Atmos. Chem. Phys., 11, 7781–7816, 2011
www.atmos-chem-phys.net/11/7781/2011/
doi:10.5194/acp-11-7781-2011
© Author(s) 2011. CC Attribution 3.0 License.

Atmospheric Chemistry and Physics

Global dust model intercomparison in AeroCom phase I

N. Huneeus[1], M. Schulz[1,2], Y. Balkanski[1], J. Griesfeller[1,2], J. Prospero[3], S. Kinne[4], S. Bauer[5,6], O. Boucher[8,*], M. Chin[9], F. Dentener[10], T. Diehl[11,12], R. Easter[13], D. Fillmore[14], S. Ghan[13], P. Ginoux[15], A. Grini[16,17], L. Horowitz[15], D. Koch[5,6,7], M. C. Krol[18,19], W. Landing[20], X. Liu[13,21], N. Mahowald[22], R. Miller[6,23], J.-J. Morcrette[24], G. Myhre[16,25], J. Penner[21], J. Perlwitz[6,23], P. Stier[26], T. Takemura[27], and C. S. Zender[28]

[1]Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, IPSL, Gif-sur-Yvette, France
[2]Meteorological Institut, Oslo, Norway
[3]Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL., USA
[4]Max-Planck-Institut für Meteorologie, Hamburg, Germany
[5]The Earth Institute, Columbia University, New York, USA
[6]NASA Goddard Institute for Space Studies, New York, NY, USA
[7]US Department of Energy, Washington, DC, USA
[8]Met Office, Hadley Centre, Exeter, UK
[9]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[10]European Comission, Joint Research Centre, Institute for Environment and Sustainability, Italy
[11]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
[12]Universities Space Research Association, Columbia, Maryland, USA
[13]Pacific Northwest National Laboratory, Richland, WA, USA
[14]NCAR, Boulder, Colorado, USA
[15]NOAA, Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA
[16]Department of Geosciences, University of Oslo, Oslo, Norway
[17]Kongsberg Oil & Gas Technologies, Norway
[18]Utrecht University, Institute for Marine and Atmospheric Research, Utrecht, The Netherlands
[19]Wageningen University, Meteorology and Air Quality, Wageningen, The Netherlands
[20]Departement of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, FL, USA
[21]Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, MI, USA
[22]Departement of Earth and Atmospheric Sciences, Cornell University, Ithaca, New York, USA
[23]Department of Applied Physics and Applied Mathematics, Columbia University, New York, USA
[24]European Centre for Medium-Range Weather Forecasts, Reading, UK
[25]Center for International Climate and Environmental Research – Oslo (CICERO) Oslo, Norway
[26]Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, UK
[27]Research Institute for Applied Mechanics, Kyushu University, Fukuoka, Japan
[28]Department of Earth System Science, University of California, Irvine, USA
[*]now at: Laboratoire de Météorologie Dynamique, IPSL, CNRS/UPMC, Paris, France

Received: 5 August 2010 – Published in Atmos. Chem. Phys. Discuss.: 12 October 2010
Revised: 29 June 2011 – Accepted: 9 July 2011 – Published: 3 August 2011

Abstract. This study presents the results of a broad intercomparison of a total of 15 global aerosol models within the AeroCom project. Each model is compared to observations related to desert dust aerosols, their direct radiative effect, and their impact on the biogeochemical cycle, i.e., aerosol optical depth (AOD) and dust deposition. Additional comparisons to Angström exponent (AE), coarse mode AOD and dust surface concentrations are included to extend the assessment of model performance and to identify common biases present in models. These data comprise a benchmark dataset that is proposed for model inspection and future dust model development. There are large differences among the global models that simulate the dust cycle and its impact on climate. In general, models simulate the climatology of vertically integrated parameters (AOD and AE) within a factor of two

Correspondence to: N. Huneeus
(nicolas.huneeus@lsce.ipsl.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

whereas the total deposition and surface concentration are reproduced within a factor of 10. In addition, smaller mean normalized bias and root mean square errors are obtained for the climatology of AOD and AE than for total deposition and surface concentration. Characteristics of the datasets used and their uncertainties may influence these differences. Large uncertainties still exist with respect to the deposition fluxes in the southern oceans. Further measurements and model studies are necessary to assess the general model performance to reproduce dust deposition in ocean regions sensible to iron contributions. Models overestimate the wet deposition in regions dominated by dry deposition. They generally simulate more realistic surface concentration at stations downwind of the main sources than at remote ones. Most models simulate the gradient in AOD and AE between the different dusty regions. However the seasonality and magnitude of both variables is better simulated at African stations than Middle East ones. The models simulate the offshore transport of West Africa throughout the year but they overestimate the AOD and they transport too fine particles. The models also reproduce the dust transport across the Atlantic in the summer in terms of both AOD and AE but not so well in winter-spring nor the southward displacement of the dust cloud that is responsible of the dust transport into South America. Based on the dependency of AOD on aerosol burden and size distribution we use model bias with respect to AOD and AE to infer the bias of the dust emissions in Africa and the Middle East. According to this analysis we suggest that a range of possible emissions for North Africa is 400 to 2200 Tg yr^{-1} and in the Middle East 26 to 526 Tg yr^{-1}.

1 Introduction

Desert dust plays an important role in the climate system. Models suggest that dust is one of the main contributors to the global aerosol burden (Textor et al., 2006) and has a large impact on Earth's radiative budget due to the absorption, scattering and emissions of solar and infrared radiation (Sokolik et al., 2001; Tegen, 2003; Balkanski et al., 2007). The deposition of desert dust to the ocean is an important source of iron in high-nutrient-low-chlorophyll (HNLC) regions (Mahowald et al., 2009). This iron contribution may be crucial for the ocean uptake of atmospheric CO_2 through its role as an important nutrient for phytoplankton growth (Jickells et al., 2005; Aumont et al., 2008; Tagliabue et al., 2009). Dust also plays a significant role in tropospheric chemistry mainly through heterogeneous uptake of reactive gases such as nitric acid (Bian and Zender, 2003; Liao et al., 2003; Bauer et al., 2004) and heterogeneous reactions with sulfur dioxide (Bauer and Koch, 2005). Furthermore, mineral aerosols are important for air quality assessments through their impact on visibility and concentration levels of particulate matter (Kim et al., 2001; Ozer et al., 2007; Jimenez-Guerrero et al., 2008). Links between the occurrence of meningitis epidemics in Africa and dust have been suggested (Thomson et al., 2006). Impacts on climate and air quality are intimately coupled (Denman et al., 2007).

Many global models simulate dust emissions, its transport and deposition in a coherent manner (e.g. Guelle et al., 2000; Reddy et al., 2005b; Ginoux et al., 2001; Woodage et al., 2010). A large diversity has been documented between models in terms of e.g. dust burden and aerosol optical depth introducing uncertainties in estimating the direct radiative effect, and even more difficult the anthropogenic component of it (Zender et al., 2004; Textor et al., 2006; Forster et al., 2007). On the other hand, inter-model differences in simulated dust emission and deposition fluxes make estimating the impact of dust on ocean CO_2 uptake in HNLC regions difficult (Textor et al., 2006; Tagliabue et al., 2009).

An exhaustive comparison of different models with each other and against observations can reveal weaknesses of individual models and provide an assessment of uncertainties in simulating the dust cycle. Uno et al. (2006) compared multiple regional dust models over Asia in connection to specific dust events. They concluded that even though all models were able to predict the onset and ending of a dust event and were able to reproduce surface measurements, large differences existed among them in processes such as emissions, transport and deposition. Todd et al. (2008) conducted an intercomparison with five regional models for a 3-day dust event over the Bodélé depression. The analyzed model quantities presented a similar degree of uncertainty as reported by Uno et al. (2006). Kinne et al. (2003) compared aerosol properties from seven global models to satellite and ground data. The largest differences among models were found near expected source regions of biomass burning and dust. Further global model intercomparisons have been conducted within the AeroCom project (http://nansen.ipsl.jussieu.fr/AEROCOM/). Textor et al. (2006) conducted an intercomparison between global models of the life cycle of the main aerosol species. Large differences (diversity) were found in emissions, sinks, burdens and spatial distribution for the different aerosol species simulated. These diversities reveal uncertainties in simulating aerosol processes that have large consequences for estimating the radiative impact of dust. However, no comparisons against observations were made in that study. Kinne et al. (2006) extended the study of Kinne et al. (2003) and compared the aerosol properties from all AeroCom models to satellite and ground data. None of these AeroCom studies however, focused exclusively on dust particles. Tegen (2003) compared the dust cycle simulated by two global dust models to satellite climatology of TOMS aerosol index (AI). Zender et al. (2004) compared the emission fluxes and burdens for different models. Prospero et al. (2010) conducted a more exhaustive intercomparison, comparing and evaluating the temporal and spatial variability of African dust deposition in Florida simulated by models within the AeroCom initiative. The comparison shows that in

general models reproduce the seasonal variability but most yield weak summer maxima.

This work represents a broader dust model intercomparison. Global dust models within AeroCom are compared against each other and against different datasets. By using one homogeneous model data compilation (model versions in AeroCom and documented by Textor et al., 2006) we demonstrate the use of a benchmark data test set for across model inspections and for future developments of dust models. We compare each model to observations focusing on variables related to the uncertainties in the estimation of the direct radiative effect and the dust impact on the ocean biogeochemical cycle, i.e. aerosol optical properties and dust deposition as well as surface concentration. The article is structured as follows. We start by describing the data used in the validation and the different models considered in this work (Sect. 2). The results are presented in Sect. 3 while the discussion of these results is given in Sect. 4. Finally in Sect. 5 the conclusions of this work are presented.

2 Data and models

We evaluate the models described in Sect. 2.4 against in-situ measurements of dust deposition (Sect. 2.1) and surface concentration (Sect. 2.2) as well as retrievals of aerosol optical depth (AOD, Sect. 2.3) and Angström exponent (AE, Sect. 2.3). A brief description of each of these datasets follows together with a brief description of the AeroCom models used in this work.

2.1 Dust deposition

Deposition at sites remote from sources serves as a powerful constraint on the overall global dust budget. Total deposition fluxes are most useful when accumulated over long time periods. In this way direct dust deposition data have been used in the validation of global dust models.

We first use three compilations giving deposition fluxes over land. We use the measured deposition fluxes given in Ginoux et al. (2001) based partly upon measurements taken during the SEAREX campaign (Prospero et al., 1989; capital letters in Fig. 1). Only those data corresponding to actual measurements were considered. Most sites are located in the Northern Hemisphere and far away from source regions. The measured values range from 450 [$g\,m^{-2}\,yr^{-1}$] in the Taklimakan desert to 0.09 [$g\,m^{-2}\,yr^{-1}$] in the equatorial Pacific; measurement periods vary according to the site. Mahowald et al. (2009) present a compilation with a total of 28 sites measuring iron and/or dust deposition, mostly in the last two decades (non-italic numbers in Fig. 1). We assume a 3.5 % iron content in dust to infer dust deposition fluxes from iron deposition. This value is the average iron content of the Earth's crust and is widely used in studies deriving iron inputs to the ocean from dust aerosols (Mahowald et al., 2005; Hand et al., 2004). The iron content in soils varies according to the source region but studies suggest that uncertainties in dust deposition and iron solubility are more important to understand than the variability of iron content in different source regions (Mahowald et al., 2005). In addition we use deposition fluxes derived from ice core data (lower case letters in Fig. 1). These depositions have proven to be accurate to represent the current climate (Mahowald et al., 1999).

We then use deposition fluxes measured from sediment traps and collected in the Dust Indicators and Records in Terrestrial and Marine Paleoenvironments (DIRTMAP) database (Tegen et al., 2002; Kohfeld and Harrison, 2001; italic numbers in Fig. 1). We follow Tegen et al. (2002) and only use those stations with deployment period larger than 50 days and sites without contamination from suspected fluvial inputs or hemipelagic reworking. This database contains a set of comparable deposition fluxes providing a picture of the gradients in the intensity of the dust deposition to the Atlantic Ocean and the Arabian Sea. In addition, we also follow Tegen et al. (2002) and Mahowald et al. (2009) and do not use DIRTMAP deposition data derived from marine sediment cores because they represent the integrated dust flux to the ocean over a time span of hundreds to possibly thousands of years and are thus inadequate to be used in the evaluation of simulation of the dust cycle for specific years (Tegen et al., 2002).

The total deposition data used in this study comes from 84 sites with yearly dust deposition fluxes that are not coincident in time with the model simulated year (Table S1 in the Supplement). Model yearly deposition fluxes were computed using all days. Except for the ice core data, the sites have been grouped regionally. To facilitate the comparison with model data, each of these regions is identified with a different colour in Fig. 1. Given the characteristics described above, we suggest that these datasets represent to first order a modern or present-day climatology of dust deposition observations. However, some of these measurements do not cover a sufficiently long period to qualify as "climatological" in a strict sense. The impact of this assumption on the model evaluation will be considered in the discussion (Sect. 4). Deposition data from the same locations were averaged in order to provide one climatological data set.

Dust particles are efficiently removed by wet scavenging, especially over the open ocean (Prospero et al., 2010; Hand et al., 2004; Gao et al., 2003). To test the wet deposition simulated by the models we compare simulated deposition against data from the Florida Atmospheric Mercury Study (FAMS) network (Prospero et al., 2010) and from a compilation of estimates of the fraction of wet deposition (Mahowald et al., 2011). For the former a total of nine stations measured wet and total deposition during almost three years (April 1994 till end of 1996). These data have already been used to evaluate some AeroCom models in Prospero et al. (2010). Nevertheless, we include this dataset to extend the comparison to the expanded set of AeroCom models.

