spatial representativeness of observations are tested. In addition, light absorbing interfering species and uncertainties in the BC optical properties might strongly affect the interpretation of the observations. Equivalent black carbon measurements performed at remote and regional background sites are then used to validate TM5/M7 model; mathematical and statistical tools to evaluate model/experiment agreement and to highlight reasons of model failure are suggested and discussed.

2 Observation and model data

In the following sections we describe the measurement sites used for the model/observation comparison, the physical meaning of equivalent black carbon based on light attenuation measurements, and the model details. The choice of the measurement sites was based on the availability of long time series of observations. Unfortunately, the number of locations that offer public data access and long time coverage is limited and this study could not be extended to more than six sites.

2.1 Monitoring sites

Light attenuation measurements from the WMO Global Atmosphere Watch (GAW) sites at Alert (Canada), Bondville (Illinois), Ispra (Italy), Jungfraujoch (Switzerland), Mace Head (Ireland), and Trinidad Head (California) have been used in this study. Ispra, Jungfraujoch, and Mace Head are also part of the European Monitoring and Evaluation Programme (EMEP) network. Site coordinates, periods during which measurements are available (till 2005), and the details of the measurements made are summarised in Table 1. This set of sites is chosen as representative of both remote sites and rural background sites. In addition it gives us the opportunity to illustrate diverse issues that make black carbon modeling challenging. Table 2 lists sources of light absorbing carbonaceous aerosol that affect the sites with their corresponding seasonal cycles.

Alert is a continental remote site operated by Environment Canada, located on the edge of Lincoln Sea and Ellesmere Island, in the Canadian Arctic. The air masses circulation is driven by a semi-permanent pressure system able to transport continental airflow from Europe and Asia to the Arctic. Light absorbing carbon observed in the Arctic region is mainly produced by anthropogenic activities and biomass burning emissions transported from northern Europe, Siberia and southern Asia (Sharma et al., 2004, 2006; Koch and Hansen, 2005); open fires in the boreal forest of North America contribute to the black carbon levels during the summer season (Stohl et al., 2006).

Bondville (IL) is a rural monitoring site and GAW regional station that is part of the NOAA/ESRL and IMPROVE networks, located 8 km south of Champaign (population 75 000). The prevailing sources of light absorbing carbon are emissions from local traffic, coal-fired power plants located in Tennessee Valley, Ohio River Valley and western Ohio, Canadian forest fires, local crop harvesting and agricultural field burning (Sheridan et al., 2000; Kim et al., 2005; Malm et al., 2007).

Ispra site is located in northern Italy, 50 km north of Milan (population 1 300 000); although Ispra is classified as a rural background site, concentration of carbonaceous aerosol is comparable to those of urban European background sites, likely due to the impact of regional air pollution from the nearby Po Valley (Yttri et al., 2007). The main sources of light absorbing carbonaceous aerosols are central Europe continental pollution, local traffic and residential heating (Lazaridis et al., 2002; Putaud et al., 2002).

Mace Head Atmospheric Research Station is situated on the west coast of Ireland, and is one of the most westerly locations of Europe; the closest urban area is Galway City at 88 km to the east (population 70 000). Air masses that reach Mace Head are most of the time from the west i.e. clean air masses from the North Atlantic Ocean; less frequently, the site is affected by anthropogenic air pollution from continental Europe and United Kingdom (Jennings

Table 1. Site description, model and experimental details.

Site	ALT	BND	IT04	JFJ	MHT	THD
Location	Alert	Bondville	Ispra	Jungfraujoch	Mace Head	Trinidad Head
Lat (°)	82.5	40.1	45.8	46.6	53.3	41.1
Long (°)	−62.5	−88.4	8.6	8.0	−9.9	−124.2
Elev (m)	210	213	209	3450	5	107
Period	2000–2005	2000–2005	2004–2005	2000–2005	2000–2002	2002–2005
Parameter	EBC	B_{ATN}	B_{ATN}	EBC	B_{ATN}	B_{ATN}
Instrument	AE–6	PSAP	AE–31	AE–31	AE–9	PSAP
Wavelength (nm)	broadband	550	660	370	broadband	550
C			2.14	2.15		
R			1	1		
$\sigma^*(m^2\,g^{-1})$	19/28	10	10	18	19	10
RH		<40%				<40%
$T\,(°C)$	Ambient		Ambient		25	
TM5 resolution	6° × 4°	6° × 4°	1° × 1°	1° × 1°	1° × 1°	6° × 4°
A_0 (ng m^{-3})	97 ± 1	134 ± 5	1920 ± 30	83 ± 1	57 ± 4	58 ± 5

Table 2. Sources and seasonality of light absorbing aerosol at the monitoring sites.

Site	Long-range transport	Vertical transport	Local traffic	Residential heating	Coal power plant	Forest fire	References
ALT	All year					Summer	Koch et al. (2005) Sharma et al. (2004) Sharma et al. (2006)
BND			All Year		All Year	Summer	Kim et al. (2005) Malm et al. (2007)
IT04	All year		All Year	Winter			Putaud et al. (2002) Yttri et al. (2007)
MHT	All year					Summer	Cooke et al. (1997) Derwent et al. (2001) Forster et al. (2001) Jennings et al. (1997) Kleefeld et al. (2002) Yttri et al. (2007)
JFJ	All year	Summer				Summer	Baltensperger et al. (1997) Nyeki et al. (1998) Real et al. (2007)
THD	All year					Summer	Goldstein et al. (2004) Ogren et al. (2002) Ramanathan et al. (2007)

et al., 1997; Derwent et al., 2001; Kleefeld et al., 2002; Yttri et al., 2007). Investigation of temporal trend of carbonaceous aerosol shows that removal mechanisms (i.e. precipitation) play a major role in modulating light absorbing carbon concentration (Cooke et al., 1997).

Jungfraujoch is a remote high elevation site located in the southern part of Switzerland. Local emissions are expected to be negligible, in fact the site can be reached only through an electrical railway, tourist facilities are heated electrically, and local wastes are transported to the valley (Baltensperger et al., 1997). Most of the time the site is representative of continental European background, while slope winds are responsible for planetary boundary layer influence during summer (Nyeki et al., 1998).

Trinidad Head is located on the northern coast of California, about 40 km north of Eureka (population 25 000). This remote maritime site with negligible anthropogenic influence and prevailing maritime airflow is also part of the NOAA/ESRL network. Trinidad Head is affected by North American continent pollution, Asian outflow including dust and anthropogenic emissions, and smoke from forest fires (Ogren et al., 2002; Goldstein et al., 2004; Perry et al., 2004; Ramanathan et al., 2007; Zhao et al., 2008).

2.1.1 Light attenuation and equivalent black carbon measurements

Determination of BC from visible light attenuation implies three assumptions that are seldom verified:

– carbonaceous species are the only absorbing species at the measurement wavelength,

– organic species do not contribute to light absorption, and

– the relationship between light attenuation and BC concentration is well known and predictable.

Iron oxides, like hematite, are common components of aerosols containing dust and absorb visible light at wavelengths shorter than 600 nm (Sokolik and Toon, 1999; Alfaro et al., 2004; Yang et al., 2009). Some organic species also absorb radiation at wavelength shorter than 600 nm (Kirchstetter et al., 2004; Alexander et al., 2008). Light absorbing organic carbon, also known as brown carbon, biases BC measurements especially in regions strongly affected by biomass burning where brown carbon concentrations are higher than BC concentrations (Andreae and Gelencser, 2006; Barnard et al., 2008). The conversion of light attenuation into BC requires the knowledge of the attenuation cross section σ, which defines the ability of a species to absorb radiation at a certain wavelength. σ can be determined theoretically based on the Mie theory (Liousse et al., 1993). For ambient aerosol σ is more often determined experimentally using the Lambert-Beer equation and assuming that the BC concentration is equal or proportional to the elemental carbon (EC) concentration measured by thermo-optical methods. σ values reported in literature for combustion aerosols range over one order of magnitude (2–25 m^2 g^{-1}), while values reported for BC ambient aerosol at 550 nm are between 7 and 12 m^2 g^{-1}(Rosen and Hansen, 1984; Sharma et al., 2002; Bond and Bergstrom, 2006; Corrigan et al., 2006; Junker et al., 2006; Yang et al., 2009). σ depends on particle morphology, composition of internally mixed soot particles and soot aging (Liousse et al., 1993; Bond and Bergstrom, 2006), consequently it varies from location to location, and in the same location, from season to season. To take into account this dependency and variability we refer to the light absorbing carbon as equivalent black carbon (EBC) and not BC. The term BC implies to have optical properties equivalent to those of soot, while EBC is operationally defined as the amount of light absorbing carbon that would give the same signal of soot in an optical instrument like the aethalometer (Andreae and Gelencser, 2006). Caution must be exercised when EBC concentrations are calculated from light attenuation measurements assuming a known and constant σ.

Light attenuation measurements are made available through the World Data Centre for Aerosol (WDCA), as part of the GAW program, by the investigators responsible at each site (originally: http://wdca.jrc.ec.europa.eu, now http://www.gaw-wdca.org/). Light attenuation measurements are performed with an aethalometer (Hansen et al., 1984) at Alert, Ispra, Jungfraujoch, and Mace Head and with a Particle Soot Photometer PSAP (Bond et al., 1999) at Bondville and Trinidad Head; time resolution of observation data is 1 h. The aethalometer and PSAP are filter-based measurement techniques. They measure attenuation of light (B_{ATN}) due to particles accumulated on a filter, according to:

$$B_{ATN} = -100\ln(I/I_0) \tag{1}$$

I and I_0 are the intensities of light transmitted through a loaded filter and a blank, respectively. The presence of a collection substrate might affect light attenuation: some of the aerosol particles can be embedded into the filter fibrous structure and do not contribute to light attenuation; at the same time filter fibres might be able to enhance aerosol absorption due to the scattering of the light beam and consequently increase of the beam path.

During aethalometer measurements light absorption (B_{abs}) is calculated from light attenuation through the equation

$$B_{abs} = B_{ATN} \frac{1}{C \cdot R(ATN)} \tag{2}$$

where C and R are empirical calibration factors (dimensionless): C describes the multiple scattering of light produced by the filter fibers and depends on the filter properties, R depends on the amount of particles embedded in the filter structure and the particle optical properties (Weingartner et al., 2003). EBC concentration (in µg m^{-3}) can be calculated from light attenuation (B_{ATN} in Mm^{-1}) according to the equation reported by Weingartner et al. (2003) and Liousse et al. (1993)

$$B_{ATN} = EBC \cdot \sigma \tag{3}$$

where σ is the attenuation cross–section (in m^2 g^{-1}) and varies with the radiation wavelength λ. For particles small compared to the radiation wavelength, according to the Mie theory, it can be proved (Jennings and Pinnick, 1980) that the attenuation coefficient is inversely proportional to the radiation wavelength (λ); Jennings and Pinnick (1980) calculated that σ of spherical pure black carbon particles is equal to 14625/λ (nm). To calculate EBC from light absorption, Eqs. (2 and 3) can be written as:

$$B_{abs} = EBC \cdot \sigma^* \tag{4}$$

We define σ^* (in m^2 g^{-1}) the empirical calibration coefficient that relates light absorption and EBC, and for aethalometer it is equal to σ divided by C and R.

At Alert the relationship between light absorption and EBC concentration is determined experimentally through a long-term comparison of aethalometer and thermo-optical measurements performed from 1998 to 2001 (Sharma et al.,

2002, 2004); σ^* determined for the period November–April is $19\,\mathrm{m^2\,g^{-1}}$ and for May–October is $28\,\mathrm{m^2\,g^{-1}}$. At Jungfraujoch C is empirically determined (Lavanchy et al., 1999) and σ^* suggested by the data owner is $18\,\mathrm{m^2\,g^{-1}}$. No calibration information is available at Ispra and Mace Head. At Ispra we assume C equal to 2.14 (Bond and Bergstrom, 2006) and the resulting σ^* is $10\,\mathrm{m^2\,g^{-1}}$. At Mace Head σ^* is assumed equal to $19\,\mathrm{m^2\,g^{-1}}$ as suggested by the instrument manufacturer and by Junker et al. (2006) from the analysis of long-term measurements (1989–2003) at the same site.

Light attenuation measurements reported by PSAP already include an empirical calibration that takes into account absorption by the filter medium and nonlinear response due to filter overloading. They are corrected using the method described by Bond et al. (1999) for flow rate and spot size inaccuracy. σ^* values reported in literature for PSAP span over a fairly wide range (5 to $25\,\mathrm{m^2\,g^{-1}}$) depending on the optical properties of ambient aerosols and on the thermo-optical method used for calibration (Snyder and Schauer, 2007). The measurements are available at the two sites located in the United States. The comparison of EBC from light integrating plate measurements (LIPM) and EC from thermo-optical measurements at rural sites of the United States (Huffman, 1996; Malm et al., 1994) suggests that σ^* is equal to $20\,\mathrm{m^2\,g^{-1}}$; nevertheless Malm et al. (1994) used $10\,\mathrm{m^2\,g^{-1}}$ during their discussion of the IMPROVE network data, as suggested by theoretical calculations; in fact, the value $20\,\mathrm{m^2\,g^{-1}}$ would result from the misidentification of OC/EC split during the thermo-optical measurements and the consequent underestimation of EC. EBC concentrations presented here for Bondville and Trinidad Head are calculated assuming σ^* equal to $10\,\mathrm{m^2\,g^{-1}}$.

A summary of σ^* values used at each site is reported in Table 1, together with C and R correction factors when σ^* was calculated from the theoretical σ value of $14625/\lambda$.

At Bondville, Junfraujoch, and Trinidad Head EBC concentrations are available for PM_1 and PM_{10} size fractions, while at Alert, Ispra, and Mace Head EBC concentrations are available for total suspended particles (TSP).

2.1.2 Black carbon modeling

The global model TM5/M7 simulated BC concentrations at the monitoring sites from July 2002 to June 2003, with 1-hour time resolution.

The TM5 model is an off-line global transport chemistry model (Krol et al., 2005) that uses the ECMWF IFS (Integrated Forecast System) meteorological data. It has a spatial global resolution of $6° \times 4°$ and a two-way zooming algorithm that allows regions (e.g. Europe, North America, Africa and Asia) to be resolved at a finer resolution of $1° \times 1°$. To smooth the transition between the global $6° \times 4°$ region and the regional $1° \times 1°$ domain, a domain with a $3° \times 2°$ area resolution has been added. In the present application the zoom is over Europe, therefore outside the Eu-

ropean domain the resolution of the model is $6° \times 4°$. In the current version, the model has a vertical resolution of 25 layers, defined in a hybrid sigma-pressure coordinate system with a higher resolution in the boundary layer and around the tropopause. The height of the first layer is approximately $50\,\mathrm{m}$.

The model transport has been extensively validated using 222Rn and SF6 (Krol et al., 2005; Peters et al., 2004) and its further validation was performed within the EVERGREEN Project (Bergamaschi et al., 2006).

Gas phase chemistry is calculated using the CBM-IV chemical mechanism (Gery et al., 1989; Gery, 1989) modified by Houweling et al. (1998), solved by means of the EBI method (Hertel et al., 1993). Dry deposition is calculated using the ECMWF surface characteristics and the resistance method (Ganzeveld and Lelieveld, 1995).

Wet deposition is the dominant removal process for most aerosols. Removal occurs in convective systems (convective precipitation $-2.14\,\mathrm{Tg\,C\,year^{-1}}$) and in large scale systems ($5.86\,\mathrm{Tg\,C\,year^{-1}}$) that are associated with weather fronts. The in-cloud removal rates, which depend on the large scale precipitation and are differentiated for cumulus and stratiform precipitation, are calculated following Guelle et al. (1998) and Jeuken et al. (2001). Aerosol below-cloud scavenging is parameterised accordingly to Dana and Hales (1976). TM5 is coupled to the microphysical aerosol model M7 (Vignati et al., 2004, 2010a,b) that allows the resolution of particle masses and numbers. The model ability to evaluate BC concentration has been evaluated in Vignati et al. (2010a). The particles are represented by seven internally mixed classes, using a pseudo-modal approach. Four classes are for soluble mixed particles representing nucleation, Aitken, accumulation, and coarse mode, and three are for the insoluble (Aitken, accumulation, and coarse mode). Nucleation, condensation of sulphuric acid and coagulation between the particles are included. BC can be present in the insoluble and soluble Aitken modes, and in the soluble accumulation and coarse modes. The ageing is accomplished by considering condensation of H_2SO_4 and coagulation with soluble particles, which form a soluble shell around the hydrophobic core and the particles are moved from the insoluble to the soluble/mixed modes. The other components in M7 are mineral dust, primary organic carbon (OC), sulfate, and sea salt. All particles are removed in case of convective wet removal. In presence of large scale precipitations only the soluble accumulation and coarse modes are scavenged by rain, while the remaining modes (insoluble Aitken, accumulation, coarse and soluble nucleation and Aitken) form interstitial aerosols and they are not in-cloud removed. The soluble accumulation and coarse modes are assumed to form cloud droplets where the oxidation of SO_2 by O_3 and H_2O_2 takes place; the resulting sulphate is partitioned between the two modes as function of number of particles present in the modes (Stier et al., 2005). Below cloud scavenging removes all the particles as function of their size. The BC emission

inventories used in the present application are from Bond et al. (2004) for the anthropogenic contributions (fossil and bio fuels) and from van der Werf et al. (2004) for large scale biomass burning areas. Black carbon is assumed to be insoluble when emitted. The number of BC emitted particles is calculated assuming the freshly emitted particles with number median radii of 0.03 and 0.075 μm, for fossil/bio fuel and biomass burning, respectively, and emitted in the insoluble Aitken mode with standard deviation equal to 1.59 (adapted from Dentener et al. 2006)

3   Results and discussion

This study uses observations from July 2002 to June 2003 where possible, to be in agreement with time period of the TM5/M7 model runs. At Ispra, where light attenuation measurements are available from January 2004 onwards, the comparison is performed using observations for the period July 2004–June 2005. In recognition of this, montly data are used in the comparisons at all sites. Since no size fractionation is applied to BC model output, TSP or PM_{10} experimental measurements are used for the comparison.

3.1   Temporal representativeness of experimental measurements

First, we want to verify that the time period chosen for model evaluation is representative for the site based on longer time scale observations, when available. Figure 1 shows the time series of EBC concentrations at the different sites from January 2000 to December 2005. The variability of observations often ranges over one order of magnitude, with larger fluctuations (defined by the ratio 95th percentile to 5th percentile) at Jungfraujoch, Ispra, and Mace Head.

The black lines in the six panels of Fig. 1 correspond to the fit functions defined as a sinusoidal curve:

$$\mathrm{EBC}(t) = \mathrm{EBC}_0 + \frac{A_0}{2}\sin(\omega t + \varphi) \qquad (5)$$

t is the time variable, EBC_0 is the mean value, A_0 is the amplitude of the sine wave, ω is the frequency, and φ is the phase. The model fitting uses a Levenberg-Marquardt algorithm to search for the coefficient values that minimize chi-square. To verify the presence of seasonality in the time series, the initializing value of ω is calculated assuming the periodicity equal to one year. The fitting parameters A_0 are reported in Table 1.

At Alert, Jungfraujoch, and Ispra the resulting values of the amplitude (A_0) are significantly larger than the signal fluctuation, indicating the presence of a consistent seasonal cycle. To verify that the year July 2002–June 2003 is representative at each site, the standard deviation from the fit function over this period is compared to the standard deviation calculated for all the available measurements over 2000–2005. At all sites the standard deviation difference is
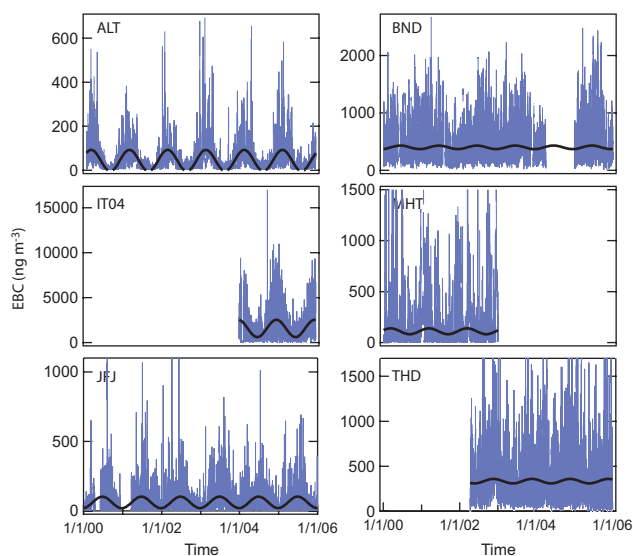

Fig. 1. Time series of EBC concentration (bleu) and sinusoidal fits (black) for available years at Alert (ALT), Bondville (BND), Ispra (IT04), Mace Head (MHT), Jungfraujoch (JFJ), and Trinidad Head (THD).

smaller than 15%, and the largest differences are observed at Trinidad Head (12%), and Alert (15%).