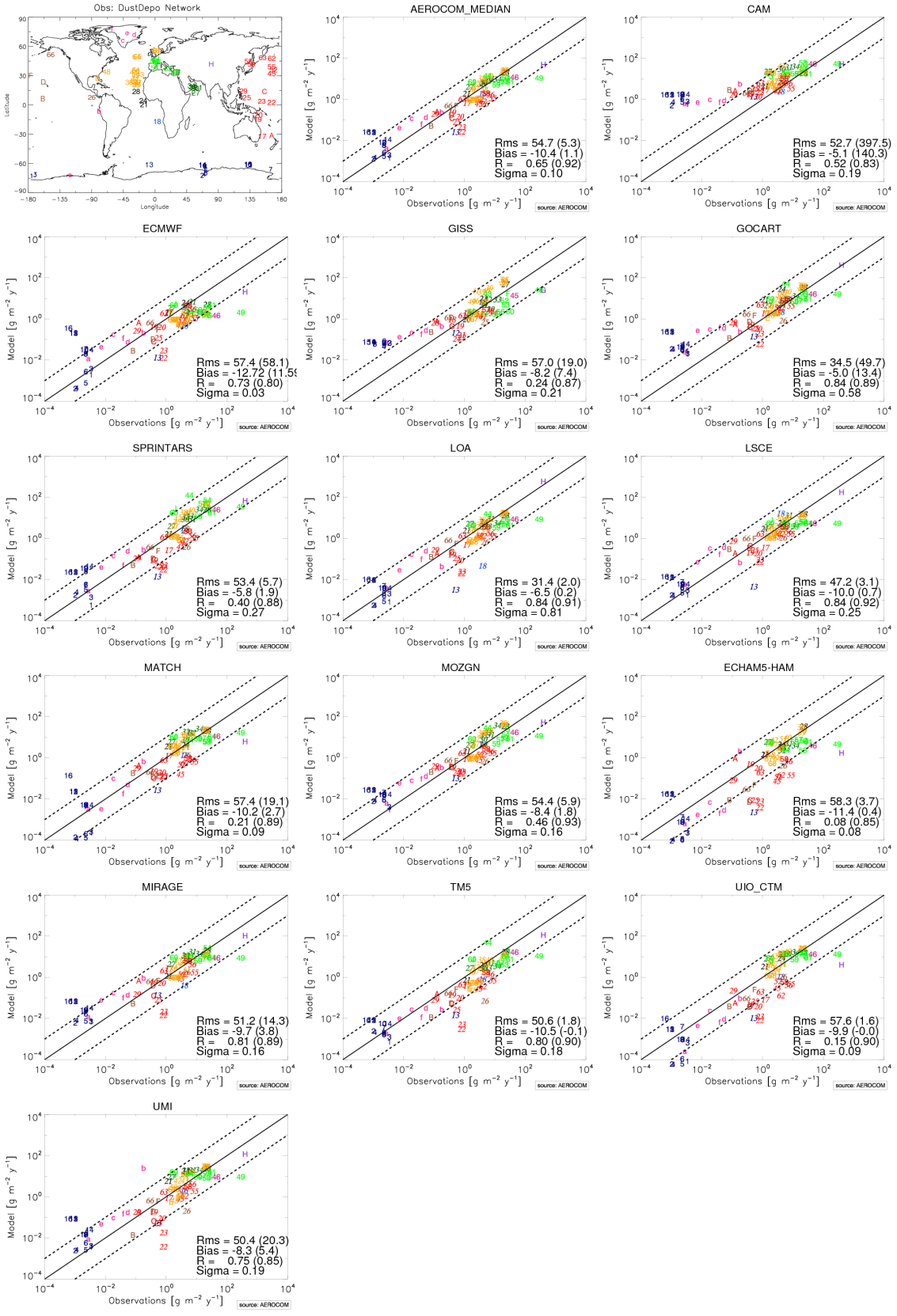

Fig. 1. Measured yearly deposition fluxes versus modeled ones; units are g m^{-2} yr^{-1}. Location for each data point in the scatter plot is given in the upper left subfigure. Number and letters are coloured regionally for West/East Pacific (red/brown), North/Tropical/South Atlantic (orange/black/light-blue), Middle East/Asia/Europe (violet/purple/light green), Indian/Southern Ocean (dark green/dark blue) and pink ice core data in Greenland, South America and Antarctica. Data from Ginoux et al. (2001)/Mahowald et al. (2009)/DIRTMAP/Mahowald et al. (1999) are indicated by letters/non-italic numbers/italic numbers/lower-case letters. Root mean square error (RMS), bias, ratio of modeled and observed standard deviation (sigma) and correlation (R) are indicated for each model in the lower right part of the scatterplot. Mean normalized bias and normalized root mean square error are given in parenthesis next to RMS and mean bias, respectively. The correlation with respect to the logarithm of the model and of the observation is also given in parenthesis next to R. Black continuous line is the 1:1 line whereas the black dotted lines correspond to the 10:1 and 1:10 lines. Names and locations for each selected station are given in Table S1 in the Supplement.

In general the cut-off size of the deposition measurements is not provided in model evaluation studies and is difficult to find. This cut-off size depends on the instrument used and it can be as high as several tenths of micrometers (Goossens and Rajot, 2008). No size distribution data of the deposited dust for the period of measurements are provided. However Reid et al. (2003) and Li-Jones and Prospero (1998) reported measurement diameters of Saharan dust particles mainly smaller than 10 μm across the Atlantic Ocean on the eastern limit of the Caribbean Sea. Since most of our deposition data corresponds to measurements in remote regions and most models only simulate dust particles up to 10 μm, we do not consider the cut-off size as a significant source of bias in the results.

2.2 Surface concentration

Surface concentrations are an alternative mean to evaluate the transport and dispersion of simulated dust. We compare the AeroCom models with monthly dust concentration measurements taken at 20 sites managed by the Rosenstiel School of Marine and Atmospheric Science at the University of Miami (Prospero et al., 1989; Prospero, 1996; Arimoto et al., 1995). The measurements taken in the Pacific Ocean are from the sea/air exchange (SEAREX) program (Prospero et al., 1989) whereas the measurements from the northern Atlantic are from the Atmosphere-Ocean chemistry experiment (AEROCE, Arimoto et al., 1995). Both experiments were designed to study the large-scale spatial and temporal variability of aerosols. Most measuring sites are located far downwind of dust emission sources (Fig. 2). A list of the stations and their location is given in Table S2 in the Supplement. The dust concentrations are derived from measured aluminium concentrations assuming an Al content of 8 % in soil dust (Prospero, 1999) or from the weights of filter samples ashed at 500 °C after extracting soluble components with water. This database has been largely used for the evaluation of dust models (e.g. Ginoux et al., 2001; Cheng et al., 2008; Tegen et al., 2002) and in the reports of the Intergovernmental Panel of Climate Change (IPCC) of 2001 and 2007. The measurements were taken for the most part in the 1980s and 1990s with varying measurement periods at each station. We extend this data set with monthly dust concentrations at Rukomechi, Zimbabwe (Maenhaut et al., 2000a; Nyanganyura et al., 2007) and Jabiru, Australia (Maenhaut et al., 2000b; Vanderzalm et al., 2003). The primary goal of these measurements was to study aerosol composition in Rukomechi and the impact of biomass burning in northern Australia. Nevertheless, we include these data because dust was one of the species measured during these long term measurements.

We have separated the sites in three distinctive groups according to the range of measured data. The first group corresponds to stations with monthly mean surface concentrations lower than 1 μg m^{-3} throughout the year (LOW). These stations are located in the Antarctica and in the Pacific Ocean below 20° N far from any dust sources. Orange numbers and dots illustrate them in Fig. 2. The second group (in violet in Fig. 2) corresponds to stations under the influence of minor dust sources of the Southern Hemisphere or remote sites in the Northern Hemisphere (MEDIUM). Finally, the third group corresponds to locations downwind of African and Asian dust sources, presented by blue numbers and dots in Fig. 2 (HIGH). In each of these groups the stations are ordered from south to north. A list of the stations with their location, identifier used in Fig. 2 and attributed data range group is given in Table S2 of the Supplement. The simulated monthly averages of surface concentrations for all models are computed using all days.

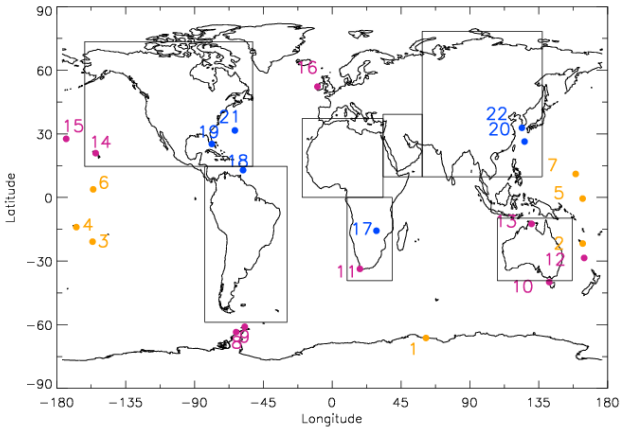

Fig. 2. Network of stations measuring surface concentration (Sect. 2.2). Stations are grouped according to the regime of measured data into remote stations (orange), stations under the influence of minor dust sources of the Southern Hemisphere or remote sites in the Northern Hemisphere (violet) and finally locations directly downwind of African and Asian dust source (blue). Stations within each group are numbered from south to north. Names and locations for each selected station are given in Table S2 in the Supplement. Rectangles illustrate regions defined to compute the emissions presented in Table 5.

In addition, we complement the monthly averages with the data set of surface concentrations presented in Mahowald et al. (2009). These data correspond to measurements taken mostly during cruises but include also long term measuring stations. The measurements taken during cruise campaigns will be compared to yearly averages even though they represent short-term data. Mahowald et al. (2009) show that 30–90 % of the annually averaged deposited dust is seen on a few days (5 %). In order to account for the error of comparing model yearly averaged surface concentration with short-term measurements we follow Mahowald et al. (2008) and show the range of values representing the median 66 % of the daily averaged model concentration as an error bar on the model and annual mean (vertical dashed line) for each cruise data. Because the long-range transport of dust is an important

attribute and we do not have monthly mean values at many locations, we include this cruise data with large uncertainty bars until better data is available.

We consider all the above described data sets as "climatology" even though they do not cover a long enough period to be termed climatology in a strict sense.

We also use measurements from the year 2000 at Barbados station and at Miami consistent with the model output from the AeroCom models used in this study and presented below (Sect. 2.4). This is the most extensive long-term record of surface dust concentration available. Concentrations have been measured under on-shore wind conditions almost continuously since 1965 in an equivalent manner as described above (Prospero, 1999; Prospero and Lamb, 2003). The Barbados data have been used to study the long-range transport from African dust over the Atlantic and the factors influencing its variability (Prospero and Nees, 1986; Prospero and Lamb, 2003; Chiapello et al., 2005). We will compare these measurements to the climatological cycle described above and evaluate how representative the climatology is from the year 2000.

The instruments used to measure surface concentrations efficiently captured particles below 40 μm in stations managed by the Rosenstiel School of Marine and Atmospheric Science at the University of Miami. While this cut-off limit could be critical for model evaluation close to sources (provided coarse dust particles are present) it is safe to assume that it is less important in remote stations. Measurements on the eastern limit of the Caribbean Sea reported diameters of Saharan dust particles mainly smaller than 10 μm (Reid et al., 2003; Li-Jones and Prospero, 1998). Furthermore most models considered in this study only simulate dust particles up to 10 μm.

2.3 Aerosol optical depth and Angström exponent

The widespread deployment of sun photometers in the last ten years has provided very reliable global information about dust, although limited to times when dust dominates the AOD. When full inversions of multiple-angle sky observations are available, coarse mode AOD may provide a better estimate of dust optical depth. Note that the measurements are biased towards daytime, clear-sky conditions. AOD retrievals may also miss very dusty situations because of cloud discrimination problems. The AErosol RObotic NETwork (AERONET) is a global network of photometers that delivers numerical data to monitor and characterize the aerosols in a regional and/or global scale. The network has more than 300 stations distributed in the world measuring aerosol in both remote and polluted areas (Holben et al., 1998, 2001). We use here AERONET total AOD and coarse mode AOD at 550 nm and Angström exponent (AE). Typically, the uncertainty in AOD under cloud-free conditions is of 0.01 at 550 and 865 nm and 0.02 at 440 nm (Holben et al., 1998). The coarse mode AOD requires almucantar and azimuth plane measurements; these requirements limit the amount of data since these scans cannot be accomplished nearly so regularly as the direct sun radiances which also allow the retrieval of the total AOD. The AE is calculated from multi-wavelength direct sun observations and delivers useful information on the aerosol size distribution. The simulated AE is computed from the estimated AOD at 550 and 865 nm whenever the model does not provide it.

Although AERONET provides daily averaged data of the above mentioned parameters we focus solely on the monthly mean. This provides a comprehensive picture of the seasonal dust cycle but precludes the evaluation of the frequency and intensity of dust events. The evaluation of the ability of global dust models to simulate individual dust events is beyond the scope of this work. Model monthly averages are constructed from daily means by selecting those days when observations are available. Note that an overall average from these monthly aggregates will be different than that of all daily data. We use all available stations with measurements for the year 2000 and a climatology constructed from the multi-annual database 1996–2006.

The AERONET network has stations spread around the world delivering aerosol data under various different atmospheric aerosol loads. In order to evaluate the models with respect to dust only, we selected those stations dominated by dust. We refer hereafter to these stations as "dusty" sites. We use a selection method based upon Bellouin et al. (2005) to differentiate between stations influenced by coarse, fine, or a mixture of these aerosol modes. In contrast to the authors who used the accumulation-mode fraction to discern between these three cases, we use the AE. We assume that AERONET stations with AE smaller than 0.4 are dominated by natural or coarse mode aerosols whereas those with values higher than 1.2 are dominated by anthropogenic or fine mode aerosols. Stations with values within these boundaries are exposed to a mixture of fine and coarse aerosols. Assuming that the AOD (at 440 nm) of oceanic aerosols does not exceed 0.15 (Dubovik et al., 2002), we filter out the oceanic aerosol stations from stations dominated by dust aerosols by eliminating those stations with monthly AOD (at 550 nm) smaller than 0.2. It should be noted that in remote stations fine mode desert dust can be mixed with other fine mode aerosols (sulphate, black carbon, organic matter) and thus have AE larger than 1.2. However, since we cannot separate these stations from those dominated by other fine mode aerosols based only on AE, we base our filtering criteria solely on the coarse mode. Therefore we define an AERONET station as "dusty" if it has simultaneously at least two months in the year (not necessarily consecutive) where the monthly average AE is smaller than 0.4 and where the monthly average of total AOD is larger than 0.2. We require at least two months in order to avoid selecting sites where a monthly average could be biased by a single day of low AE not necessarily linked to desert dust. For comparisons purpose however we consider all months with AE smaller

Table 1. Description of the global models considered in this study. Aerocom Median is not included in this table since it is constructed at every grid point and for every month by computing the local median from the models specified in Table 2. Models have been grouped according to their size ranges; models CAM to UMI simulate dust aerosols in the size range 0.1–10 μm, models ECMWF and LOA in the size range 0.03–20 μm and UIO_CTM in the range 0.05–25 μm. Models LSCE, TM5, ECHAM5-HAM and MIRAGE describe dust aerosols through 1, 2, 2 and 4 modes respectively.

N	Model	Resolution	Characteristics of size distribution	Winds[5]	Emissions	Reference of emission scheme	Model Reference
1	CAM	2.8°×2.8°×26 levels	4 bins 0.1–1–2.5–5–10 μm	Prescribed by GCM	Interactive[d]	Zender et al. (2003); Mahowald et al. (2006)	Mahowald et al. (2006)
2	GISS	5°×4°×20 layers	4 bins 0.1–1–2–4–8 μm	NCEP[b] reanalysis	Interactive	Cakmur et al. (2006)	Schmidt et al. (2006); Bauer and Koch (2005); Miller et al. (2006)
3	GOCART	2°×2.5°×30 layers	5 bins[1] 0.1–1.0–1.8–3.0–6.0–10.0	GEOS-3 DAS[c] Analysis	Interactive	Ginoux et al. (2001)	Chin et al. (2000)
4	SPRINTARS	1.125°×1.125°×20 layers	6 bins 0.1–0.22–0.46–1.0–2.15–4.64–10.0	NCEP[b] reanalysis	Interactive[d]	Takemura et al. (2009)	Takemura et al. (2005)
5	MATCH	1.9°×1.9°×28 layers	4 bins 0.1–1.0–2.5–5.0–10	NCEP[b] reanalysis	Interactive[d]	Zender et al. (2003)	Zender et al. (2003)
6	MOZGN	1.9°×1.9°×28 layes	5 bins 0.1–1.0–1.8–3.0–6.0–10.0	NCEP[b] reanalysis	Interactive[d]	Ginoux et al. (2001)	Horowitz et al. (2003); Tie et al. (2005)
7	UMI	2.5°×2°×30 layers	4 bins 0.05–0.63–1.25–2.5–10 μm radius	NASA DAO[a] reanalysis	Off-line	Ginoux et al. (2001)	Liu and Penner (2002); Liu et al. (2007)
8	ECMWF	0.7°×0.7°×60 levels	3 bins 0.03–0.55–0.9–20 μm	ECMWF reanalysis	Interactive[d]	Morcrette et al. (2009)	Morcrette et al. (2009)
9	LOA	3.75°×2.5°×19 layers	2 bins[2] 0.03–0.5–20 μm	ECMWF reanalysis	Off-line	Balkanski et al. (2004)	Reddy et al. (2005a, b)
10	UIO_CTM	2.8°×2.8°×40 layers	8 bins 0.03–0.07–0.16–0.37–0.87–2.01–4.65–10.79–25	ECMWF reanalysis	Interactive[d]	Grini et al. (2005)	Berglen et al. (2004); Myhre et al. (2007)
11	LSCE	3.75°×2.5°×19 layers	1 mode mmr=1.25 μm $\sigma 0 = 2.0$	ECMWF reanalysis	Interactive[d]	Balkanski et al. (2004)	Schulz (2007)
12	ECHAM5-HAM	1.8°×1.8°×31 layers	2 modes mmr=0.37,1.75 μm $\sigma 0 = 1.5, 2.0$	ECMWF reanalysis	Interactive	Tegen et al. (2002)	Stier et al. (2005)
13	MIRAGE	2.5°×2.0°×24 layers	4 modes[3] mmr=0.03,0.16,2.1,2.5 μm σ0=1.6,1.8,1.8,2.0	ECMWF reanlysis	Off-line	Ginoux et al. (2001)	Ghan and Easter (2006)
14	TM5	6°×4° global 1°×1° North America and Europe 25 layers	2 modes mmr=0.22,0.59–0.86 μm[4] $\sigma 0 = 1.59, 2.0$	ECMWF 12 h forecast	Off-line	Dentener et al. (2006)	Krol et al. (2005); de Meij et al. (2006)

[1] For optical calculations the first bin is distributed into 4 bins (0.1–0.18–0.3–0.6–1.0) by assuming a mass fraction.