3.2   Variability within a grid box

Ground observations are performed at a specific location and thus, from a spatial point of view, they describe the behaviour of a parameter at a point, while models simulate the investigated parameter over a surface whose size is defined by the model grid resolution. Point observations thus need to be representative of the entire model grid cell in order to be used in model evaluation; such a requirement can be a critical issue when the spatial resolution of the model is low. Model data at Ispra, Jungfraujoch and Mace Head are available with a resolution of $1° \times 1°$, while at Alert, Bondville, and Trinidad Head the model resolution is only $4° \times 6°$, that is about 500×500 km. The large grid cell is a non-optimal condition that we want to test in order to discuss model evaluation. EBC concentrations are compared to light absorption and elemental carbon measurements collected at other measuring stations to verify that Alert, Bondville and Trinidad Head are representative sites of the model grid cells.

Alert is a remote site, located 1700 km to the north of the Polar Circle and at more than 2000 km from Iqaluit, the closest populated area. Figure 2a compares EBC monthly medians at Alert with EBC monthly medians at Ny-Alesund and Barrow, two polar remote sites located in Northern Europe and Alaska, respectively. EBC is calculated from light absorption measurements at 880 nm with an aethalometer AE-31 at Ny-Alesund, and at 550 nm with a PSAP at Barrow.
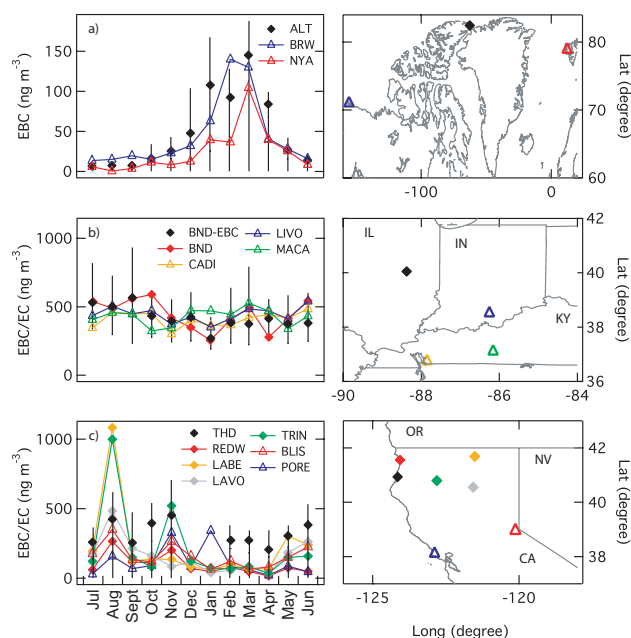
Fig. 2. 50th percentile (black diamonds) and 25th–75th percentile range (black lines) of EBC concentrations at Alert **(a)**, Bondville **(b)**, and Trinidad Head **(c)**, together with EBC and EC experimental values at Point Barrow (BRW), Ny–Alesund (NYA), Cadiz (CADI), Livonia (LIVO), Mammoth Cave (MACA), Redwood National Park (REDW), Lava Beds (LABE), Lassen Volcano (LAVO), Trinity (TRIN), Bliss State Park (BLIS), and Point Reyes National Seashore (PORE); closed diamonds represent locations inside the same model grid cell of the study site.

The EBC trend observed at Alert is common to the other polar locations, even if farther than 1000 km, indicating that sources, sinks, and transport pathways are similar for the three sites. This is consistent with the fact that the pristine environment surrounding Alert for hundreds of kilometres makes it representative of the $4° \times 6°$ cell.

Figure 2 compares EBC concentrations at Bondville (panel b) and Trinidad Head (panel c) to elemental carbon (EC) concentrations at different sites in the same $4° \times 6°$ grid cell or in the closest adjacent cell. The EC sites belong to the IMPROVE network, where $PM_{2.5}$ aerosol samples are collected twice a week for 24 h. Although corresponding to a different size fraction, EC measurements are used for the comparison assuming that the amount of LAC present in particles larger than 2.5 μm is negligible compared to the amount of LAC in smaller particles. EC quantification is based on aerosol thermal properties, while EBC measurements are based on aerosol optical properties. We do not expect a perfect agreement between optical and thermal measurements (Watson et al., 2005), but at Bondville, where both EBC and EC measurements are available, their monthly medians agree within 20%, suggesting that the comparison is accurate enough for the purposes of this study.

The IMPROVE sites compared to Bondville are Cadiz KY (36.7° N 87.8° W), Livonia IN (38.5° N 86.3° W), and Mammoth Cave KY (37.1° N 86.1° W). The three sites are located in the grid cell adjacent to Bondville and are less than 380 km away. The four sites do not show a seasonal cycle over the study period. Other than January medians at Mammoth Cave, EC medians are within the 25th–75th percentile of EBC measurements, and on average the difference between EC and EBC medians is less than 5%.

EBC monthly medians at Trinidad Head are compared to EC monthly medians at Redwood National Park CA (41.6° N 124.1° W), Lassen Volcanic National Park CA (40.5° N 121.5° W), Lava Beds CA (41.7° N 121.5° W), Trinity CA (40.7° N 122.8° W), Bliss State Park CA (38.9° N 120.1° W), and Point Reyes National Seashore CA (38.1° N 122.9° W). Point Reyes is located to the south of Trinidad Head cell at 350 km distance. Redwood National Park and Point Reyes National Seashore are marine sites, situated along the northern coast of California like Trinidad Head. EBC monthly medians during December 2002 and January 2003 are not reported because data capture is smaller than 30%. The Trinidad Head grid cell thus demonstrates greater spatial variability than found at Bondville. The variability of EBC and EC measurements is larger during the period July–November compared to the following months; in particular, during February, March, and April the EC monthly averages differ by less than 10% of the variability range. Although EC medians are usually within the 25th–75th percentile range of EBC measurements, EBC medians are consistently higher then EC medians with higher discrepancies during spring. The use of a default specific absorption coefficient rather than a site calibrated one may be responsible for the systematic overestimation. In addition, dust can interfere with EBC measurements, especially during the spring season when Asian dust plumes might reach the west coast of the United States (Perry et al., 2004; Zhao et al., 2008).

Assuming that the concentration of dust in PM_{10} at Trinidad Head is equal to 10 μg m^{-3} the corresponding light attenuation would be about 0.3 Mm^{-1} (Yang et al., 2009); this would lead to an overestimation of EBC of 30 ng m^{-3}, which is not enough to explain the discrepancy observed between EBC and EC measurements. To further investigate the role of absorption cross section coefficient uncertainty, EBC concentrations are calculated assuming σ^* equal to 20 m^2 g^{-1}, as suggested by Huffman (1996). Figure 3 shows a better agreement between EC and EBC when a larger σ^* is employed and the differences between EC and EBC measurements are reduced to 30%. The larger absorption cross section coefficient of Trinidad Head compared to Bondville could be justified by the presence of more processed aerosols at the coastal site; aged aerosol is mixed with sulfate and other organic species that enhance ability of BC to absorb visible light through the lens effect (Liousse et al., 1993).
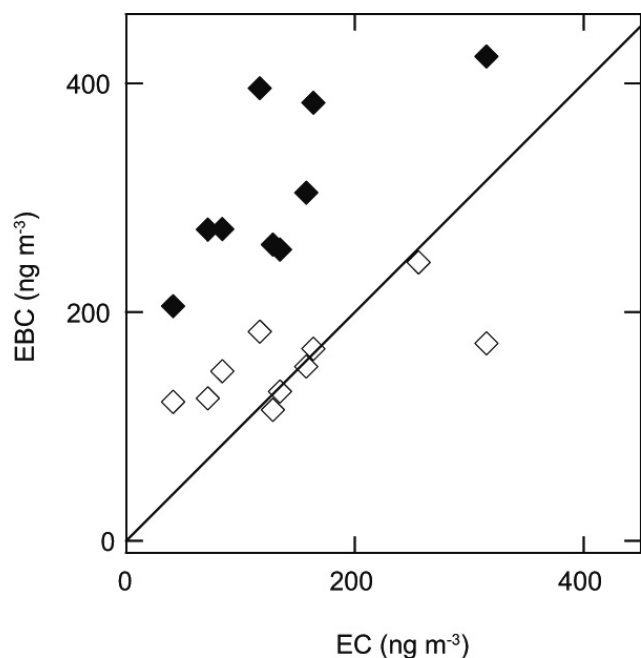
Fig. 3. Scatter plot of EBC concentration at Trinidad Head calculated with σ^* equal to $10\,\mathrm{m}^2\,\mathrm{g}^{-1}$ (full diamonds) and $20\,\mathrm{m}^2\,\mathrm{g}^{-1}$ (open diamonds) versus the average EBC concentration at the IMPROVE sites; black line indicates the 1:1 ratio.

EBC at Bondville compares well with EC at the same site and at the nearest IMPROVE sites, indicating both similarity among rural and regional background sites, and an adequate choice of σ^* to estimate EBC from optical measurements. On the contrary, the Trinidad Head grid cell is characterized by a larger variability of EC and EC is systematically smaller than EBC; the discrepancy is likely due to an incorrect σ^* rather than dust interference on optical measurements; at Trinidad Head the uncertainty of black carbon optical properties and the variability of the grid cell compromise the use of EBC observations for model evaluation.

4 Statistical methods

The statistical tools used in the following paragraphs include Fast Fuorier Transformation (FFT) analysis, skewness analysis, median comparison, variability analysis, probability distribution curve comparison, Student's t-test, and Mann-Whitney test. The FFT analysis is used to identify the frequency and amplitude of data modulation. The skewness analysis investigates the symmetry of data distribution (of model or observations); when skewness is zero the data are symmetrically distributed around their average value. Median comparison and variability analysis are used to test the similarity of two data distributions: the median comparison focuses on the similarity of central values of each distribu-

tion, while the variability analysis investigates how data are spread around the central values. The similarity of two data distributions can further be investigated with the overlap of the probability distribution curves, the Student's t-test, and the Mann-Whitney test; in the first two tests, higher value of the output (overlap area and test probability) corresponds to higher similarity of data distributions, while the Mann-Whitney test shows lower output parameter Z for higher similarity of distributions.

In the following sections we will use the different statistical tools within three tests with three different objectives. The comparison of monthly medians is used to evaluate the ability of the model to simulate the time trend of measurements over a longer time period. The variability analysis illustrates how data spread around the median of model and observations, to understand the time scale of possible model failure. The comparison of model and observation probability distributions summarizes the information from the first two tests, quantifying the agreement of measurements and simulations during a time period shorter than one year, i.e. a month or a season.

4.1 Comparison of model and observation time series

Figure 4 shows box whisker plots of 1-hour resolution data of observed EBC concentrations (on the left side), and TM5/M7 model BC concentrations (on the right side), from July 2002 to June 2003. Ispra panel reports measurements from July 2004 to June 2005. Boxes and whiskers are not reported during months characterized by capture smaller than 30%. 5th, 25th, 50th, 75th, and 95th percentiles are calculated for each month. The median rarely falls in the middle of the 25th–75th percentile interval or the 5th–95th percentile interval, indicating that the 1-hour resolution data (both observations and model) when considered month by month are not normally distributed. Skewness is used to quantify the deviation relative to the normal Gaussian curve. If the skewness is zero, the distribution is symmetric like the bell shaped normal curve, with a positive skew the distribution curve shows a tail towards larger values, and with a negative skew the distribution curve shows a tail towards smaller values. The observations are characterized by positive skew that averages between 1 (Bondville and Ispra) and 3.5 (Mace Head). The model data distributions are characterized by lower skew parameters, varying between 0 (Ispra) and 2.2 (Mace Head). The mean value of an asymmetric data distribution is strongly affected by even a small number of extremely high or extremely low values. For this reason we use the monthly medians rather than the monthly means to compare the time series of model and observations.

At Alert, Ispra and Jungfraujoch, where a seasonal cycle is detected (see Sect. 3.1), the model reproduces this cycle, with maxima during winter at Alert and Ispra, and during summer at Jungfraujoch. At Bondville and Mace Head no specific trend is observed in either the measurements or the model
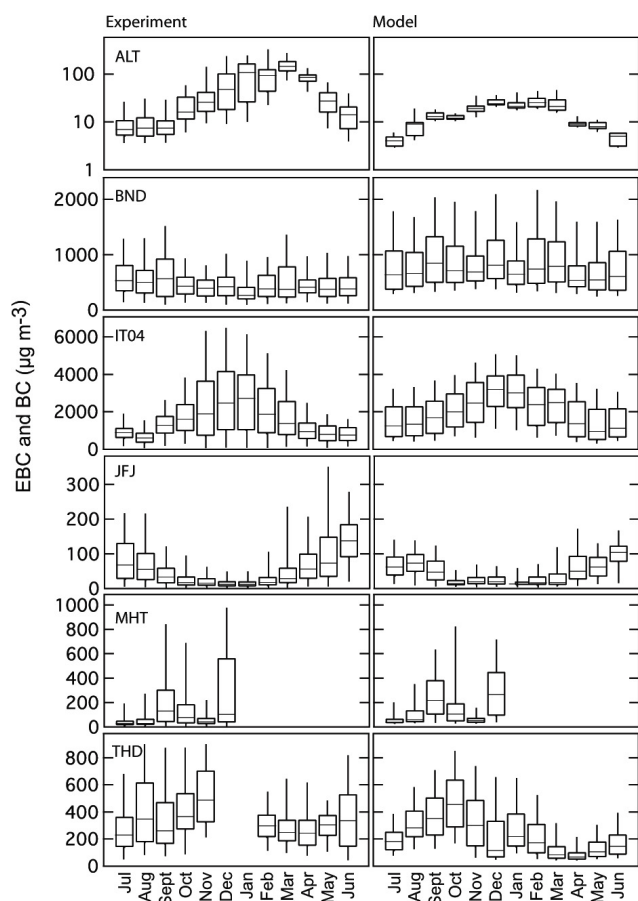
Fig. 4. Time series of experimental EBC and model BC concentrations, from July 2002 to June 2003; central point of the boxes correspond to median values, boxes indicate 25th–75th percentile intervals, and whiskers indicate 95th-5th percentile intervals.

during the year, while at Trinidad Head observations and model shows a similar trend with slightly higher EBC/BC concentrations during fall compared to spring.

At Alert, observed monthly medians range between 7 and 140 ng m^{-3}, while model monthly medians range between 4 and 25 ng m^{-3}. The winter maxima are due to fewer precipitation events and atmospheric circulation patterns that control long-range transport (Sharma et al., 2006); the modelled BC concentration during winter and spring is more than 5 times smaller than observed EBC levels; the similarity of Alert EBC time series with those recorded at other polar sites (Fig. 3a) suggests that the model/observation disagreement is due to the inability of the model to simulate the Arctic haze. During early summer and fall (July–November) the disagreement between model and observations averages 40% of observations.

The highest concentrations are recorded at Bondville and Ispra. Bondville has a limited month-to-month variation; median values of observations and model vary in the range

275–570 ng m^{-3}, and 540–850 ng m^{-3}, respectively; on average model monthly medians are 70% higher than observations, with larger overestimation during January–March. The higher concentrations calculated by the model might be attributed to the model grid size, which includes rural areas together with large cities (i.e. Chicago), whose plume rarely reach Bondville. At Ispra model monthly medians vary between 1000 and 3000 ng m^{-3}, while observed monthly medians vary between 600 and 3000 ng m^{-3}. Winter measurements are three times higher compared to summer, most probably due to larger emissions from residential heating, and more frequent events of temperature inversion (Yttri et al., 2007). The model overestimates observations significantly during August and March, when modelled BC is roughly twice as large as measured EBC.

Experimental medians at Jungfraujoch vary between 10 and 140 ng m^{-3}. The range is well simulated by the model (10–110 ng m^{-3}), which on average differs from the observation by less than 5%. Higher concentrations are observed and simulated during summer, when polluted air masses from lower altitude are transported to the site by thermal convection (Baltensperger et al., 1997; Cozic et al., 2008).

At Mace Head, observation and model medians vary in the interval 25–130 ng m^{-3} and 30–270 ng m^{-3}, respectively. The highest concentrations are reported during September, October, and December, both by observations and by model. On average model medians overestimate observations by 70%.

Observed monthly medians at Trinidad Head are in the range 230–480 ng m^{-3}, while model medians are 60–450 ng m^{-3}. EBC and EC concentration maxima are in October and November due to a larger contribution of North American emissions in the cold season. Model underestimates by 30% the observation medians, and larger disagreement takes place during spring (on average 50%).

4.2 Comparison of variability

In this section we illustrate how to investigate the ability of model to simulate observation variability, and to verify if the model fails in simulating the daily modulation. First we compare 1-hour and 24-hour resolution data to identify cases when variability is dominated by daily modulation, and then we use the FFT analysis to verify how model describes this modulation.

The comparison of monthly median values alone does not describe in detail the ability of the model to simulate the observations; the data variability around the median plays a crucial role in defining the agreement of model and experiment. Since model and observation data are not normally distributed, the standard deviation is not a meaningful parameter to represent data variability, instead the variability is expressed as the difference between 95th and 5th percentiles of 1-hour resolution data. The monthly variability of observations ranges from 75% of the median at Jungfraujoch to

750% at Mace Head, and variability of model ranges from 90% of the median at Alert to 500% at Mace Head. Figure 5a shows the ratio of model variability to observed variability per each month. The ratio is effectively an indication of how much variability of the observations is explained by the model. The ratio is expected to be smaller than one; primarily due to the effects of local sources, which are not represented in the model, but are responsible for fluctuations in pollutant concentration on a short time scale and can be responsible for larger variability of observations compared to model simulations. In addition, variability of meteorological parameters is not resolved on the large model grid. Most of the time the ratio is in the range 0.5–1 indicating that the model explains more than 50% of measurement variability. Bondville in winter and Ispra in summer show ratios larger than 1.5, while at Alert the ratio is always below 0.5.

To understand the causes of model failure, the variability of 1-hour resolution and 24-hour resolution data are investigated. Figure 6 illustrates this with two examples. When the 1-hour resolution data variability is dominated by the diurnal fluctuations (case A), the variability of 24-hour resolution data is much smaller than the variability of 1-hour resolution data and their ratio tends to 0. On the contrary, when the daily modulation is smoother (case B), the variability of 24-hour and 1-hour resolution datasets are similar and their ratio tends to 1. Figure 5 shows the variability ratio of 24-hour resolution data to 1-hour resolution data of measurements (panel b) and simulations (panel c). Generally, the ratio ranges between 0.5 and 1 indicating that the daily modulation does not dominate the 1-hour resolution data variability. Conversely the modelled concentrations at Ispra and Bondville have a stronger daily modulation, and at the two sites the ratio shows a seasonal trend with values smaller than 0.5 during spring and summer.

The daily variability is further investigated with the Fast Fourier Transformation (FFT) analysis, which identifies the frequency and amplitude of data modulations. Table 3 reports the intensity of the daily modulations of modelled concentrations and observations. At Jungfraujoch the daily modulation is perfectly simulated by the model. At Alert, the model fails to simulate the daily variability, but since the 1-hour resolution data variability is larger than the daily modulation, this is not enough to explain the low ratio observed in Fig. 5a. At Bondville and Ispra the model overestimates the daily variability. Bondville does not show any seasonal differences between warm and cold seasons, while the model overestimation of daily variability at Ispra is 5 times higher during summer compared to winter.

Model overestimates observation variability at Bondville and Ispra, and underestimates it at Alert. At Ispra the overestimation is limited to the summer months, when the model variability is dominated by the daily modulation and the model overestimates daily modulation of observations. At Bondville and Alert the model fails in reproducing both daily modulations and 1-hour resolution data variability.
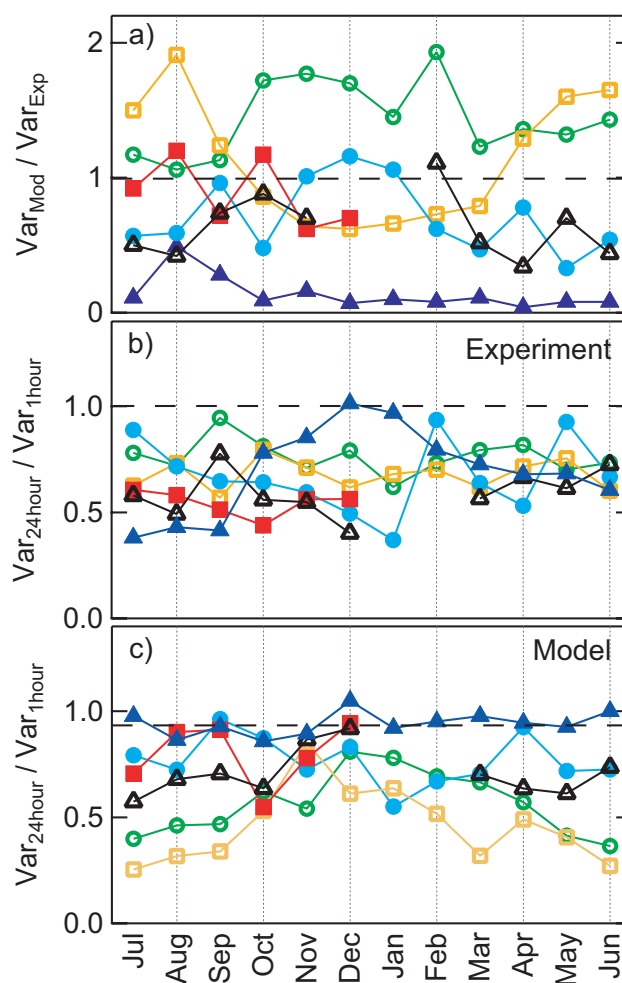

Fig. 5. Panel **(a)**: explained variability, defined as ratio of model variability to experiment variability, from July 2002 to June 2003, at Alert (blue triangle), Bondville (green open circles), Ispra (yellow open squares), Jungfraujoch (light blue circles), Mace Head (red squares), and Trinidad Head (black triangles). Panel **(b)** and **(c)**: variability ratio of 24-hour resolution data to 1-hour resolution data for experiment **(b)** and model **(c)**.