[2] Emission follow lognormal with mmr 1.25 μm and $\sigma 0 = 2.0$.

[3] The mmr values are global annual averages. They vary spatially and temporally with the mode volume and number mixing ratios.

[4] The coarse mode diameter was varied for a fixed sigma in order to fit the data in Ginoux et al. (2001).

[5] Unless otherwise specified, the winds correspond to the year 2000.

[a] National Aeronautics and Space Agency Data Assimilation Office.

[b] National Centers for Environmental Prediction.

[c] Goddard Earth Observing System version 3 Data Assimilation System.

[d] Some tunning was done in the emission flux, in general to fit a given dataset of observations.

than 1.2. In addition, in view of the coarse resolution of the models (Table 1) and their difficulties to reproduce high altitude sites, we exclude stations above 1000 m a.s.l. Additional comparisons at each AERONET site between each model and AOD and AE are documented as time series in http://nansen.ipsl.jussieu.fr/AEROCOM/.

AERONET dusty sites are grouped regionally into Africa (AF), Middle East (ME) and Caribbean-American (C-AM) sites. Stations not belonging to any of the defined groups are considered separately. In each one of these groups stations are ordered from south to north. A list of the selected dust sites based on the measurements for the year 2000 and on the climatology constructed considering the multi-annual database 1996–2006 is given in Table S3 in the Supplement.

2.4 AeroCom models

We use fifteen model outputs from the AeroCom aerosol model intercomparison initiative (http://nansen.ipsl.jussieu.fr/AEROCOM/). This initiative is a platform for detailed evaluation of aerosol simulation by global models. It seeks to advance the understanding of global aerosol and its impact on climate by performing a systematic analysis and comparison of the results among global aerosol models including a comparison with a large number of satellite and surface observations (Textor et al., 2006). The comparisons conducted throughout the AeroCom project have revealed important differences among models in describing the aerosol life cycle at all stages from emission to optical properties (Kinne et al., 2006; Schulz et al., 2006; Textor et al., 2006, 2007; Koch et al., 2009; Prospero et al., 2010). The first of the comparisons considered a total of sixteen global models. Each model simulated the year 2000 using independently-selected simulation conditions. This experiment "A" is documented in Textor et al. (2006) and Kinne et al. (2006). A second experiment, "B", was conducted in which the same emissions were used in all models (Textor et al., 2007) and where radiative forcing was assessed (Schulz et al., 2006). In this present study we use the model outputs for the year 2000 of experiment A. For model TM5, which did not submit results for experiment A, we used results from experiment B instead.

The model features that are important for this work are presented in Table 1. For additional information on the models see Textor et al. (2006) and references therein. Four models from experiment A were excluded (ARQM, DLR, ULAQ, UIO_GCM) because their configuration was not meant to simulate the dust cycle. Furthermore, some models that joined the AeroCom project after the initial publication of experiments A and B were included (CAM, Community Atmosphere Model). Model names have changed with respect to previous AeroCom publications; MPI-HAM is now ECHAM5-HAM, KYU is now SPRINTARS and PNNL is now MIRAGE. We use also the AeroCom median model constructed at every grid point and for every month by computing the local median from the state-of-the-art AeroCom A models. Since some variables are not available from all models, the number of models used to construct the AeroCom median changes from variable to variable. Table 2 lists the models used to compute each variable. In the following comparisons and assessment the AeroCom median "model" will be treated as any other model in this study. Its performance with respect to the other models will be discussed in Sect. 4.

Table 2. Models used to compute the AeroCom median for each variable are indicated by an x. The variables are aerosol optical depth at 550 nm (AOD), Angström exponent (AE), dust surface concentration (SCONC) and dust total deposition (DEPO).

Model	AOD	AE	SCONC	DEPO
CAM				
GISS	x		x	x
GOCART	x		x	x
SPRINTARS	x	x	x	x
MATCH	x	x	x	x
MOZGN	x	x	x	x
UMI	x	x	x	x
ECMWF				
LOA	x	x	x	x
UIO_CTM	x	x	x	x
LSCE	x	x	x	x
ECHAM5-HAM	x	x	x	x
MIRAGE	x	x		x
TM5	x	x	x	x

We also include in this study the aerosol model developed within the Global and regional Earth-system Monitoring using Satellite and in-situ data (GEMS) project (Hollingsworth et al., 2008). This model fully describes the atmospheric life cycle of the main aerosol species; organic and black carbon, dust, sea salt and sulphate (Morcrette et al., 2009). It is now fully integrated in the operational four-dimensional data assimilation apparatus from the European Centre for Medium Range Weather Forecast (ECMWF). Hereafter, we refer to this model as ECMWF. Aerosol optical depth products from the Moderate resolution Imaging Spectroradiometer (MODIS) are assimilated to better estimate the aerosol fields (Benedetti et al., 2009). In this study we only consider simulations without data assimilation. For the evaluation of the impact of data assimilation on the model performance see Benedetti et al. (2009) and Mangold et al. (2011).

We evaluate the models in their performance to capture the yearly mean and the seasonal variability. For the yearly mean, we use scatter plots to analyse the model performance and we quantify this performance by computing the root mean square errors (RMS), the mean bias, the ratio of the modelled and observed standard deviation (sigma) and the correlation (R). Considering the different range of

measurement of the variables used in the study and in order to allow the intercomparison of the model performance for the different variables, we include the normalized root mean square (NRMS) error and mean normalized bias (MNB). We use the NRMS to quantify the average model-observations distance and the MNB for the average over- and underestimation. These statistics are computed as follows:

$$\mathrm{NRMS} = \sqrt{\frac{1}{S}\sum_{i=1}^{S}\sum_{j=1}^{T}\left(\frac{m_{ij}-o_{ij}}{o_{ij}}\right)^2} \quad (1)$$

$$\mathrm{MNB} = \frac{1}{S}\sum_{i=1}^{S}\sum_{j=1}^{T}\frac{m_{ij}-o_{ij}}{o_{ij}} \quad (2)$$

where S is the number of stations considered and T the total number of month used in the analysis for each station, o_{ij} is the observed value at the station i and month j and m_{ij} is the corresponding model monthly average at the closest grid point to the station. For the seasonal analysis we use Hovmoller-like diagrams where each row corresponds to a given station. These diagrams are usually designed to indicate spatial propagation of features with time. However, we choose to group the stations not in a geographically meaningful way as is usually done in Hovmoller diagrams but regionally to ease the assessment of the dust cycle. To evaluate the model performance to reproduce the observed seasonal cycle we use the MNB and the centred pattern root mean square (CPRMS) error. The latter corresponds to the RMS error when the bias has been removed (Taylor, 2001) and thus illustrates the average difference between the models and the observations. We compute the CPRMS as the difference between the NRMS and MNB and obtain in fact a normalized CPRMS (NCPRMS). The NRMS and MNB are computed as follows:

$$\mathrm{NRMS} = \sqrt{\frac{1}{N}\sum_{l=1}^{N}\left(100\times\frac{\mathbf{M}_l-\mathbf{O}}{\mathbf{O}}\right)^2} \quad (3)$$

$$\mathrm{MNB} = \frac{1}{N}\sum_{i=1}^{N}\left(100\times\frac{\mathbf{M}_l-\mathbf{O}}{\mathbf{O}}\right) \quad (4)$$

where N is the number of models used in the study, $\mathbf{O}$ is the array containing the elements o_{ij} defined above and $\mathbf{M}_l$ is the equivalent array of each model l with the elements m_{ij}. Both of these arrays have dimensions of $S\times T$. We highlight that in Eqs. (3) and (4) the sum is conducted over to the total number of models as opposed to Eqs. (1) and (2) where the sum is conducted with respect to the stations and months and the operation is repeated for each model. The observation array $\mathbf{O}$ contains the data for each station and each month and remains therefore constant in Eqs. (3) and (4). To highlight this fact we decide to omit the indexes on both arrays indicating stations (i) and months (j). This characteristic of $\mathbf{O}$ allows us to continue computing CPRMS as the difference between RMS and bias in spite of the normalization. The NCPRMS is then calculated from

$$\mathrm{NCPRMS} = \sqrt{\mathrm{NRMS}^2-\mathrm{MNB}^2} \quad (5)$$

In addition, we use the normalized standard deviation (NSTD) to assess the spread among the models to simulate the seasonal cycle or model diversity. The normalized standard deviation is computed as follows:

$$\mathrm{NSTD} = \sqrt{\frac{1}{N-1}\sum_{l=1}^{L}\left(100\times\frac{\mathbf{M}_l-\bar{\mathbf{M}}}{\bar{\mathbf{M}}}\right)^2} \quad (6)$$

where $\bar{\mathbf{M}}$ is an array of $S\times T$ elements with the average over all models for each station and month. Finally, we also include the Hovmollers of the individual models in the Supplement. Throughout the text we use the term "diversity" as employed in Textor et al. (2006) namely "to describe the scatter of model results".

We computed the global model dust budgets for each one of the models (Table 3). The annual emissions of the AeroCom models in Phase I are between 500 and 4400 Tg yr^{-1}. This range exceeds the range of 1000–3000 Tg yr^{-1} usually attributed to global models (e.g. Zender et al., 2004). The global averaged dust AOD ranges from 0.01 to 0.053 with 80 % of the models having a value between 0.02 and 0.035. The lifetime of dust aerosols is between 1.6 and 7.1 days for most models.

Note that the model results used in the present analysis correspond mostly to simulations submitted before the year 2005. Many of these models have been significantly improved since submitting their simulations. Therefore the results presented in this study do not necessarily represent the current state of the models.

3 Results

The ability of each model to reproduce different aspects of the desert dust cycle is evaluated by comparing them against the data sets described above. We conduct the analysis on a station by station basis. We use the AeroCom tools developed at the Laboratoire du Climat et de l'Environnement (LSCE) to conduct the comparison and evaluation. The global annual distribution of total deposition, surface concentration, AOD and AE of the AEROCOM median model have been included in the Supplement (Fig. S1). The corresponding figures of the remaining models can be found via the AeroCom web interfaces (http://nansen.ipsl.jussieu.fr/AEROCOM/data.html).

3.1 Dust deposition

The comparison of total annual deposition and simulated deposition flux is presented in Fig. 1. See Table S1 in the Supplement for further information on the stations. The bias at

Table 3. Mass balance for each one of the models. NaN represents variables not provided by model. MEE corresponds to the mass extinction efficiency.

N	Model	Size [μm]	Emission [Tg yr^{-1}]	Load [Tg]	Deposition [Tg yr^{-1}]	Wet Depo [Tg yr^{-1}]	Dry Depo [Tg yr^{-1}]	Sedim [Tg yr^{-1}]	OD550 Dust	MEE [m^2 g^{-1}]	Life Time [days]
1	CAM	0.1–10	4313	25.7	4359	1382	2300	675	0.035	0.69	4.6
2	GISS		1507	29.0	1488	456	352	680	0.034	0.60	7.1
3	GOCART	(bins)	3157	29.5	3178	583	120	2475	0.035	0.60	3.4
4	SPRINTARS		3995	17.2	3984	628	2791	565	0.024	0.72	1.6
5	MATCH		981	17.3	1070	517	431	122	0.033	0.96	5.9
6	MOZGN		2371	21.1	2368	425	1943	NaN	0.022	0.52	3.3
7	UMI		1688	19.3	1691	619	1073	NaN	0.021	0.56	4.2
8	ECMWF	0.03–20	514	6.8	750	406	322	22	0.027	0.25	3.3
9	LOA	(bins)	1276	13.7	1275	417	521	336	0.034	1.28	3.9
10	UIO_CTM	0.03–25 (bins)	1572	21.7	1571	681	890	NaN	0.026	0.61	5.1
11	LSCE		1158	20.3	1156	616	310	231	0.031	0.77	6.4
12	ECHAM5-HAM	modes	664	8.2	676	374	37	265	0.010	0.60	4.4
13	MIRAGE		2066	22.0	2048	1361	687	NaN	0.053	1.22	3.9
14	TM5		1683	9.3	1682	295	592	794	0.013	0.68	2.0
15	AEROCOM_MEDIAN	–	1123	15.8	1257	357	396	314	0.023	0.72	4.6

most stations is within a factor 10 of the observations. All the models in this study present a positive mean normalized bias (MNB) in the deposition fluxes ranging from 0.1 to 140.3. However, if the model CAM is not considered the maximum MNB decreases to 13.4. In addition, if the measurements from remote regions of the Southern Ocean and close to the Antarctica (dark blue in Fig. 1), mostly overestimated by the models, are excluded, seven from the 15 models produce a negative MNB. While most models mainly overestimate the deposition data from Mahowald et al. (2009) in the Antarctica, most of them underestimate the deposition in the Weddell Sea (13) in Antarctica (DIRTMAP; Tegen et al., 2002). This difference in performance will be discussed in Sect. 4. Most models (12 out of 15) underestimate the deposition in the Pacific and the South Atlantic Ocean, while eight models overestimate the deposition in Europe (green) and the North Atlantic (orange) and nine in the Indian Ocean (dark green). There is only one data set of deposition measurements in the Taklimakan desert in central Asia (station H, purple in Fig. 1). The model estimates of deposition at this site vary over a large range, yet mainly underestimating the observations.

We expand the analysis conducted on 9 AeroCom models in Prospero et al. (2010) to estimate the wet and total deposition of the FAMS network in Florida. Measurements were conducted during almost three years and represent an invaluable source of data to evaluate not only the simulated wet and total deposition but also the simulated dust transport across the Atlantic. As in that study, to illustrate the model performance we chose three representative stations from the nine stations in the FAMS network. These stations are oriented from south (Little Crawl Key) to north (Lake Barco) and therefore provide a latitudinal gradient of deposition in Florida. The general conclusions from that study are also valid for the entire AeroCom model set of 15 models. Most models capture the seasonality of the deposition and the dominance of wet deposition in the summer months, from July to September, but only a few models capture the magnitude of the deposition (wet and total) in this period - most underestimate it (Fig. 3). The model performance deteriorates from south to north, reflecting model difficulties in transporting the dust northward. In general, the models overestimate the role of the wet deposition. They manage to reproduce the fraction of wet deposition in regions where the wet deposition dominates but fail to do so in sites dominated by dry deposition (Table 4).

3.2 Surface concentration

We analyze the correspondence between observed and modelled yearly average surface concentrations at each site first (Figs. 4 and 5) and then evaluate the simulated seasonality (Fig. 6).