4.3 Model/observation agreement

We compare here four different statistical tools to quantify the agreement between model and observations. The most accurate way to quantify this agreement is the overlap of the probability distribution curves. Considering that this method is elaborate, we discuss limitations and advantages of alternative and more common statistical tools: comparison of medians, Student's t-test, and Mann-Whitney test.

Probability distributions (PD) of 1-hour resolution model data and observations are calculated for each month. The distribution curves are normalized to 100 and the overlap of model and observation curves is calculated. The overlap is
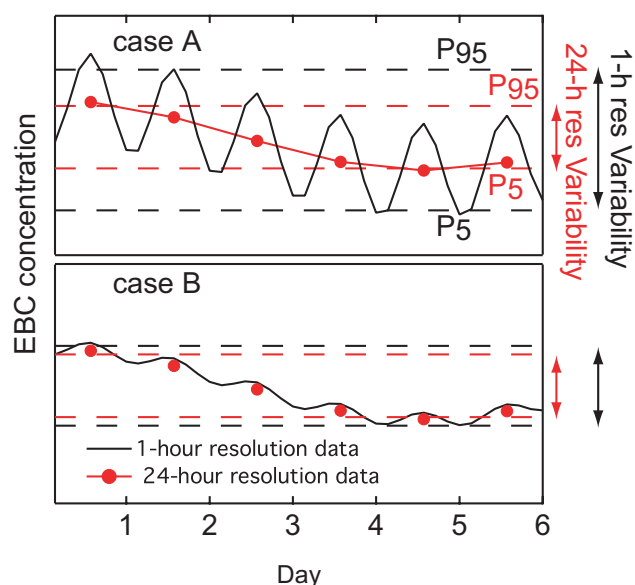
Fig. 6. Example of overall variability and day-to-day variability defined as the difference between 95th percentile (P_{95}) and 5th percentile (P_5) of 1-hour resolution data and 24-hour resolution data, respectively.
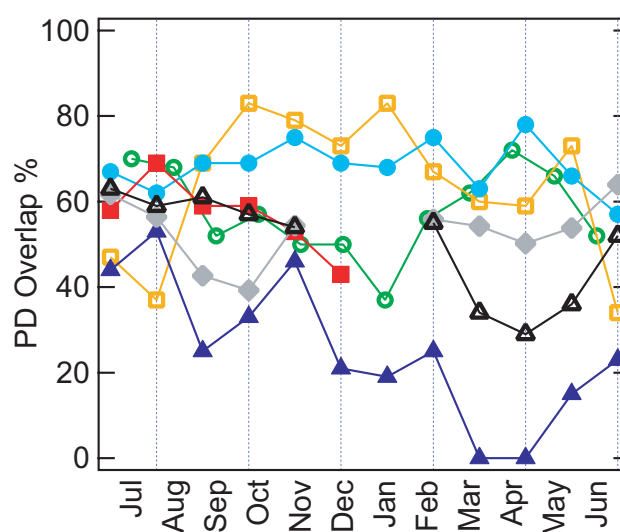

Fig. 7. Agreement of model BC and experimental EBC concentrations determined by overlap of probability density curves at Alert (blue triangle), Bondville (green pen circles), Ispra (yellow open squares), Jungfraujoch (light blue circles), Mace Head (red squares), and Trinidad Head (black triangles). Trinidad Head agreement is calculated with σ^* equal to $10\,m^2\,g^{-1}$ (black triangles) and $20\,m^2\,g^{-1}$ (gray diamonds).

Table 3. Intensity of daily modulation from FFT data analysis ($ng\,m^{-3}$).

Site	Experiment	Model
ALT	4.0	–
BND	60	401
IT04	565	1058
JFJ	6.9	5.3
MHT	100	27
THD	59	96

100% for perfectly identical distribution curves and 0% for distribution curves that do not overlap at all. The PD overlap quantifies the model/observation agreement in a more accurate way than the simple comparison of monthly medians, because it takes into account, not only the central values of model and observations data but also variability and probability distribution shapes. PDs of 24-hour resolution data are used at Bondville and Ispra: at Bondville the variability of 1-hour resolution data has already been shown not to be simulated adequately by the model, while at Ispra the use of 1-hour resolution data could be misleading since meteorological data used by the model (July 2002–June 2003) do not correspond to the measurement period (July 2004–June 2005). Figure 7 shows the PD overlap of monthly data at each site and during each month. PD overlap values at Trinidad Head are reported for σ^* equal to $10\,m^2\,g^{-1}$ and $20\,m^2\,g^{-1}$ to understand how the uncertainty in σ^* af-

fects the agreement between model and observations. The PD overlap averages 25% (0–53%) at Alert, 57% (37–72%) at Bondville, 64% (34–83%) at Ispra, 68% (57–78%) at Jungfraujoch, 57% (43–69%) at Mace Head, and 53–57% (39–64%) at Trinidad Head. To validate the comparison at Ispra, the period July 2002–June 2003 has been compared to the period July 2004–June 2005 using 24-hour resolution sulfate measurements performed at the same site. The PD overlap of sulfate data, on average 68%, indicates that the two periods do not differ significantly for sulfate concentration and support the assumption that the black carbon measurements can be used for model evaluation. The worst agreement between model and experiment are observed at Alert, with minimum during the spring haze period, as expected from the median comparison. The highest values are recorded at Jungfraujoch where the agreement is constantly above 55% during the entire year. Bondville and Ispra show the same seasonal trends already seen, with higher agreements during the warm and the cold season, respectively. At Trinidad Head the model/experiment agreement is lower than 50% in winter if σ^* is assumed equal to $20\,m^2\,g^{-1}$ and in spring if σ^* is assumed equal to $10\,m^2\,g^{-1}$. This result might indicate a seasonal dependence of σ^* or the model failure during one season. A better knowledge of black carbon optical properties at the site is mandatory to use light attenuation measurements for model evaluation.
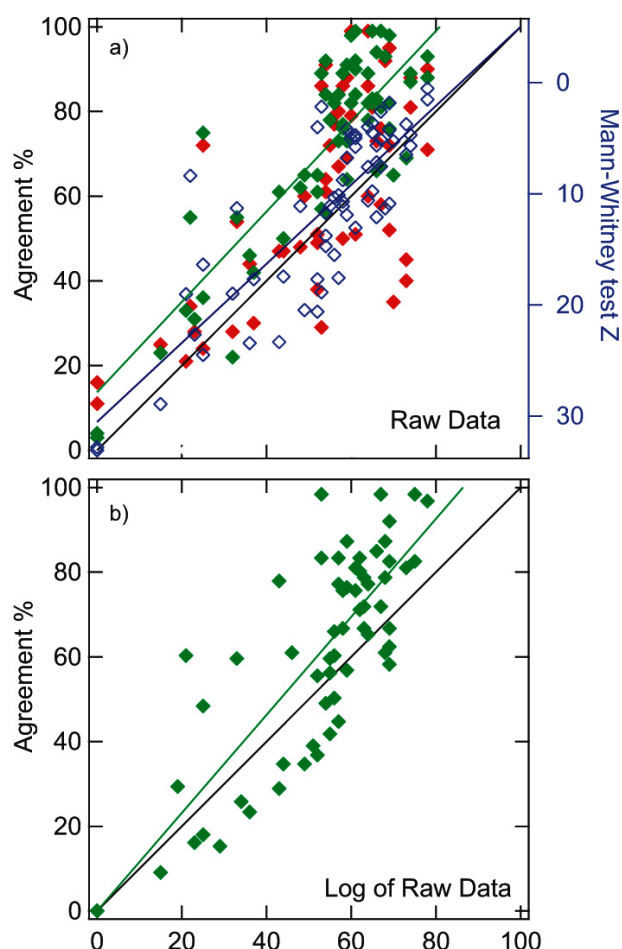
Fig. 8. Model experiment agreement determined with Student's t-test (green) and median comparison (red) plotted versus the PD overlap percentage; rank test Z (blue) is plotted on the right vertical axis. Linear fits are shown for Student's t-test and Mann-Whitney test; black line is the 1:1 ratio. Panels show agreement of original data **(a)** and agreement of logarithmic transformation of original data **(b)**.

Figure 8 compares the PD agreement to the probability of agreement evaluated with parametric (comparison of medians and Student's t-test) and nonparametric (Mann-Whitney test) statistical tools. The PD overlap is first compared to the median agreement, defined as:

$$\text{Median agreement} = 100 - 100 \frac{|\text{median}_{\text{exp}} - \text{median}_{\text{mod}}|}{\text{median}_{\text{exp}}} \quad (6)$$

The median agreement and PD overlap do not co-vary linearly: when the PD agreement is larger than 50%, the median comparison might overestimate or underestimate the agreement by more than 20%.

The agreement between two datasets characterized by normal distribution is usually calculated with Student's t-test, which evaluates the difference of the means based on the dis-

tribution standard deviations. Although model and experiment data are not normally distributed, we compare the t-test probability to the PD overlap to assess the error introduced by the normal distribution assumption. Observations and model data are treated as two independent samples with means X_{obs} and X_{mod}; t is calculated as:

$$t = \frac{X_{\text{obs}} - X_{\text{mod}}}{\sqrt{s_{\text{obs}}^2/n_{\text{obs}} + s_{\text{mod}}^2/n_{\text{mod}}}} \quad (7)$$

where s and n indicates standard deviations and number of data points, respectively. The null hypothesis states that the population means of observations and model data are the same. The comparison of t with the Student's t-distribution table defines the probability that the null hypothesis is true. Student's t-test probabilities co–vary with PD overlap ($r^2 = 0.70$) (Fig. 8a), although the t-test overestimates the agreement when the PD agreement is higher than 60%. The t-test percentages show a positive bias equal to 14% $\pm$ 5%.

EBC measurements and modelled BC data are not normally distributed. Skew parameters indicate that the probability distributions are characterized by a tail towards larger values, so logarithmic transformation of the raw data should obtain a distribution that is closer to a Gaussian distribution. Monthly means and standard deviations of the logarithm of EBC and BC concentrations are used to recalculate the agreement probability with the Student's t-test (Fig 8b). The t-test probability of logarithmic data on average is not biased, but lower agreements are underestimated and higher agreements are overestimated.

The comparison of two datasets characterized by nonnormal distributions can be performed with the Mann-Whitney U-test, which is also known as Mann-Whitney-Wilcoxon test or rank sum test. The null hypothesis states that the two samples are drawn from the same population, thus their distributions are equal. Observations and model data are put into a single array in increasing order, rank are assigned to each data (the smaller value is ranked 1), the sum of ranks (R) is computed separately for model and observations, and the parameters U are calculated as:

$$U_{\text{mod}} = R_{\text{mod}} - \frac{n_{\text{mod}}(n_{\text{mod}} + 1)}{2} \quad (8)$$

$$U_{\text{exp}} = R_{\text{exp}} - \frac{n_{\text{exp}}(n_{\text{exp}} + 1)}{2} \quad (9)$$

n_{mod} and n_{exp} are the number of model and experiment data points. Similar model and observation distributions lead to similar U_{obs} and U_{mod}. The difference between the two distributions is quantified by Z:

$$Z = \frac{|\mu - U_{\text{mod}}|}{\sigma} \quad (10)$$

where μ is the mean value of U_{obs} and U_{mod} (it is uninfluential if U_{obs} or U_{mod} is used in the formula), and σ is defined as:

$$\sigma = \sqrt{\frac{n_{\mathrm{mod}}\, n_{\mathrm{exp}}}{N\,(N-1)} \left[\frac{N^3 - N}{12} - \sum_{j=1}^{g} \frac{t_j^3 - t_j}{12} \right]} \qquad (11)$$

where N is the sum of n_{obs} and n_{mod}, g is the number of groups of ties, and t_j is the number of tied ranks in group j. Z is used to calculate the significance probability associated to the null hypothesis; larger Z corresponds to larger difference between U_{mod} and U_{exp} and thus to lower agreement between model and observations.

The Mann-Whitney test is applied to model and observations: 1-hour resolution monthly data are analyzed. The smallest values are calculated for Jungfraujoch while the highest values (larger than 20) are observed at Alert, Trinidad Head, and Bondville. At Alert the largest Z values occur during spring (March and April) when the model fails in simulating the Arctic haze. At Bondville Z is smaller during summer and higher during winter, in agreement with the time trend reported in Fig. 7 where higher PD similarity is observed in summer and lower similarity in winter. At Trinidad Head the highest Z values occur in spring when model and observations mostly disagree. Z values calculated at all sites are compared to PD overlap in Fig. 8a. Z anti-correlates with the PD overlap, as expected, but occasionally distributions characterized by a satisfactory agreement (larger than 60%) are associated with high Z values (larger than 20). The low correlation ($r^2 = 0.67$) is likely due to the different shapes of model and observation distribution curves: although Mann–Whitney test does not assume that the two populations of data are normally distributed, it requires that the shapes of their distributions are identical. For this reason we re-applied the Mann-Whitney test only to those months for whom model does not overestimate observation variability (see Sect. 4.2). Considering only months when the model to observed variability ratio is smaller than 1, the correlation of Z values and PD overlap area is improved ($r^2 = 0.79$).

Monthly medians, Student's t-test, and Mann-Whitney test are employed to quantify the agreement between model and observation data distributions and compared to the probability distribution curve overlap method. The three tools are based on different sets of assumptions that are not always encountered in real population samples. If these assumptions are not discussed and the methods are applied without discrimination to the entire dataset, the result can be misleading.

In summary, the comparison of median is inaccurate and it should be discouraged because ignores the data distribution around the medians. The Student's t-test is a good choice only when the data distribution is similar to a normal distribution; this assumption is rarely verified and its testing can be time consuming. The Mann-Whitney test is a satisfactory alternative to the comparison of PD curves, but it requires that observation and model data distribution have the same shape; this can be easily verified by comparison of observation and model variability.
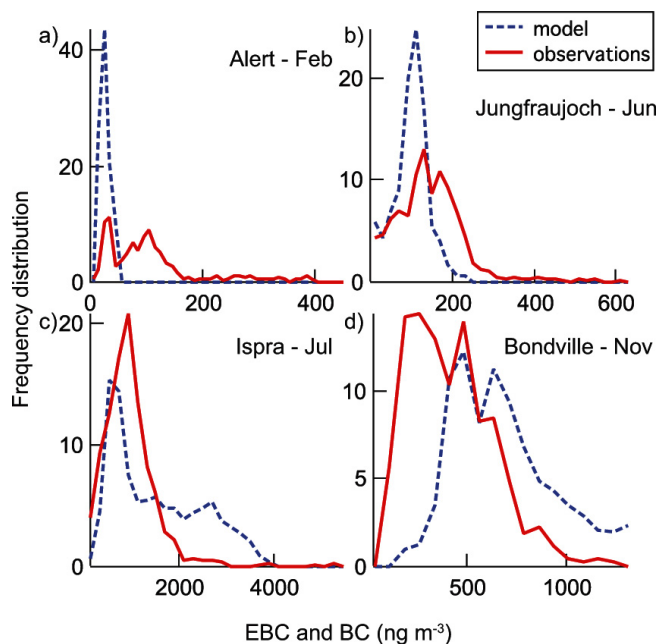

Fig. 9. PD curves of model (blue) and observations (red) at Alert in February (**a**), at Jungfraujoch in June (**b**), at Ispra in July (**c**), and at Bondville in November (**d**)
.

4.4 PD overlap summary

We explore here the physical meaning of PD overlap and how PD overlap analysis can be used to discuss model performance. For this purpose we focus on four cases that were already discussed in the previous sections.

At Alert the model is not able to reproduce the transport of pollutants from lower latitudes, which leads to PD overlap ranging between 0 and 25%. In March and April model distribution data is narrow and centered around BC concentrations smaller than $50\,\mathrm{ng\,m^{-3}}$, while observations are shifted towards higher concentrations; from December to February observations distribution is bimodal and the model simulates only the lower concentration mode, again below $50\,\mathrm{ng\,m^{-3}}$ (Fig. 9a). The model does not show higher BC concentrations since it fails in describing pollutant transport taking place at the end of winter and in spring. A similar effect is observed at Jungfraujoch in summer, when local sources might affect the sampling site. The model fails in describing the transport of emissions from lower altitudes (Fig. 9b), and thus it does not simulate the occasionally higher BC concentrations in May–August. This results in PD overlap slightly smaller during summer than in winter.

Figure 9c shows that at Ispra in summer the model data distribution has a bimodal shape and the mode at concentrations larger than $2000\,\mathrm{ng\,m^{-3}}$ is not present in the observations. This result is likely due to the model

underestimation of dilution processes, which also leads to the overestimation of daily cycle amplitude, as seen during variability analysis.

At Bondville the model does not simulate the lower BC concentrations (Fig. 9d). This could be due to underestimation of dilution, but more likely by the large model grid cell that contains industrial, urban, and background areas.

5 Conclusions

Model evaluation studies are based on the comparison of model data and observations. The goal of this paper is to discuss, compare and evaluate a variety of statistical and analytical tools in order to better inform such evaluation.

Temporal representativeness and variability within a grid box of observations must be assessed before observations are used to evaluate model simulations. Temporal representativeness is tested when longer time series of measurements are available. Seasonal trend over multiple years is defined by a fitting sinusoidal curve and the scatter of data points relative to the fitting function is used to define the similarity of the investigated time series to previous and following years. Site measurements should be representative of the entire grid cell defined by the model. Investigation of spatial representativeness is recommended especially for lower model resolution. During the present study EBC measurements are compared to EC and light attenuation measurements collected at different sites inside the same grid cell (or close to the investigated location).

This study illustrates a suite of tests that should be applied to evaluate black carbon model with EBC measurements. Three tests are proposed: the comparison of monthly medians, the variability analysis, and the comparison of monthly probability distribution curves. The median comparison illustrates the ability of the model to simulate the seasonal trend of black carbon measurements. The variability explained by the model is defined as the ratio of measurement variability to model variability of 1-hour resolution data; it allows a preliminary evaluation of model performance. The comparison of normalized frequency distribution curves (or probability distribution PD curves) leads to a more complete and accurate model evaluation, compared to the previous tests. The agreement, defined by the overlapping area of the two curves, takes into account both the similarity of the median values and the variability of the data distributions. In addition the comparison of PD curves reflects the similarity of the data distributions.

Student's t-test and Mann-Whitney test are compared to the PD agreement to evaluate their limitations. Student's t-test requires that data are normally distributed and Mann-Whitney test requests that data distributions have the same shape. Typically neither EBC measurements nor modelled BC concentration time series fulfil such requirements and these tests might overestimate the agreement when PD over-

laps are higher than 60%. Applying a logarithmic transformation to EBC measurements and BC simulations to reduce skewness does not improve the accuracy of Student's t-test.

Study of long measurement time series and comparison of observations collected over a larger geographical area are required to guarantee temporal and spatial representativeness of experimental data. The agreement between model and observation should be defined by the PD curve overlapping areas; the use of Student's t-test and Mann-Whitney is discouraged. When model fails, the comparison of model and data variability, together with Fast Fourier Transform analysis, helps to identify the reason of model/observation disagreement.

Edited by: K. Carslaw

References

Alexander, M. L., Crozier, P. A., and Anderson, J. R.: Brown carbon spheres in East Asia outflow and their optical properties, Science, 833–836, 2008.

Alfaro, S. C., Lafon, S., Rajot, J. L., Formenti, P., Gaudichet, A., and Maillei, M.: Iron oxides and light absorption by pure desert dust: An experimental study, J. Geophys. Res., 109, 08201–08209, 2004.

Andreae, M. O. and Gelencsér, A.: Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols, Atmos. Chem. Phys., 6, 3131–3148, doi:10.5194/acp-6-3131-2006, 2006.

Andreae, M. and Merlet, P.: Emission of trace gases and trace gases and aerosols from biomass burning, Global Biogeochemi. Cy., 15, 955–966, 2001.

Baltensperger, U., Gaggeler, H. W., Jost, D. T., Lugauer, M., Schwikowski, M., Weingartner, E., and Seibert, P.: Aerosol climatology at the high-alpine site Jungfraujoch, Switzerland, J. Geophys. Res., 102, 19707–19715, 1997.