We first compare the simulated surface concentration to short-term measurements from cruises (squares and filled-in circles in Fig. 4) and long term measuring stations (diamonds in Fig. 4). Because major dust events occur on a relatively small number of days per year (∼5 %, Mahowald et al., 2009) and because of the limited number of ship measurements, it is possible that the measurements miss one (or more) of the events or that they actually coincided with an event. The error in the measurements associated with missing a dust event or coinciding with one is represented by the vertical lines in Fig. 4. For each model these

Table 4. Fraction of wet deposition [%]. The compilation is taken from Mahowald et al. (2010).

	Obs [%]	AeroCom Median	CAM	ECMWF	GISS	GOCART	SPRINTARS	LOA	LSCE	MATCH	MOZGN	-HAM ECHAM5	MIRAGE	TM5	UIO_CTM	UMI
Bermuda	17–70	79	34	90	86	78	59	88	83	86	64	83	91	87	91	95
Amsterdam Island	35–53	80	1	96	77	59	76	93	96	82	66	83	85	88	96	91
Cape Ferrat	35	79	61	91	53	60	78	77	89	82	78	79	86	90	88	83
Enewetak Atoll	83	89	7	95	89	77	71	90	77	94	83	81	92	92	93	94
Samoa	83	93	3	96	88	86	72	95	92	96	85	86	95	92	94	97
New Zealand	53	85	2	90	79	68	81	90	94	88	59	82	92	94	84	91
Midway	75–85	80	27	88	NaN	67	60	91	85	84	65	78	92	92	96	94
Fanning	75–85	75	9	97	84	70	43	86	84	91	75	65	87	93	94	93
Greenland	65–80	87	58	98	72	68	92	96	97	95	75	82	95	93	87	92
Coastal Antartica	90	60	0	100	31	20	85	90	87	85	71	84	94	90	96	81

errors correspond respectively to 96 % and 20 % of the model yearly average. In spite of the large uncertainties, these observations deliver valuable information in remote regions that are seldom sampled (e.g. the Southern Ocean and South Atlantic Ocean) but where dust could have a great impact on the biogeochemical cycle because of the high concentrations of primary nutrients. All models mainly overestimate the surface concentrations, exceeding in most of the cases two orders of magnitudes with respect to the observed concentrations; the MNB varies between 34.08 and 1249.6. The cases with large overestimation correspond mainly to short-term cruise measurements in regions downwind of the main dust sources. However, the models perform equally well against cruise data in remote regions of the Southern Hemisphere (i.e. South Atlantic and Indian Ocean) as they do against long-term measurements in other regions (diamonds in Fig. 4); over and underestimation is mostly within two orders of magnitude. All models agree in mostly overestimating the cruise data in the Indian Ocean. In the South Atlantic however, one third of the models underestimate the surface concentration, one third overestimate it and finally one third both under and overestimates the surface concentration.

We next compare the models to yearly averages of the SEAREX and AEROCE data. The over and underestimation is mostly within a factor 10 (Fig. 5), except for Antarctica (stations 1, 8 and 9). The correspondence between modelled and measured surface concentration in most models improves in stations with higher values; the agreement is much better in stations downwind of major dust sources (HIGH, stations 17 to 22) than in the other two groups (Fig. 5). Likewise, the correspondence is better for stations of the second group (MEDIUM, stations 8 to 16) than for the first one (LOW, stations 1 to 7). Half of the models present larger differences with the observed surface concentrations at sites associated with the Asian sources (stations 15, 20 and 22) than at stations measuring the trans-Atlantic dust transport from the Sahara (stations 18 and 19). The above suggests difficulties in simulating simultaneously the magnitude of the dust emissions from Sahara and Asia (Tegen et al., 2002). The remaining models produce similar performance in reproducing surface concentrations associated with both deserts. All models underestimate the surface concentration at Rukomechi in Zimbabwe (17), which measures the dust emitted from the Kalahari Desert.

The yearly cycle of measured surface concentration and a measure of the model performance to reproduce these observations (Sect. 2.4) are presented in Fig. 6. Each row corresponds to the monthly surface concentration of a network station illustrated in Fig. 2. The measurements at each numbered station in Fig. 2 are presented in the numbered row in Fig. 6. As in Figs. 2 and 5, we continue to group the stations as LOW, MEDIUM and HIGH according to their surface concentration regime.

In all three groups the underestimation is smaller than the overestimation and in general no significant differences in the MNB are observed between the groups. Likewise, no significant difference in the errors (CPRMS) is seen between the three groups. The standard deviation reveals large spread in the models to simulate the surface concentration, exceeding in most of the cases 100 % and in some cases up to 500 %. Yet important diversity exists between the models in the different group of stations. The largest diversity among the stations is seen in the Antarctica followed by the station on the western Atlantic Ocean. This diversity will be discussed below in more detail.

The models on average underestimate the surface concentrations in the Antarctica stations (1, 8 and 9) throughout the year in coherence with Fig. 5. In Mawson (1) and Palmer (8) the largest errors coincide mostly with the period of low surface concentration from March till September for the former and austral summer and early autumn for the latter. In King George (9), on the contrary, large errors occur in both, months with low and high surface concentration. The large model diversity seen in these stations occur mainly from late austral spring till early autumn in Mawson and throughout most of the year in Palmer and King George. In Mawson, periods of large diversity coincide mostly with month with large errors.

The stations New Caledonia (2), Norfolk Island (12), Cape Grim (10) and Jabirun (13) illustrate the Australian dust cycle. While New Caledonia belongs to the LOW group, characterized by surface concentrations below 1 $\mu g\,m^{-3}$, the

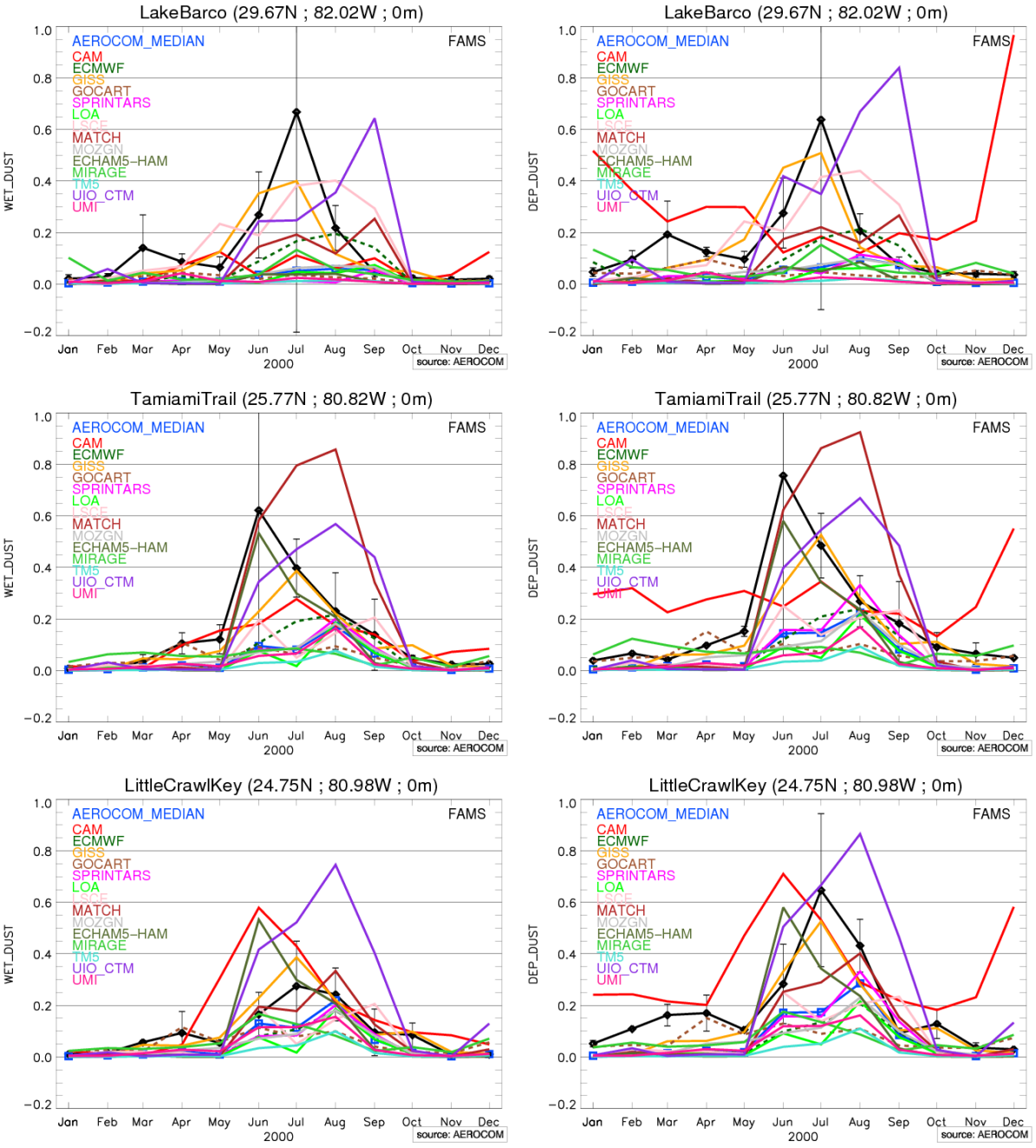

Fig. 3. Modeled and observed wet (left) and total (right) dust deposition rates at three sites from the Florida Atmospheric Mercury Study (FAMS) network: Lake Barco (LB), Tamiami Trail (TT), and Little Crawl Key (LCK). The black line is the mean of the 3 years of FAMS data from 1994–1996. Vertical lines correspond to one standard deviation of the 3 yr average. Units are g m^{-2} month^{-1}.

stations of Cape Grim, Norfolk Island and Jabirun belong to the MEDIUM group. It is interesting to compare the yearly average at New Caledonia and Norfolk Island. These stations are relatively close to one another (800 km) but they lie in quite different dust regimes. Measurements suggest that Norfolk Island is impacted by Australian dust while New Caledonia lies outside of the northeast dust transport pathway from Australia (Mackie et al., 2008). Most models do not reproduce the different dust regimes in both stations and attribute to New Caledonia the same range of measurement and seasonality as in Norfolk Island. This is illustrated by the overestimation throughout most of the year in New Caledonia and large errors in Norfolk Island. In addition, important model diversity is seen in these two stations mainly during austral

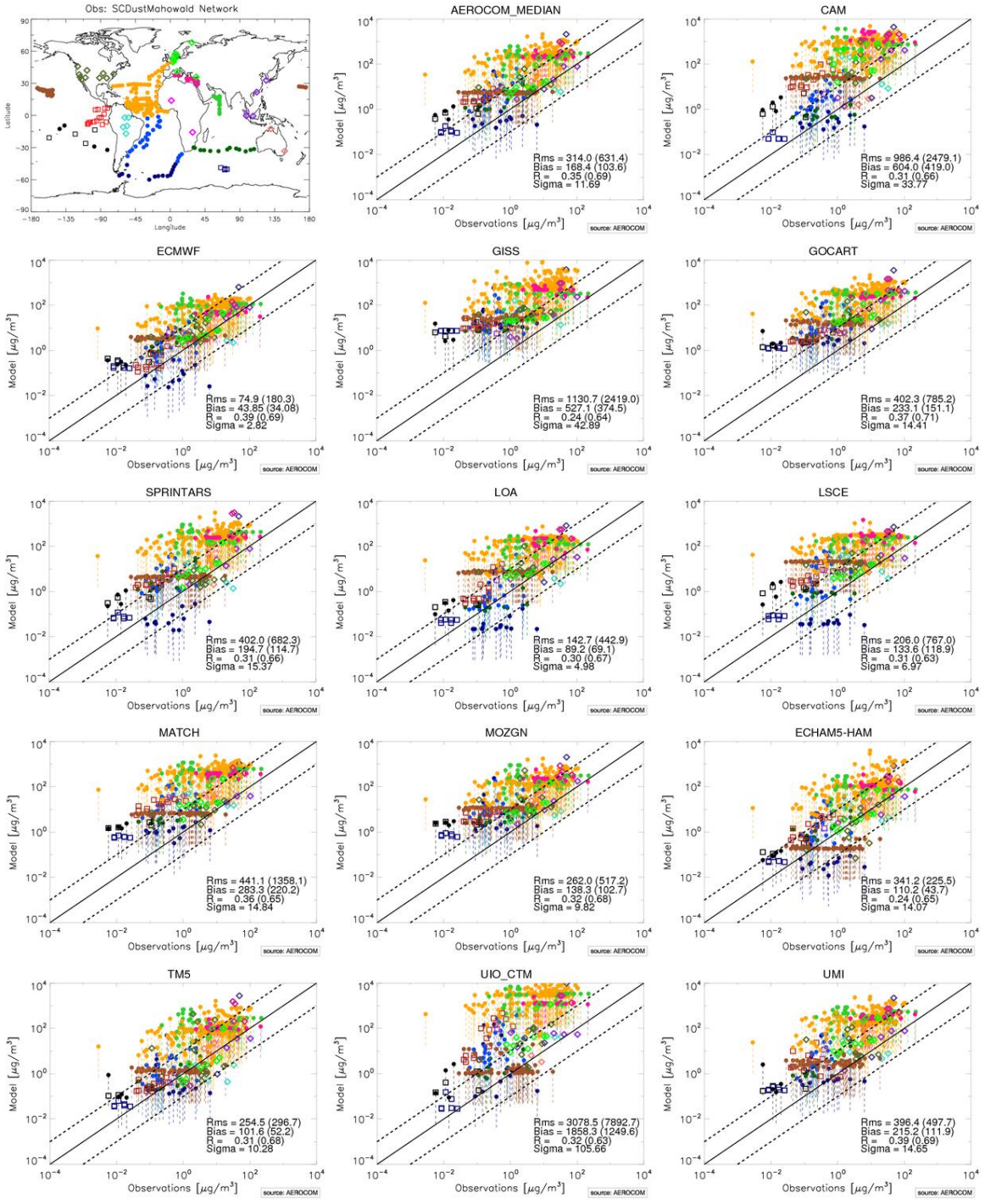

Fig. 4. Scatterplot of surface concentration from Mahowald et al. (2009) versus modeled one. Short-term measured Fe (and converted to dust by assuming a 3.5 % Fe in dust) during cruise are represented by filled-in circle. Data corresponding to long term measurements are illustrated with diamonds while measurements of Aluminium or dust during cruise are indicated by squares. The colored dotted lines are estimates of the error in the model-data comparison when the model represents the annual mean, while the data is taken on a few days. The methodology is discussed in the text (Sect. 2.2). Root mean square error (RMS), mean bias, ratio of modeled and observed standard deviation (sigma) and correlation (R) are indicated for each model in the lower right part of the scatterplot. Mean normalized bias and normalized root mean square error are given in parenthesis next to RMS and mean bias, respectively. The correlation with respect to the logarithm of the model and of the observation is also given in parenthesis next to R. Black continuous line is the 1:1 line whereas the black dotted lines correspond to the 10:1 and 1:10 lines.