Barnard, J. C., Volkamer, R., and Kassianov, E. I.: Estimation of the mass absorption cross section of the organic carbon component of aerosols in the Mexico City Metropolitan Area, Atmos. Chem. Phys., 8, 6665–6679, doi:10.5194/acp-8-6665-2008, 2008.

Bergamaschi, P., Meirink, J. F., Müller, J. F., Körner, S., Heimann, M., Bousquet, P., Dlugokencky, E. J., Kaminski, U., Vecchi, R., Marcazzan, G., Meinhardt, F., Ramonet, M., Sartorius, H.,

and Zahorowski, W.: Model Inter-comparison on Transport and Chemistry – report on model inter-comparison performed within European Commission FP5 project EVERGREEN, EUR 22241-EN, 2006.

Bond, T. C. and Bergstrom, R. W.: Light absorption by carbonaceous particles: An investigative review, Aerosol Sci. Technol., 40, 1–41, 2006.

Bond, T. C., Anderson, T. L., and Campbell, D.: Calibration and intercomparison of filter-based measurements of visible light absorption by aerosols, Aerosol Sci. Technol., 30, 582–600, 1999.

Bond, T. C., Streets, D. G., Yarber, K. F., Nelson, S. M., Woo, J.-H., and Kilmont, Z.: A technology-based inventory of black carbon and organic carbon emissions from combustion, J. Geophys. Res., 109, D14 203, 2004.

Boylan, J. W. and Russell, A. G.: PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models, Atmos. Environ., 40, 4946–4959, 2006.

Cachier, H.: Atmospheric particles, chap. Carbonaceous combustion aerosol, p. 622, John Wiley & sons, Inc., 1998.

Chang, J. C. and Hanna, S. R.: Air quality model performance evaluation, Meteorol. Atmos. Phys., 87, 167–196, 2004.

Cooke, W. F., Jennings, S. G., and Spain, T. G.: Black carbon measurements at Mace Head, 1989–1996, J. Geophys. Res., 102, 25339–25346, 1997.

Corrigan, C. E., Ramanathan, V., and Schauer, J. J.: Impact of monsoon transitions on the physical and optical properties of aerosols, J. Geophys. Res., 111, D18208, doi:10.1029/2005JD006, 2006.

Cozic, J., Verheggen, B., Weingartner, E., Crosier, J., Bower, K. N., Flynn, M., Coe, H., Henning, S., Steinbacher, M., Henne, S., Collaud Coen, M., Petzold, A., and Baltensperger, U.: Chemical composition of free tropospheric aerosol for PM1 and coarse mode at the high alpine site Jungfraujoch, Atmos. Chem. Phys., 8, 407–423, doi:10.5194/acp-8-407-2008, 2008.

Dana, M. T. and Hales, J. M.: Statistical aspects of the washout of polydisperse aerosols, Atmos. Environ., 10, 45–50, 1976.

Dentener, F., Kinne, S., Bond, T., Boucher, O., Cofala, J., Generoso, S., Ginoux, P., Gong, S., Hoelzemann, J. J., Ito, A., Marelli, L., Penner, J. E., Putaud, J.-P., Textor, C., Schulz, M., van der Werf, G. R., and Wilson, J.: Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom, Atmos. Chem. Phys., 6, 4321–4344, doi:10.5194/acp-6-4321-2006, 2006.

Derwent, R. G., Ryall, D. B., Jennings, S. G., Spain, T. G., and Simmonds, P. G.: Black carbon aerosol and carbon monoxide in European regionally polluted air masses at Mace Head, Ireland during 1995-1998, Atmos. Environ., 35, 6371–6378, 2001.

Eleftheriadis, K., Vratolis, S., and Nyeki, S.: Aerosol black carbon in the European Arctic: Measurements at Zeppelin station, Ny-Ålesund, Svalbard from 1998–2007, Geophys. Res. Lett., 36, L02809, doi:10.1029/2008GL035741, 2009.

Ganzeveld, L. and Lelieveld, J.: Dry deposition parameterization in a chemistry general circulation model and its influence on the distribution of reactive trace gases, J. Geophys. Res., 100, 20999–21012, 1995.

Gery, M. W.: Development and testing of the CBM-IV for urban and regional modelling, EPA, 1989.

Gery, M. W., Whitten, G. Z., Killus, J. P., and Dodge, M. C.: A photochemical kinetics mechanism for urban and regional scale computer modeling, J. Geophys. Res., 94, 12925–12956, 1989.

Goldstein, A. H., Millet, D. B., McKay, M., Jaegle, L., Horowitz, L., Cooper, O., Hudman, R., Jacob, D. J., Oltmans, S., and Clarke, A.: Impact of Asian emissions on observations at Trinidad Head, California, during ITCT 2K2, J. Geophys. Res, 109, 1–13, 2004.

Guelle, W., Balkanski, Y. J., Schulz, M., Dulac, F., and Monfray, P.: Wet deposition in a global size-dependent aerosol transport model. 1. Comparison of a 1 year 210Pb simulation with ground measurements, J. Geophys. Res., 103, 11429–11445, 1998.

Hansen, A. D. A., Rosen, H., and Novakov, T.: The aethalometer – An instrument for the real-time measurement of optical absorption by aerosol particles, Sci. Total Environ., 36, 191–196, 1984.

Hansen, A. D. A., Bodhaine, B. A., Dutton, E. G., and Schnell, R. C.: Aerosol black carbon measurements at the South Pole: initial results, 1986–1987, Geophys. Res. Lett., 15, 1193–1196, 1988.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, M., Bauer, S., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., Wielicki, B., Wong, T., Yao, M., and Zhang, S.: Efficacy of climate forcings, J. Geophys. Res., 110, 1–45, 2005.

Hara, K., Osada, K., Yabuki, M., Hayashi, M., Yamanouchi, T., Shiobara, M., and Wada, M.: Measurement of black carbon at Syowa station, Antarctica: seasonal variation, transport processes and pathways, Atmos. Chem. Phys. Discuss., 8, 9883–9929, doi:10.5194/acpd-8-9883-2008, 2008.

Haywood, J. M. and Shine, K. P.: The effect of anthropogenic sulfate and soot aerosol on the clear sky planetary radiation budget, Geophys. Res. Lett., 22, 603–606, 1995.

Hertel, O., Berkowicz, R., Christensen, J., and Hov, O.: Test of two numerical schemes for use in atmospheric transport-chemistry models, Atmos. Environ., 27, 2591–2611, 1993.

Houweling, S., Dentener, F., and Lelieveld, J.: The impact of non-methane hydrocarbon compounds on tropospheric photochemistry, J. Geophys. Res., 103, 10673–10696, 1998.

Huffman, H. D.: Comparison of the light absorption coefficient and carbon measures for remote aerosols: An independent analysis of data from the improve network – I, Atmos. Environ., 30, 73–83, 1996.

Jacobson, M. Z.: Global direct radiative forcing due to multicomponent anthropogenic and natural aerosols, J. Geophys. Res., 106, 1551–1568, 2001.

Jennings, S. G. and Pinnick, R. G.: Relationaships between visible extinction, absorption and mass concentration of carbonaceous smokes, Atmos. Environ., 14, 1123–1129, 1980.

Jennings, S. G., Geever, M., McGovern, F. M., Francis, J., Spain, T. G., and Donaghy, T.: Microphysical and physico-chemical characterization of atmospheric marine and continental aerosol at Mace Head, Atmos. Environ., 31, 2795–2808, 1997.

Jeuken, A., Veefkind, J. P., Dentener, F., Metzger, S., and Robles Gonzalez, C.: Simulation of the aerosol optical depth over Europe for August 1997 and a comparison with observations, J. Geophys. Res., 106, 28295–28311, 2001.

Junker, C., Jennings, S. G., and Cachier, H.: Aerosol light absorption in the North Atlantic: trends and seasonal characteristics during the period 1989 to 2003, Atmos. Chem. Phys., 6, 1913–1925, doi:10.5194/acp-6-1913-2006, 2006.

Kim, E., Hopke, P. K., Kenski, D. M., and Koerber, M.: Sources of fine particles in a rural Midwestern US area, Environ. Sci. Technol., 39, 4953–4960, 2005.

Kirchstetter, T. W., Novakov, T., and Hobbs, P. V.: Evidence that the spectral dependence of light absorption by aerosols is affected by organic carbon, J. Geophys. Res., 109, 21201–21212, 2004.

Kleefeld, S., Hoffer, A., Krivácsy, Z., and Jennings, S. G.: Importance of organic and black carbon in atmospheric aerosols at Mace Head, on the West Coast of Ireland (53°19 N, 9°54 W), Atmos. Environ., 36, 4479–4490, 2002.

Koch, D. and Hansen, J.: Distant origins of Arctic black carbon: A Goddard Institute for Space Studies ModelE experiment, J. Geophys. Res., 110, 1–14, 2005.

Koch, D., Schulz, M., Kinne, S., McNaughton, C., Spackman, J. R., Balkanski, Y., Bauer, S., Berntsen, T., Bond, T. C., Boucher, O., Chin, M., Clarke, A., De Luca, N., Dentener, F., Diehl, T., Dubovik, O., Easter, R., Fahey, D. W., Feichter, J., Fillmore, D., Freitag, S., Ghan, S., Ginoux, P., Gong, S., Horowitz, L., Iversen, T., Kirkevåg, A., Klimont, Z., Kondo, Y., Krol, M., Liu, X., Miller, R., Montanaro, V., Moteki, N., Myhre, G., Penner, J. E., Perlwitz, J., Pitari, G., Reddy, S., Sahu, L., Sakamoto, H., Schuster, G., Schwarz, J. P., Seland, Ø., Stier, P., Takegawa, N., Takemura, T., Textor, C., van Aardenne, J. A., and Zhao, Y.: Evaluation of black carbon estimations in global aerosol models, Atmos. Chem. Phys., 9, 9001–9026, doi:10.5194/acp-9-9001-2009, 2009.

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, Atmos. Chem. Phys., 5, 417–432, doi:10.5194/acp-5-417-2005, 2005.

Kulmala, M., Asmi, A., Lappalainen, H. K., Carslaw, K. S., Pöschl, U., Baltensperger, U., Hov, Ø., Brenquier, J.-L., Pandis, S. N., Facchini, M. C., Hansson, H.-C., Wiedensohler, A., and O'Dowd, C. D.: Introduction: European Integrated Project on Aerosol Cloud Climate and Air Quality interactions (EUCAARI) - integrating aerosol research from nano to global scales, Atmos. Chem. Phys., 9, 2825–2841, doi:10.5194/acp-9-2825-2009, 2009.