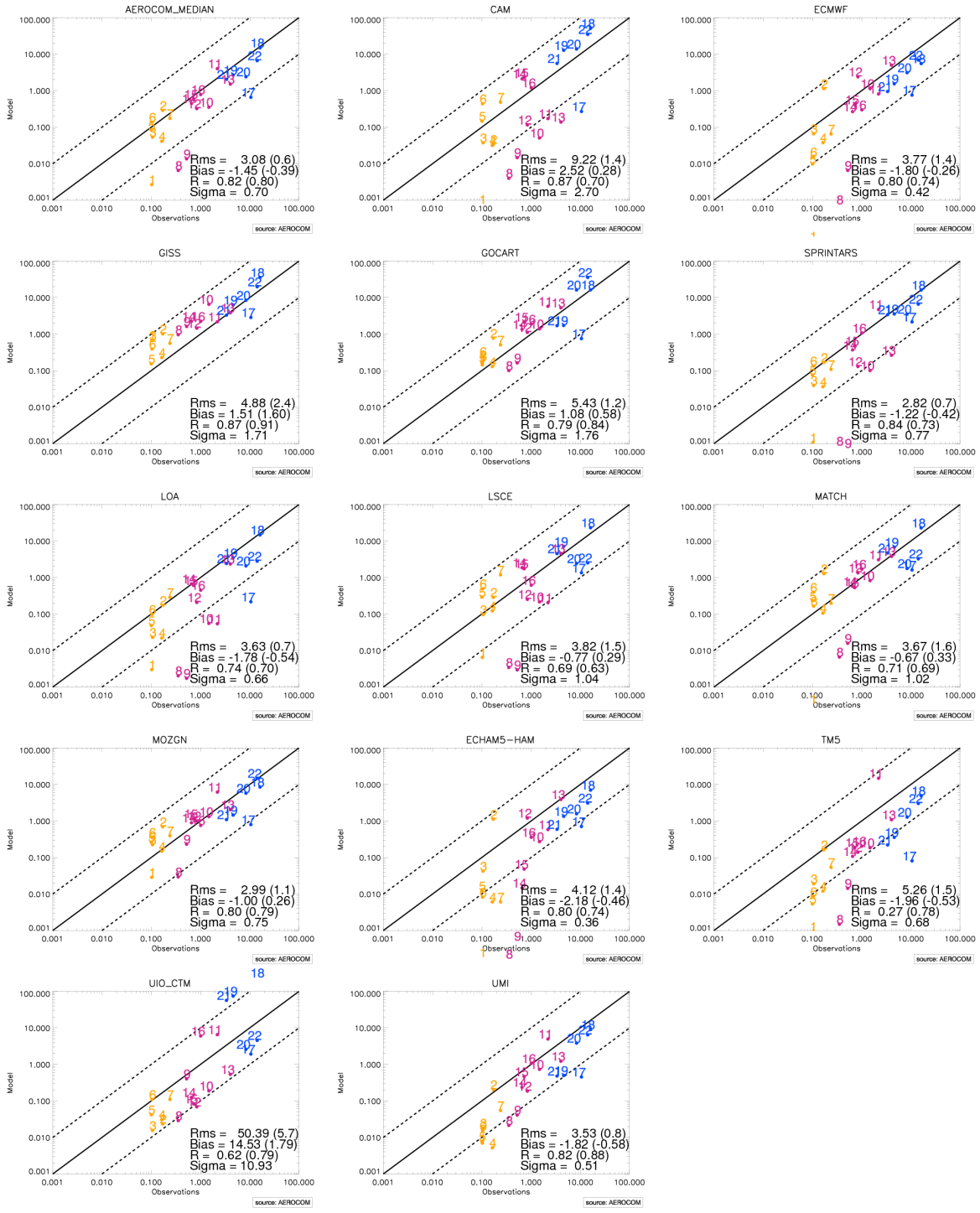

Fig. 5. Yearly averaged measured surface concentration from the network operated by the University of Miami versus modeled one at each station, units are $\mu g\,m^{-3}$. Stations are grouped according to the regime of measured data into remote stations (orange), stations under the influence of minor dust sources of the Southern Hemisphere or remote sites in the Northern Hemisphere (violet) and locations downwind of African and Asian dust source (blue). The location of each station is illustrated in Fig. 2 and given in Table S2 in the Supplement. Root mean square error (RMS), bias, ratio of modeled and observed standard deviation (sigma) and correlation (R) are indicated for each model in the lower right part of the scatter plot. Mean normalized bias and normalized root mean square error are given in parenthesis next to RMS and mean bias, respectively. The correlation with respect to the logarithm of the model and of the observation is also given in parenthesis next to R. Black continuous line is the 1:1 line whereas the black dotted lines correspond to the 10:1 and 1:10 lines.

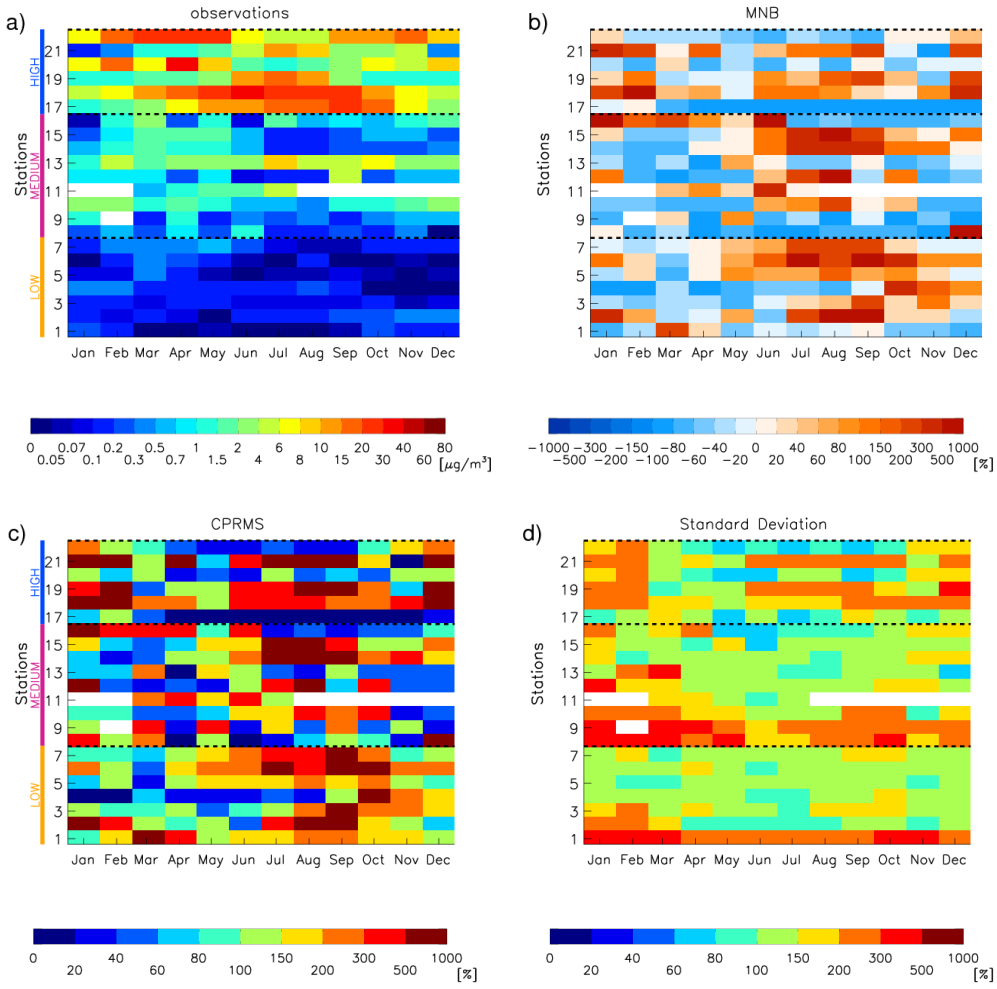

Fig. 6. Monthly averages of measured surface concentration in $\mu g\,m^{-3}$ are shown together with the mean normalized bias (MNB), the normalized centred pattern root mean square error (CPRMS) and the normalized standard deviation (Sect. 2.4). In all subfigures, each row corresponds to the seasonal cycle at one of the stations and the row for each station corresponds to the number presented in Fig. 2. The stations have been grouped into Low (orange), Medium (violet) or High (blue) surface concentration sites (Sect. 2.2) and each group is identified by a colored bar on the left side of the left hand figures. Stations are ordered from south to north within each group. White color corresponds to months without measurements. For the individual figure of each model presenting the simulated values and their differences (in %) with respect to observations see Figs. S2 and S3, respectively, in the Supplement.

summer. This may suggest difficulties by most models to correctly simulate the transport of Australian dust to the east. However, the differences between both stations could be related to the fact that the dust data are a climatology whereas the model data are for a specific year. Dust emissions in Australia (Mackie et al., 2008) are highly episodic from year-to-year; consequently the model overestimation might actually be the result of a small number of events that may have occurred in 2000 but not captured in the long-term measurements. The stations Cape Grim (10) in southern Australia, Norfolk Island (12) offshore eastern Australia and Jabirun (13) in northern Australia present all different seasonal cycles. In Cape Grim the months with maximum surface concentration are from late austral spring till early autumn while in Norfolk Island the maximum is observed in September with an additional period of large concentrations from January till March and in Jabirun large concentrations are seen from February till October with the maximum in July. The MNB reveals that the models mainly underestimate the observations throughout most of the year at these stations and the CPRMS shows that the largest errors do not necessarily coincide with months of maximum surface concentration. Likewise, the largest model spread in these stations is seen in periods of large surface concentration but not necessarily coincident with the maxima. The large values of standard deviation correspond not only to a spread in the magnitude of the simulated value but also on the duration and occurrence of period of maximum concentration (Fig. S2 in the Supplement).

The measurements in Hedo (20) and Cheju (22) present an annual cycle with maxima in spring and minima in summer, which corresponds to the maximum in dust storm activity in China (Prospero et al., 1989; Prospero, 1996). An additional peak in surface concentration exists at these stations in winter or late fall. The observations suggest that there is substantial dust transport to these costal regions throughout the year; however, some of this dust may be derived from relatively localized sources. The station in Midway Island (15), in central North Pacific and far off the east coast of Asia, is also impacted by aerosol transport from Asia (Prospero et al., 2003; Su and Toon, 2011). The measurements present a similar springtime maximum as the one in Cheju and Hedo and low concentrations throughout the rest of the year. The springtime maximum in Midway illustrates an important long-range dust transport of Asian dust. Most models mainly underestimate throughout the year the surface concentration in Hedo and Cheju whereas they mostly overestimate it in Midway. In Hedo and Cheju the largest difference with respect to the observations occur from late boreal autumn till observed early spring coinciding with the onset of the period of maximum surface concentration. In Midway however the largest errors occur from July to September after the period of maximum concentration, yet important errors are also seen in the month of May coinciding with the offset of the period of maximum surface concentration. The model spread remains mostly constant throughout the year in Midway while in Hedo and Cheju the largest diversity coincides mostly with month with large errors.

The measurements in Barbados (18) and Miami (19) capture the transatlantic transport of Saharan dust. The former presents an annual cycle with maximum between March and October while the latter has maxima from July to August. The surface concentration is mostly overestimated throughout the year and in particular at months with maximum surface concentration. However the largest errors with respect to the observations are observed in boreal winter month, outside the period of maximum transatlantic Saharan dust transport. In general terms, the models reproduce the annual cycle of surface concentration in Barbados but present important diversity in both, the extension and intensity of the observed large surface concentration from March to October. This diversity reaches its maximum in the winter months. The model performance to simulate the surface concentration deteriorates towards the north in Miami, both in terms of CPRMS and standard deviation. In general the models present larger discrepancies with the observation in Miami than in Barbados and model spread is also larger than in Barbados. All the above suggests that the models have more difficulties to reproduce the annual cycle in Miami than Barbados. The data do not allow us to conclude whether this difficulty is due to problems in simulating the processes responsible for the northward extension of the transported transatlantic Saharan dust or to difficulties in simulating aerosol removal processes.

To test the simulated seasonal cycle in dust transport across the Atlantic and its northern latitudinal extend, we compare the monthly mean model results to means of daily measurements in Barbados (18) and Miami (19) for the year 2000 (Fig. 7). At both Barbados and Miami there is a clear annual cycle in dust transport which yields a pronounced summer maximum. The Barbados record differs from Miami in that the peak concentrations are higher and the dust transport season extends through the late Spring and early Fall. At Barbados the model results differ greatly from the measurements over much of the year. The disparity is especially large in the summer. Over the reminder of the year, principally October to May, most models underestimate the surface concentrations. At Miami the model dispersion in reproducing the measurements is smaller. However, some models that reproduce the seasonal cycle at Barbados fail to do so in Miami. This suggests that these models have problems in simulating the processes responsible for the northward displacement of the dust transport. The seasonal cycle for the year 2000 is not unusual and follows the average from the 1996–2006 climatology (Fig. 7). However there are some differences, most notably the peak in surface concentration in Miami in the year 2000 lags the climatology by one month. At Barbados the climatology shows a maximum in June with steadily decreasing values thereafter; however in the year 2000 there are two maxima, one in June coincident with the climatology and one in August somewhat above climatology whereas July is well below climatology. Most models show a clear maximum in June, in agreement with the seasonality of measurements and, like the dust climatology, they decrease steadily thereafter. It is notable that a few models yielded very high monthly means at Barbados and Miami. Among the models with the highest values are UIO_CTM, CAM and GISS. While UIO_CTM reaches high monthly means at both stations (aprox. $450\,\mu g\,m^{-3}$ in Barbados and nearly $300\,\mu g\,m^{-3}$ in Miami) CAM and GISS largely overestimate the observations only in Barbados. Both models simulate surface concentrations close to $100\,\mu g\,m^{-3}$.

3.3 Total aerosol optical depth

We now compare the models to AERONET total and coarse mode AOD, first in terms of the average and then in their ability to reproduce the seasonal variability at dusty sites. The average is constructed by using only selected months (as defined in Sect. 2.3) and therefore it is not a yearly average. First we base the analysis on the climatology constructed using the multi-annual database 1996–2006 (Sect. 2.3) and then on the data of the year 2000. In both cases, dusty stations have been grouped regionally into African (AF), Middle East (ME) and Caribbean-American (C-AM) stations and stations elsewhere (Fig. 8). In each of these regions the stations are organized from south to north.

A total of 25 AERONET stations are considered as dusty sites based on the AE and the AOD when climatological data

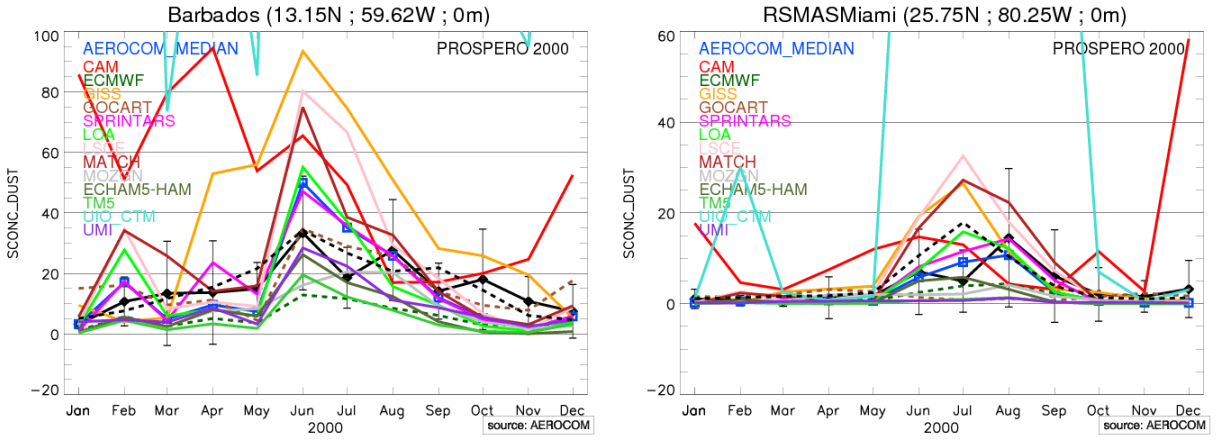

Fig. 7. Measured and simulated surface concentration in Barbados and Miami. Measurements of the year 2000 are presented by the black continuous line and the climatology (Fig. 6) is presented by the black dashed line. Units are $\mu g\,m^{-3}$.

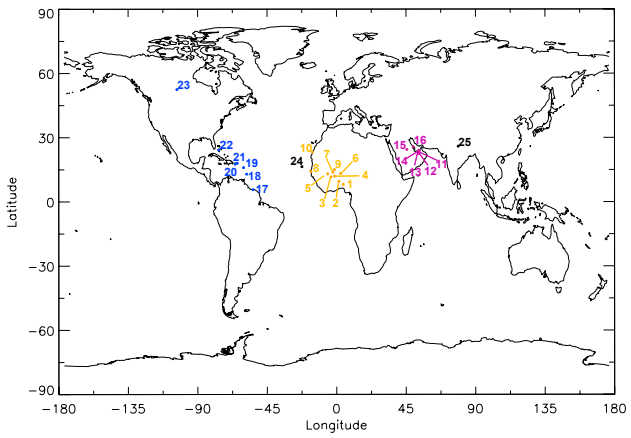

Fig. 8. Location of selected AERONET dusty sites based on the climatology built from the multi-annual database 1996–2006. Dusty stations are grouped regionally; Africa (orange), Caribbean-America (blue), Middle East (violet) and elsewhere in the world (black). Names and locations for each selected station are given in Table S3 in the Supplement.

are used (Sect. 2.3, Fig. 8). Names and locations for each one of these sites are given in Table S3 of the Supplement. In general the modeled AOD is within a two-fold range of the observations at most sites (Fig. 9). The mean normalized bias (MNB) of all models varies between −0.44 and 0.27 while the normalized root mean square error (NRMS) varies between 0.3 and 0.6. More than half of the models (8 out of 15) produce a negative MNB varying between −0.44 and −0.03 and NRMS varying between 0.3 and 0.6. For models mainly overestimating the AOD, the MNB varies between 0.02 and 0.27 and the NRMS between 0.3 and 0.5. The data show in general higher AOD at African stations than at those in the Middle East, which in turn have larger values than the American stations. In general, the models reproduce this gradient between regions. Eight of the 15 models underestimate the averaged AOD at all or almost all American stations. Some models do not reproduce the observed gradient in AOD between African and Middle East dusty stations, instead producing similar AOD in the Middle East and in Africa. Others overestimate the AOD for the African stations. Considering the closeness of the stations to the sources in both regions, the overestimation of AOD points to an overestimation of dust emissions or underestimation of the removal in the Middle East and/or Africa. Another cause could be the use of wrong size distribution with the consequent impact on the estimation of the extinction. This aspect will be further developed in Sect. 4. Finally, twelve models underestimate the AOD in Kanpur (25) in northern India, again suggesting that most models underestimate emissions of the Great Indian Desert or overestimate the removal.

The seasonality of the AOD climatology in Africa (Fig. 10) is characterized by high AOD with maximum values from December to April in the most southern stations shifting progressively to July through September in the most northern African stations. In general the underestimation coincides with the months of maximum AOD. Additional underestimation is observed from July till October in most stations. The overestimation of AOD in general corresponds to the month of late fall and early winter (November and December) and the month preceding the month of maximum AOD and presents thus also a progressive shift from late winter till late spring in stations from south to north. Exceptions to the above described behavior are the southern most stations of Ilorin (station 1) and Djougou (2) in Nigeria and Benin respectively where the underestimations extends throughout the year and in western Sahara at Dahkla (10) where the AOD is overestimated throughout most of the year. The seasonal cycle in Ilorin and Djougou is reproduced by most models and the underestimation might be indicative of

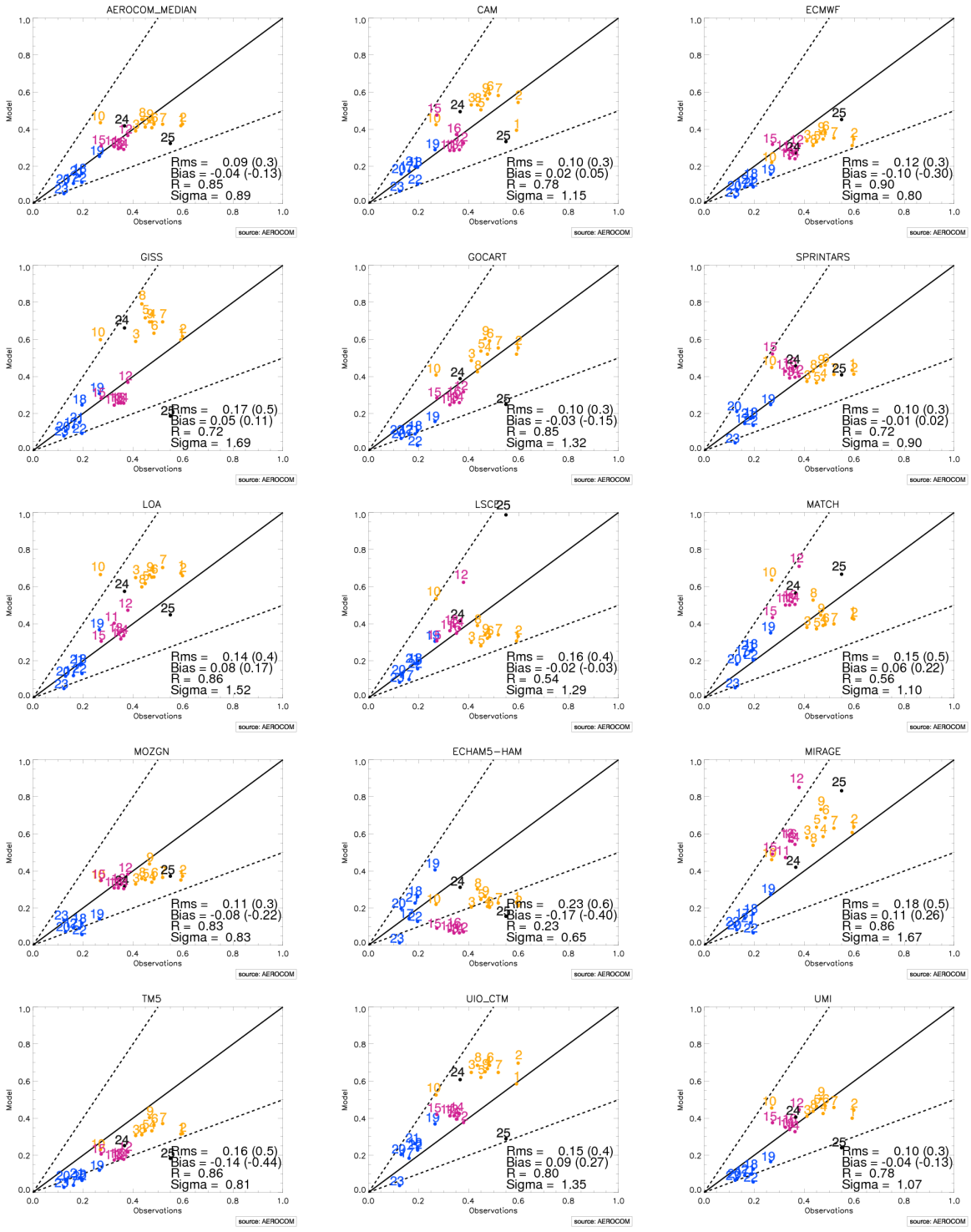

Fig. 9. Averaged AOD at 550 nm versus modeled one at dusty stations of the AERONET network. Data from the climatology based on the multi-annual database 1996–2006 are used. Stations are regionally grouped into African (orange), Middle East (Violet) and Caribbean-American stations (Blue) and stations elsewhere (black). Location of each station is illustrated in Fig. 8. Name and location of each station is given in Table S3 in the Supplement. Root mean square error (RMS), bias, ratio of modeled and observed standard deviation (sigma) and correlation (R) are indicated for each model in the lower right part of the scatter plot. Mean normalized bias and normalized root mean square error are given in parenthesis next to RMS and mean bias, respectively. Black continuous line is the 1:1 line whereas the black dotted lines correspond to the 2:1 and 1:2 lines.

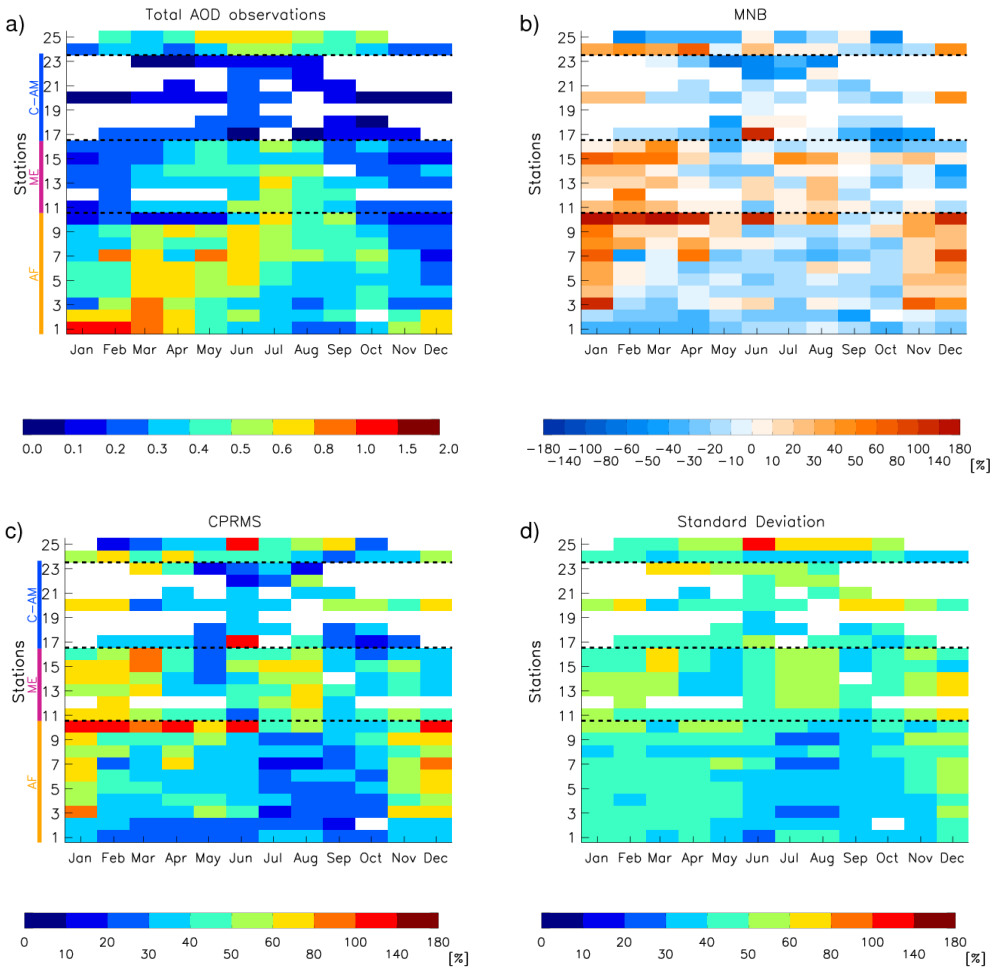

Fig. 10. AERONET AOD at 550 nm is shown together with the mean normalized bias (MNB), the normalized centred pattern root mean square error (CPRMS) and the normalized standard deviation (Sect. 2.4). In all sub-figures, each row corresponds to the seasonal cycle at one AERONET station. They have been grouped into African (AF, orange), Middle East (ME, violet) and Caribbean-American (C-AM, blue) stations and stations elsewhere in the world (OT, black). Each one of these groups is identified by a coloured bar on the left side of the left hand figure. Stations are ordered from south to north within each group. The row for each station corresponds to the number presented in Fig. 8. Name and location of each station is given in Table S3 of the Supplement. White color corresponds to month without measurements or month not complying with the selection criteria (Sect. 2.3). AERONET data correspond to the climatology based on the multi-annual database 1996–2006. For the individual figure of each model presenting the simulated values and their differences (in %) with respect to observations see Figs. S4 and S5, respectively, in the Supplement.

difficulties in simulating the emissions or removal processes. In Dahkla on the contrary the overestimation is the result of a very long period with large AOD simulated by most models. The largest differences with respect to the observations (illustrated by the CPRMS) coincide in general with the months where the AOD is overestimated. The largest errors are seen in Djougou in the first half of the year and December. The spread between the models varies mostly between 30 and 45 % with some isolated month where the spread varies between 50 and 60 %. The model diversity presents in general a seasonal cycle with minimum in summer and early autumn and maximum the rest of the time. Contrary to the cycle associated to the MNB and CPRMS, maximum diversity or standard deviation is seen in months with both, large and small AOD.

In the Middle East there is a seasonal cycle with maximum AOD from May–June to September (Fig. 10). In general the simulated AOD is mostly overestimated from January to August and underestimated afterwards. This period of overestimations in general corresponds to the months of maximum AOD and those preceding it. Again, the largest errors are mostly seen in month where the AOD is overestimated. Models present larger diversity in simulating the AOD in the Middle East than in Africa and the spread between models is

mostly coincident with periods of overestimation and large errors described above. Exceptions to this are the month of November and December where large values in standard deviation are seen in periods of small error and underestimation of the AOD.

At Caribbean-American (C-AM) stations there are large periods that have no data, mostly in the early and late months of the year (Fig. 10). The magnitude of the model diversity is in general larger than at African stations. The AOD is mostly underestimated throughout the year except for the boreal winter months in La Parguera (20) in Puerto Rico. This station presents also the largest errors (CPRMS) mostly in months with low AOD. In general large errors are observed at stations affected by transatlantic dust transport (stations 18 to 21). In addition, at these stations, the largest spread between the models is coincident to the months with largest errors. With respect to the individual stations, no model simulates the AOD in Paddockwood (station 23) in central Canada (Fig. S4 in the Supplement). At stations affected by the transatlantic dust transport (stations 18 to 21) most models capture the higher AOD in the summer month of June to September. At Surinam (17), in northern South America, a single summer month presents an overestimation, large discrepancy with the observations and large model diversity. At Capo Verde (station 24), offshore western Africa, most models (10) simulated the higher AOD from June to September; however they mainly overestimate the AOD throughout the remainder of the year. In Kanpur (station 25), northern India, on the contrary, models capture the seasonality but the magnitude is mostly underestimated (by 12 of the 15 models).

In the analysis of data for the year 2000 fewer stations are included because the number of available stations for this particular year is smaller. Only 8 AERONET stations from a total of 446 met the requirements of a "dusty" station described in Sect. 2.3 (Table S3 in the Supplement, Fig. 8). Note that we use the same numbers to identify the stations as used in Figs. 9 and 10.

The averaged AOD is again reasonably well simulated by almost all models (Fig. S6 in the Supplement). The simulated AOD is within a factor two of the observed AOD at almost all stations and for almost all models. The MNB varies between −0.38 and 0.4 and the NRMS between 0.1 and 0.5 for all models. The same 8 models that underestimate the climatology also underestimate the data of the year 2000 with MNB between −0.38 and −0.04. While the NRMS presents larger variability between models for data of the year 2000 compared to the climatology, the models produce in general smaller errors in simulating the data of the year 2000. In addition, except for two models, all models produce a larger correlation (R) when simulating the AOD at all stations for the year 2000. An AOD grouping similar to the climatology is observed among the dust regions; African stations yield the largest AOD followed by Middle East and then America. About half of the models (7) reproduce the AOD grouping observed in the three defined regions. Eight models underestimate the AOD at American sites while six of the 15 models overestimate the AOD in all African stations. A few models (4) systematically underestimate the AOD at all or almost all of the dusty stations.

Contrary to what was seen for the climatology where the underestimation coincided mostly with months of maximum AOD, for the year 2000 the AOD is mainly overestimated at all stations and throughout most of the year (Fig. S7 in the Supplement). Exceptions to this are Ouagadougou (4) in Burkina Faso and Surinam (17) in northern South America. Yet the large errors (CPRMS) are observed, as for the climatology, at months where the AOD is overestimated. The maximum CPRMS in fact coincide with the maximum MNB. In addition the largest diversity corresponds to month with overestimation and large errors.