Lavanchy, V. M. H., Gaggeler, H. W., Nyeki, S., and Baltensperger, U.: Elemental carbon (EC) and black carbon (BC) measurements with a thermal method and an aethalometer at the high-alpine research station Jungfraujoch, Atmos. Environ., 33, 2759–2769, 1999.

Lazaridis, M., Semb, A., Larssen, S., Hjellbrekke, A. G., Hov, O., Hanssen, J. E., Schaug, J., and Torseth, K.: Measurements of particulate matter within the framework of the European Monitoring and Evaluation Programme (EMEP) I. First results, Sci. Total Environ., 285, 209–235, 2002.

Liousse, C., Cachier, H., and Jennings, S. G.: Optical and thermal measurements of black carbon aerosol content in different environments: Variation of the specific attenuation cross-section, sigma, Atmos. Environ., 27A, 1203–1211, 1993.

Malm, W. C., Sisler, J. F., Huffman, D., Eldred, R. A., and Cahill, T. A.: Spatial and seasonal trends in particle concentration and optical extinction in the United States, J. Geophys. Res., 99, 1347–1370, 1994.

Malm, W. C., Pitchford, M. L., McDade, C., and Ashbaugh, L. L.: Coarse particle speciation at selected locations in the rural continental United States, Atmos. Environ., 41, 2225–2239, 2007.

Menon, S., Hansen, J., Nazarenko, L., and Luo, Y.: Climate effects of black carbon aerosols in China and India, Science, 297, 2250–2253, 2002.

Nyeki, S., Baltensperger, U., Colbeck, I., Jost, D. T., Weingartner, E., and Gaggeler, H. W.: The Jungfraujoch high-alpine research station (3454 m) as a background clean continental site for the measurement of aerosol parameters, J. Geophys. Res., 103, 6097–6107, 1998.

Ogren, J. A., Andrews, E., Dutton, E. G., Sheridan, P. J., and Jefferson, A.: Aerosol optical properties at Trinidad Head – ITCT and beyond, in: American Gephysical Union, Fall meeting 2002, 2002.

Perry, K. D., Cliff, S. S., and Jimenez-Cruz, M. P.: Evidence for hygroscopic mineral dust particles from the Intercontinental Transport and Chemical Transformation Experiment, J. Geophys. Res., 109, 1–9, 2004.

Peters, W., Krol, M. C., Dlugokencky, E. J., Dentener, F. J., Bergamaschi, P., Dutton, G., Velthoven, P. V., Miller, J. B., Bruhwiler, L., and Tans, P. P.: Toward regional-scale modeling using the two-way nested global model TM5: Characterization of transport using SF6, J. Geophys. Res., 109, D19314, doi:10.1029/2004JD005020, 2004.

Putaud, J. P., van Dingenen, R., and Raes, F.: Submicron aerosol mass balance at urban and semirural sites in the Milan area (Italy), J. Geophys. Res., 107, 8198, doi:10.1029/2000JD000111, 2002.

Ramanathan, V. and Carmichael, G.: Global and regional climate changes due to black carbon, Nat. Geosci., 1, 221–227, 2008.

Ramanathan, V., Li, F., Ramana, M. V., Praveen, P. S., Kim, D., Corrigan, C. E., Nguyen, H., Stone, E. A., Schauer, J. J., Carmichael, G. R., Adhikary, B., and Yoon, S. C.: Atmospheric brown clouds: Hemispherical and regional variations in long-range transport, absorption, and radiative forcing, J. Geophys. Res., 112, D22S21, doi:10.1029/2006JD008124, 2007.

Rosen, H. and Hansen, A. D. A.: Role of combustion-generated carbon particles in the absorption of solar radiation in the Arctic haze, Geophys. Res. Lett., 11, 461–464, 1984.

Seigneur, C., Pun, B., Pai, P., Louis, J. F., Solomon, P., Emery, C., Morris, R., Zahniser, M., Worsnop, D., Koutrakis, P., White, W., and Tombach, I.: Guidance for the performance evaluation of three-dimensional air quality modeling systems for particulate matter and visibility, J. Air Waste Manage., 50, 588–599, 2000.

Sharma, S., Brook, J. R., Cachier, H., Chow, J., Gaudenzi, A., and Lu, G.: Light absorption and thermal measurements of black carbon in different regions of Canada, J. Geophys. Res., 107, 4771, doi:10.1029/2002JD002496, 2002.

Sharma, S., Lavoué, D., Chachier, H., Barrie, L. A., and Gong, S. L.: Long-term trends of the black carbon concentrations in the Canadian Arctic, J. Geophys. Res., 109, 15201–15210, 2004.

Sharma, S., Andrews, E., Barrie, L. A., Ogren, J. A., and Lavoué, D.: Variations and sources of the equivalent black carbon in the high Arctic revealed by long-term observations at Alert and Barrow: 1989–2003, J. Geophys. Res., 111, D14208, doi:10.1029/2005JD006581, 2006.

Sheridan, P. J., Delene, D. J., and Ogren, J. A.: In situ measurement of aerosol light absorption and single-scattering albedo at the NSA and SGP CART sites, in: Tenth ARM Science Meeting Proceedings, 2000.

Snyder, D. C. and Schauer, J. J.: An inter-comparison of two black carbon aerosol instruments and a semi-continuous elemental carbon instrument in the urban environment, Aerosol Sci. Technol., 41, 463–474, 2007.

Sokolik, I. N. and Toon, O. B.: Incorporation of mineralogical composition into models of the radiative properties of mineral aerosol from UV to IR wavelengths, J. Geophys. Res., 104, 9423–9444, 1999.

Roelofs, G.-J., ten Brink, H., Kiendler-Scharr, A., de Leeuw, G., Mensah, A., Minikin, A., and Otjes, R.: Evaluation of simulated aerosol properties with the aerosol-climate model ECHAM5-HAM using observations from the IMPACT field campaign, Atmos. Chem. Phys., 10, 7709–7722, doi:10.5194/acp-10-7709-2010, 2010.

Stohl, A., Andrews, E., Burkhart, J. F., Forster, C., Herber, A., Hoch, S. W., Kowal, D., Lunder, C., Mefford, T., Ogren, J. A., Sharma, S., Spichtinger, N., Stebel, K., Stone, R., Ström, J., Tørseth, K., Wehrli, C., and Yttri, K. E.: Pan-Arctic enhancements of light absorbing aerosol concentrations due to North American boreal forest fires during summer 2004, J. Geophys. Res., 111, D22214, doi:10.1029/2006JD007216, 2006.

Tegen, I., Hollrig, P., Chin, M., Fung, I., Jacob, D., and Penner, J.: Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results, J. Geophys. Res., 102, 23895–23915, 1997.

Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Feichter, H., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Horowitz, L., Huang, P., Isaksen, I., Iversen, I., Kloster, S., Koch, D., Kirkevåg, A., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Liu, X., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, Ø., Stier, P., Takemura, T., and Tie, X.: Analysis and quantification of the diversities of aerosol life cycles within AeroCom, Atmos. Chem. Phys., 6, 1777–1813, doi:10.5194/acp-6-1777-2006, 2006.

van der Werf, G. R., Randerson, J. T., Collatz, G. J., Giglio, L., Kasibhatla, P. S., Arellano, A. F., Olsen, S. C., and Kasischke, E. S.: Continental scale partitioning of fire emissions during the 1997 to 2001 El Niño/La Niña period, Science, 303, 73–76, 2004.

Vignati, E., Wilson, J., and Stier, P. M.: An efficient size-resolved aerosol microphysics module for large-scale aerosol transport models, J. Geophys. Res., 109, 1–17, 2004.

Vignati, E., Facchini, M. C., Rinaldi, M., Scannell, C., Ceburnis, D., Sciare, J., Kanakidou, M., Myriokefalitakis, S., Dentener, F., and O'Dowd, C. D.: Global scale emission and distribution of sea spray aerosol: Sea–salt and organic enrichment, Atmos. Environ., 44, 670–677, 2010a.

Vignati, E., Karl, M., Krol, M., Wilson, J., Stier, P., and Cavalli, F.: Sources of uncertainties in modelling black carbon at the global scale, Atmos. Chem. Phys., 10, 2595–2611, doi:10.5194/acp-10-2595-2010, 2010b.

Watson, J. G., Chow, J., and Cheng, L.-W. A.: Summary of Organic and Elemental Carbon/Black Carbon Analysis Methods and Intercomparisons, Aerosol and Air Qual. Res., 5, 65–102, 2005.

Weingartner, E., Saathoff, H., Schnaiter, M., Streit, N., Bitnar, B., and Baltensperger, U.: Absorption of light by soot particles: Determination of the absorption coefficient by means of aethalometers, J. Aerosol Sci., 34, 1445–1463, 2003.

Wilson, J., Cuvelier, C., and Raes, F.: A modeling study of global mixed aerosol fields, J. Geophys. Res., 106, 34081–34108, 2001.

Yang, M., Howell, S. G., Zhuang, J., and Huebert, B. J.: Attribution of aerosol light absorption to black carbon, brown carbon, and dust in China - interpretations of atmospheric measurements during EAST-AIRE, Atmos. Chem. Phys., 9, 2035–2050, doi:10.5194/acp-9-2035-2009, 2009.

Yttri, K. E., Aas, W., Bjerke, A., Cape, J. N., Cavalli, F., Ceburnis, D., Dye, C., Emblico, L., Facchini, M. C., Forster, C., Hanssen, J. E., Hansson, H. C., Jennings, S. G., Maenhaut, W., Putaud, J. P., and Trseth, K.: Elemental and organic carbon in PM_{10}: a one year measurement campaign within the European Monitoring and Evaluation Programme EMEP, Atmos. Chem. Phys., 7, 5711–5725, doi:10.5194/acp-7-5711-2007, 2007.

Yu, S., Eder, B., Dennis, R., Chu, S.-H., and Schwartz, E.: New unbiased symmetric metrics for evaluation of air quality model, Atmos. Sci. Lett., 7, 26–34, 2006.

Zhang, Y., Liu, P., Pun, B., and Seigneur, C.: A comprehensive performance evaluation of MM5-CMAQ for the Summer 1999 Southern Oxidants Study episode-Part I: Evaluation protocols, databases, and meteorological predictions, Atmos. Environ., 40, 4825–4838, 2006.

Zhao, T. L., Gong, S. L., Zhang, X. Y., and Jaffe, D. A.: Asian dust storm influence on North American ambient PM levels: observational evidence and controlling factors, Atmos. Chem. Phys., 8, 2717–2728, doi:10.5194/acp-8-2717-2008, 2008.