In general most models reproduce the shifting of maximum AOD in African stations from March in Ouagadougou to June in Dakar (8) western Sahara, yet no model reproduces the second maximum in October in Ouagadougou and Banizoumbou (6) in Niger. They either fail to reproduce the second maximum at all, simulate it delayed by one month, or it is too long in duration (see Fig. S8 in the Supplement). All models simulate year-round dust transport off Africa at Capo Verde (24) offshore western Africa mostly overestimating it. While a large number of models simulate the two maxima present in the observations, a few models (4) produce only a single maximum. This last finding may indicate deficiencies in reproducing the mechanism responsible for transporting dust offshore. At the Caribbean-American stations in Barbados and Puerto Rico, all models reproduce the observed transatlantic dust transport as illustrated by the AOD in June and July. Observations suggest that only dust emissions responsible for the maximum in Dakar are transported across the Atlantic. None of the models reproduce the seasonal cycle observed in Surinam in northern South America which shows a maximum AOD in winter months. This winter peak is linked to the seasonal southward displacement of trans-Atlantic African dust plume during winter as seen in various satellite products (e.g. Husar et al., 1997) and characterized by measurements along the coast of French Guiana (Prospero et al., 1981) and over the Amazon (Swap et al., 1992). There is no relationship between the ability of models to reproduce the yearly cycle of AOD over Africa and Caribbean-America and the ability to reproduce the cycle in the Middle East.

3.4 Coarse mode aerosol optical depth

The coarse mode AOD corresponds to the aerosol optical depth of particles with radius larger than 1 μm, i.e. sea salt and desert dust. Its retrieval depends on concurrent multiple-angle sky observations (almucantar and azimuth plane measurements). Because these measurements are often precluded by sky conditions, less coarse mode AOD is retrieved than total AOD which requires only direct sun

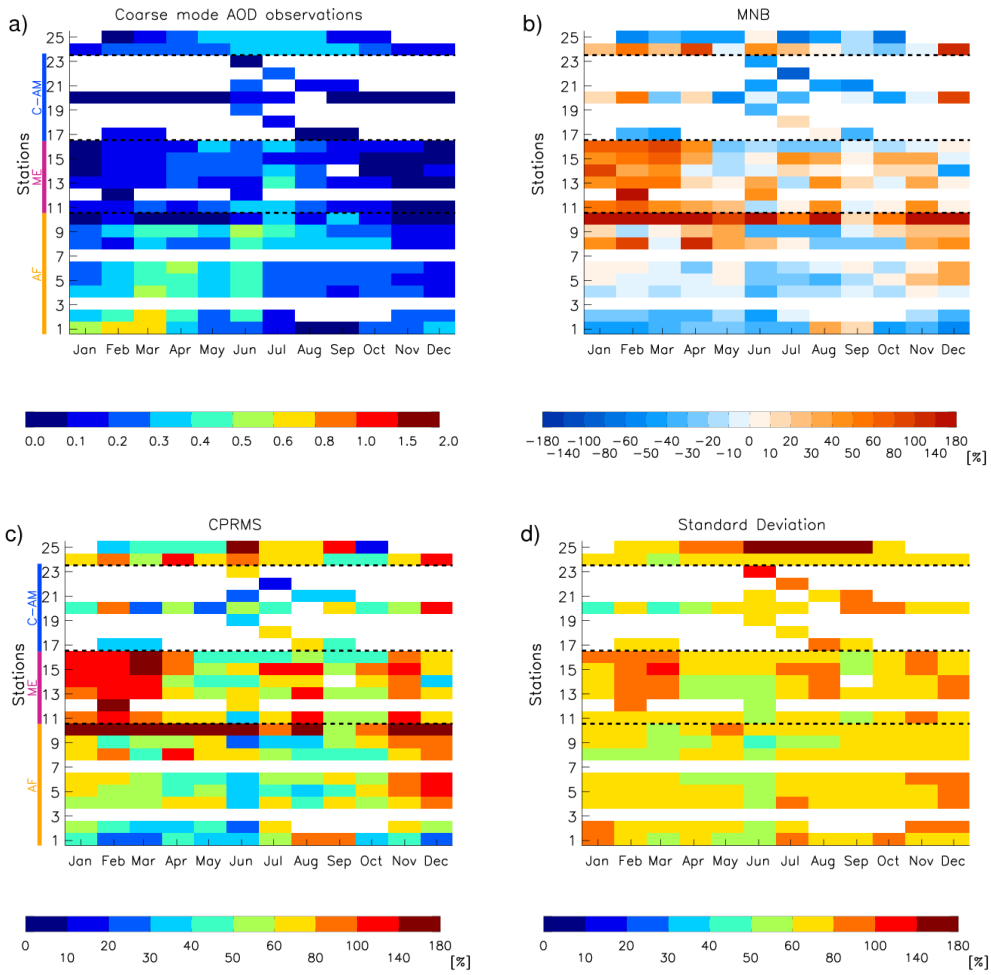

Fig. 11. Same as Fig. 10 but for coarse mode AOD. Same stations as the ones used for total AOD and illustrated in Fig. 8 are considered. Stations Bandoukoui (3) and Bidi Bahn (7) do not have coarse mode AOD for the selected period. For the individual figure of each model presenting the simulated values and their differences (in %) with respect to observations see Figs. S10 and S11, respectively, in the Supplement.

measurements. As a consequence of this difference in number of available measurements, the monthly mean coarse mode AOD can show larger values than the monthly mean total AOD.

The coarse mode AOD climatology (Fig. 11) has a seasonal cycle similar to the total AOD (Fig. 10). Note that stations Bandoukoui (3) and Bidi Bahn (7), both in Burkina Faso, do not have coarse mode AOD measurements and fewer qualifying data for the C-AM stations are available. The coarse mode AOD represents more than half of the total AOD in periods of maximum total AOD, illustrating the dominance of coarse dust particles. The models in general reproduce this dominance of coarse dust particles and produce seasonality similar to the total AOD. However the differences with respect to the observations in terms of Bias, CPRMS and standard deviation are increased compared to the total AOD. The overestimation is in general larger for the coarse mode AOD than the total AOD except for the African stations 1 to 6. In addition to a general increase in the error to reproduce the coarse mode AOD with respect to the total AOD, the large errors are not necessarily associated to an overestimation as was seen with the total AOD. Yet the maximum in CPRMS are still linked to months were the coarse mode AOD is overestimated. Finally, larger model diversity exists for all stations and throughout the year.

The observed coarse mode AOD for the year 2000 (not shown) presents the same features as the climatology in the few qualifying month available. There is a similar seasonality in coarse-AOD as total AOD and a dominance of coarse mode dust particles in months of maximum total AOD. Most models reproduce this seasonality and simulate a dominance of coarse mode particles in periods of maximum AOD. Furthermore, the models present in general larger errors and larger diversity for the coarse mode AOD than the total AOD.

Month overestimating the coarse mode AOD do not necessarily coincide with months where the total AOD is overestimated (not shown).

3.5 Angström Exponent

We now analyse the climatology of the Angstrom Exponent (AE) for dusty sites. Again, we start by analyzing the averaged AE (Fig. 12) and then the seasonal cycle at the 25 stations selected with climatological data (Fig. 13). Next we reproduce this analysis with the 8 stations selected using the data of the year 2000 (Figs. S14 and S15 in the Supplement).

In general the over/under estimation of the models is within a factor of two of the observations (Fig. 12). Yet the errors (NRMS) and bias are larger than for the AOD suggesting that models simulate better the total AOD than the AE and thus reproduce better the aerosol load than the size distribution. The sole exception to this is MATCH that shows larger NRMS for the AOD than the AE. Only four models mainly underestimate the AE, indicating that these models simulate larger particles than is observed. With the exception of MIRAGE, models overestimating the AE produce a smaller bias (MNB from 0.13 to 0.67) than the ones underestimating the AE (MNB from 0.25 to 0.75). However, the opposite is seen for the NRMS; the models underestimating the AE (0.4–0.8) have smaller errors than those overestimating the AE (0.5–0.9). Nine of the 13 models underestimate AE in the Middle East because they simulate larger particles than observed. Nearly all models overestimate the AE in a good number of Caribbean-American stations. Greater diversity is found for simulations of the AE at African stations. Except for stations Ilorin (1) and Djougou (2) in Nigeria and Benin respectively, the measurements in the Middle East show larger AE average than in Africa, thus indicating the predominance of smaller particles in the former. This larger AE could be due to the influence of anthropogenic aerosols and not necessarily dust aerosols only. The AE at the Caribbean-American stations spans the range of values observed in Africa and the Middle East. Recall that only months dominated by coarse dust aerosols or with mixtures of coarse and fine aerosols are analyzed, and that therefore observations-model discrepancies could also be due to anthropogenic aerosols. Only half of the models reproduce this difference in AE between the Middle East and Africa, while ten models simulate the wide range of AE in American stations.

The models mostly overestimate the AE throughout the year or during most of it at stations in Africa, Caribbean-America and elsewhere (Fig. 13). A few models (3) fail to reproduce the AE variability at all stations and produce rather homogenous yearly cycle (Fig. S12 in the Supplement). Only in the Middle East the models also underestimate the AE, this mainly during late fall and winter when mixture of large and fine particles dominate but also partly from March till July when large particles dominate.

Except for the two most southerly stations, the AE in Africa shows that coarse aerosols ($AE < 0.4$) dominate in months with maximum AOD; the coarse mode dominates in spring in southerly stations and shifts progressively to summer and early fall in the northerly stations. This feature is captured by a large number of models (8 out of 13), with most models underestimating the duration of period with small AE. In stations 3 to 9 the coarse aerosol dominance extends beyond the period of maximum AOD. The largest models errors to reproduce the AE are seen in month where the coarse model dominates. The simultaneous overestimation of AE and large model errors in reproducing the AE in periods where the coarse mode dominates, suggest that the models in general simulate too much or too small fine particles. This issue will be addressed in more detail in Sect. 4. The standard deviation reveals that the largest model diversity exists mainly from February till June in stations 2 to 9 mostly coincident with months of coarse mode dominance. Station 1 and 10 show the smallest and largest spread respectively extending throughout the year.

In the Middle East only a few models (6) manage to reproduce the dominance of large particles observed in the month preceding the period of high AOD and the mixture of fine and coarse particles observed during the month of high AOD. In general models overestimate the AE before and during the onset of the period of high AOD and underestimate it afterwards. Except for the stations in Solar Village and Barahin, two periods with large diversity are observed, one in July and August coincident with the months with maximum AOD and another one in March and April coincident with months dominated by coarse mode AE. Solar Village and Barahin present large diversity throughout most of the year. The error with respect to the observations coincides in general with the period of large diversity except for the months of March and April.

Most models simulate the yearly cycle at the American stations 18 to 21 consistent with a dust contribution of large African dust aerosols in the summer months. In contrast models have difficulty in reproducing the relatively small AE observed in Surinam (17) from February to May, as revealed by large errors as well as large overestimation of the AE. However, the models present small model diversity during these months. This large errors and bias suggests difficulties to reproduce the Winter-Spring transport of African dust to South America as described above.

The station at Capo Verde (24) is dominated by large particles throughout the year, illustrating the occurrence of dust transport off the coast of western Africa throughout the year. Models differ from observations mainly in the onset and duration of periods characterized by large particles. Finally most models have difficulties simulating the yearly AE cycle with the dominance of large particles from May to July in the station at Kanpur in Northern India (25).

As for the climatology, for most models the errors (NRMS) and bias in AE are larger than for the AOD.

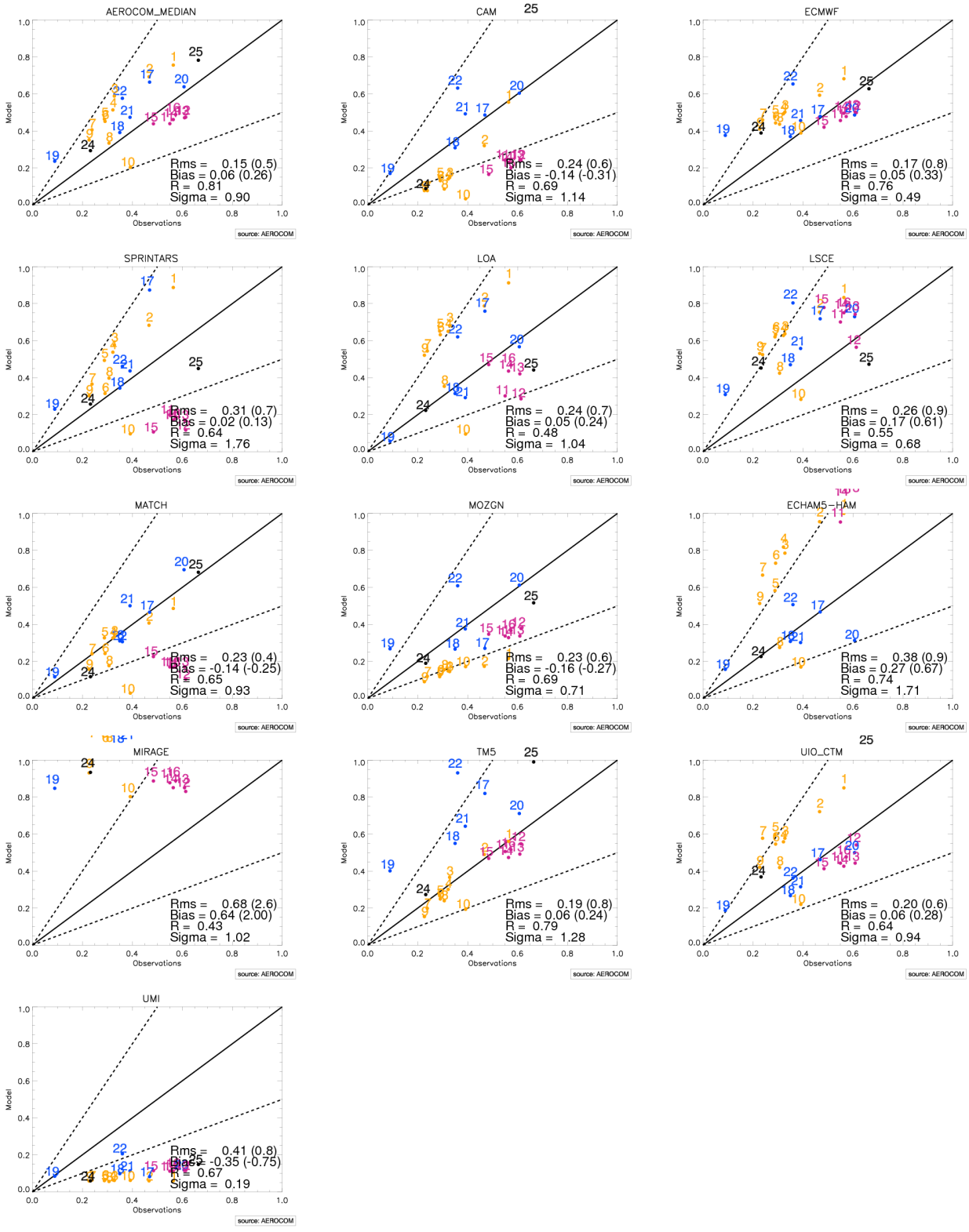

Fig. 12. Same as Fig. 9 but for Angström exponent.

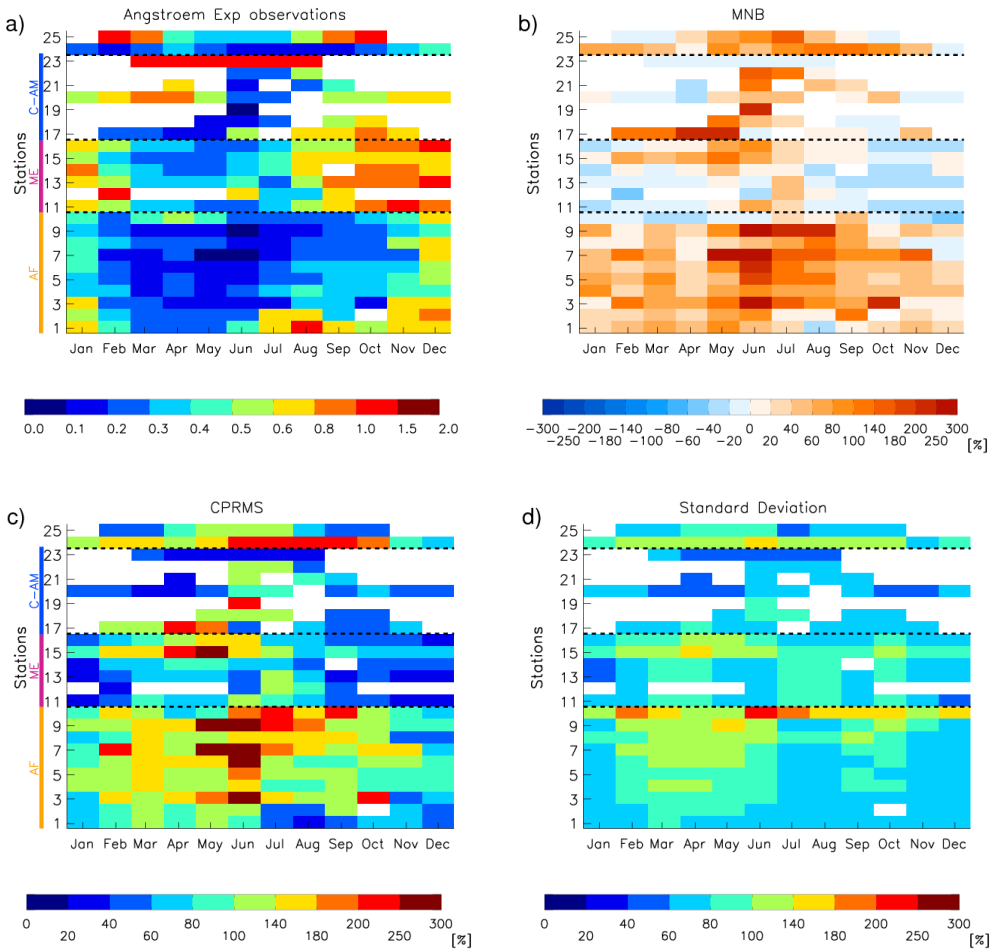

Fig. 13. Same as Fig. 10 but for Angström exponent. For the individual figure of each model presenting the simulated values and their differences (in %) with respect to observations see Figs. S12 and S13, respectively, in the Supplement.

Exceptions are the models MATCH and UIO_CTM where the AE error is smaller and ECMWF, LSCE, TM5 and UIO_CTM where the AE bias is smaller. Furthermore models overestimating the AE (excluding MIRAGE) produce in general smaller MNB than those underestimating it. The same four models yielding a negative bias with the climatology also produce one with 2000 data (Fig. S14 in the Supplement). However contrary to what is seen with the climatology, the models overestimating the AE (excluding MIRAGE) have a smaller NRMS (between 0.2–0.7) than those underestimating it (between 0.3–0.8). The averaged AE for the year 2000 shows that the smallest particles (largest AE) are observed in Solar Village (station 15) in the Middle East and Surinam (17) in northern South America while African stations present values smaller than in the Middle East but larger than the two stations in the Caribbean (Roosevelt Roads, PR and Barbados, WI). This larger AE in Africa than in the Caribbean suggests a greater ratio of large to small particles across the Atlantic than in the source regions. Possible explanations for the larger particles across the Atlantic are the influence of pollution from Europe and biomass burning from the low latitudes and/or the aging of air mass as they cross the Atlantic. In this long range transport small particles are lost due to chemical reactions (growing larger) and to agglomeration during cloud processing. However, the smaller AE average across the Atlantic can also be a numerical artifact due to fewer selected month used in the computation of the average in the Caribbean. While in African stations the average is the result of considering several months that combine large and small AE, in the Caribbean stations fewer months are considered and they are dominated by small AE. In fact the station of Surinam (17) in northern South America with a larger record presents an AE average larger than in African stations. At Capo Verde (24), offshore western Africa, the observed averaged AE is comparable to values observed in Barbados (18) and Roosevelt Roads (21). Eleven of the 13 models reproduce the observed AE for the year 2000 with absolute differences falling within a factor two

of the observations. While most models (10 out of 13) underestimate the AE in the Middle East they produce larger diversity when simulating the AE in Africa and Caribbean-America. However, many models (9) reproduce the AE in Barbados and Roosevelt Roads better than in African Stations.

The annual cycle of AE at dusty stations during the year 2000 has features similar to those seen in the climatology. Contrary to the climatology, for the year 2000 the models mainly overestimate the AE in all regions and throughout most of the year, in particular in periods when the coarse mode dominates (Fig. S15 in the Supplement). In general the models simulate better the year 2000 than the climatology as illustrated by smaller biases and errors. While the model diversity is larger in the C-AM stations for the year 2000 than for the climatology the opposite is true for Solar Village. In African stations the model spread is larger for the year 2000 in Ouagadougou (4) in Burkina Faso while in Banizoumbou (6) and Dakar (8) in general no large differences are observed. As seen with the climatology, a large number of models (7) reproduce the AE seasonality in Barbados (6) and Roosevelt Roads (7) but almost all models fail to reproduce the presence of large particles from February to April in Surinam (5). This yearly cycle is consistent with the southward displacement of the dust transport in winter months described in Ginoux et al. (2001) and as measured in French Guiana (Prospero et al., 1981) and over the Amazon (Swap et al., 1992).

4 Discussion

4.1 Surface variables

Most models simulate the dust deposition measurements within a factor 10 of the observations. Even though all the models produce a positive MNB for the total deposition, models yield both over and under-estimations that vary with the location of the data. While many models overestimate deposition in the Indian Ocean (9 out of 15) and Europe and North Atlantic (8 out of 15), most models underestimate the deposition at remote regions of the Pacific and the South Atlantic Ocean (12 out of 15). Only a few data of total deposition exist in HNLC regions to assess the model performance to reproduce deposition in regions sensible to iron contributions. In addition, the predominant model performance to reproduce deposition in these regions varies depending on the location and dataset considered. While the fluxes near the Antarctica are mostly overestimated, the one in the Southern Ocean is mostly underestimated (station 13 in Fig. 1). Different dust regimes influence each of these sites as indicated by the magnitude of the measured deposition. Difficulties in simulating these dust regimes and the dust transport to remote regions might explain this varying model performances. However, data quality cannot be discarded as source of the difference. Mahowald et al. (2009) points to the errors that can result from estimating the dust fluxes from sediment traps. On the other hand, the Antarctic dust deposition fluxes used in Mahowald et al. (2009) result from measurements of dissolved iron in snow that are known to be too low (Edwards and Sedwick, 2001). In order to reduce the uncertainty associated to the model performance to reproduce the atmospheric iron contributions in HNLC regions, further measurements and model studies are needed. The overestimation in the Northern Hemisphere may suggest a problem in representing the intensity of emissions, the size distribution of the transported dust, the transport itself and/or the representation of deposition flux. At present because of data limitations it is not possible to link the differences between models and observations to any specific process. When comparing the models against long-term measurements of total and wet deposition taken in Florida (Prospero et al., 2010), models capture the seasonality of the deposition and the dominance of wet deposition but most underestimate the magnitude. Furthermore, the performance deteriorates from south to north. These differences could be due to difficulties in simulating the northward transport of dust or the removal processes.

Observations suggest that wet deposition dominates over dry deposition over most ocean and remote regions of the world (Mahowald et al., 2011). Models are able to capture this dominance of wet deposition, but tend to overestimate it at many locations, especially in those where it is not the dominant removal process (Table 4). We agree with Wagener et al. (2008), Mahowald et al. (2009) and Prospero et al. (2010) that more measurements of deposition fluxes are needed, in particular in the HNLC regions of the Southern Hemisphere, to better estimate the atmospheric iron contribution into the oceans. Ideally such measurements should extend for a year or more considering that the large fraction of the annual deposition occurs in episodic events of just a few days (Prospero et al., 2010; Mahowald et al., 2009). In addition, these measurements should also split between wet and total deposition, as done in Prospero et al. (2010), considering the uncertainty of the contribution of wet deposition in total deposition over ocean (Jickells et al., 2005). However it should be noted that there is a severe problem in measuring dry deposition. The use of buckets or surrogate surfaces as collectors does not reflect real world conditions; the aerodynamics of these collectors and their surface properties are very different from natural surfaces such as bare soil, grasses or the ocean surface (Prospero et al., 2010). As a result dry deposition is typically calculated based on particle size distributions; such estimates are prone to large uncertainties which are typically quoted as plus/minus a factor of three (Duce et al., 1991) but which could well be larger. In the meantime, before such long-term measurements are available alternative techniques to evaluate deposition may be necessary. One such method is to simulate the deposition and advection of dissolved aluminium in the surface ocean and to compare against surface ocean aluminum measurements (Han et al., 2008). These

could also be inverted to estimate deposition. But this technique is complicated by uncertainties in the solubility of dust aluminium and the properties of the dust itself.

The model performance in simulating surface concentration depends on the data sets used. For example when using measurements from cruises, all models agree in mainly overestimating the surface concentration by mostly a factor of ten up to a hundred whereas the underestimation is mostly limited to a factor ten. The cases with large overestimation correspond mainly to short-term cruise measurements in regions downwind of the main dust sources. However, for the cruise measurements in remote regions of the Southern Hemisphere the models perform equally well as they do against long-term measurements in other regions. When using long-term measurements of the SEAREX and AEROCE network, on the other hand, the overestimation is within a factor of ten with respect to the observations. It has to be noted however that the cruise measurements correspond to short-term measurements and if the sampling error due to missing dust events is taken into account the large overestimation is reduced (up to 96 %) and the performance resembles the one observed with long-term measurements. In spite of the large uncertainties, these observations deliver valuable information in remote regions that are seldom sampled (e.g. the southern Ocean and South Atlantic ocean). While all models agree in overestimating the cruise data in the Indian Ocean, large model diversity exist in simulating the surface concentration in the South Atlantic varying from some models overestimating the observations, other underestimating it and some of the models both over and under estimating the surface concentration. Much of this region is characterized as HNLC; consequently dust deposition can have a great impact in the biogeochemical cycle.

Recall that for both surface concentration and deposition the period when the data were taken is not coincident with the simulated year, a factor which could explain part of the model-observation differences since most models constrain the dust cycle with reanalyzed winds of the year 2000 (Table 1). However, the large over/under estimation by most models points to other issues. Because of the episodic nature of dust events and the few days in which they occur (Prospero et al., 2010; Mahowald et al., 2009), short-duration measurements risk missing dust events and should therefore be applied with care for model evaluation.

Particle size is also an important factor and a source of discrepancies when comparing deposition and/or surface concentration to model outputs. The representation of size distribution of mineral dust is a fundamental parameter to simulate and understand its impact; while the fine mode controls the direct impact on radiation and cloud processes, the coarse mode governs deposition and hence its biogeochemical impact (Formenti et al., 2010). Variables integrated over all size classes, as available for this study, prevented us from exploring the impact of the different representation of the size distribution in each model on its performance in simulating the different observations. Therefore, knowledge of the size distribution of both measurements and model would allow a more in-depth model evaluation and assessment of its performance. We therefore suggest that size-resolved surface concentration and deposition be archived in future model experiments.

4.2 Vertically integrated variables

The models reproduce the retrieved AOD and AE within a factor of two. Furthermore, most models present a better performance in simulating the total aerosol load than the size distribution of dust particles as revealed by smaller errors and bias associated to the averaged AOD. In general in Africa and Caribbean-America the models underestimate the AOD climatology in months of maximum AOD and overestimate the AE throughout most of the year. While the models present the largest errors in AOD mainly in months with low values the largest error in reproducing the AE occur in months of maximum AOD dominated by coarse particles. In contrast to stations in Africa and Caribbean-America, in the Middle East models not only overestimate the AOD during month with maximum AOD but also in the month preceding it. The AE is overestimated before and during (May to July) the onset of the period of high AOD and underestimated the rest of the year. In general the models present larger diversity in simulating the AOD in the Middle East than in Africa. When compared to the year 2000, both AOD and AE are mostly overestimated in all considered dust regions.

Models capture the transport of dust across the coast of West Africa to the Atlantic throughout the year as illustrated by comparisons with measurements at Capo Verde, located 600 km to the west of the African coast. The models also reproduce the trans-Atlantic dust transport as characterized by measurements at Capo Verde, and Barbados, 4000 km to the west. While all models reproduce the AOD seasonality in Barbados, only 7 reproduce the seasonality of AE in this station.

The trans-Atlantic dust plume undergoes a seasonal displacement that is linked to movements of the Intertropical Convergence Zone (ITCZ). During the boreal summer the ITCZ reaches its most northern position, and winds carry dust to the Caribbean. During the winter the ITCZ reaches its southernmost position, and dust is carried to South America (Prospero et al., 1981; Swap et al., 1992; Ginoux et al., 2001). This seasonal transport cycle is reflected in the AOD record in Surinam (northern South America) which has a minimum in the summer at the time when the AOD at Barbados reaches the annual maximum. Most models successfully simulate the AOD seasonal cycle in Barbados but they do not reproduce the minimum AOD confined to the summer month in Surinam. This might indicate problems in simulating the general circulation in the tropics and/or removal process coincident with this southward shift of the transatlantic dust cloud.

Recall that the AOD and AE analysis is based on months dominated by coarse aerosols or the mixture of fine and coarse aerosols. Therefore the discrepancies between observations and model can also be explained by the influence of anthropogenic aerosols. However, since in African, Caribbean-American and Other stations the months of maximum AOD are also characterized by coarse particles we are confident that the atmospheric aerosols, at least in these months, are dominated by desert dust and therefore the model performance is associated to the models ability to simulate the dust cycle. In the Middle East, in contrast, the period of maximum AOD is influenced by both large particles and mixtures of fine and coarse particles and these fine particles are most likely due to the presence of anthropogenic aerosols. Eck et al. (2008), in studies in a network of 14 AERONET photometers in the United Arab Emirates, observed increases in AE coincident with the presence of increased concentrations of fine particles which they attributed to sources in the petroleum industry.

4.3 Emissions

There are no datasets of measured dust emissions that could be used in this study. Still, evaluation of the simulated combination of AOD and AE allows us to make inferences about the simulated emissions. Since the scattering efficiency varies according to the size, the AOD is not only dependent on the aerosol burden but also on the size distribution; smaller dust aerosol particles scatter light more efficiently than larger ones, i.e. for the same burden air masses containing higher concentrations of smaller particles will yield larger AOD. Based on the latter factor, the combination of AE and AOD measurements can be used to infer whether the emissions are over- or under-estimated. To illustrate this, let's suppose that a model simultaneously overestimates the AOD and underestimates the AE close to the source. In order to increase the AE and thus reduce the underestimation, a larger fraction of fine particles is necessary. This can be achieved by either augmenting the emissions of fine mode particles, which would increase even more the AOD, or by reducing the emissions of coarse particles, leading to a reduction of the AOD. Therefore, a simultaneous overestimation of the AOD and underestimation of the AE points to an overestimation of the mass emissions especially of the coarse dust particles if interference from other aerosol components can be excluded. Likewise, the simultaneous underestimation of the AOD and overestimation of the AE, points to an underestimation of the coarse dust emissions. In both cases, however, fine mode dust emission adjustments might additionally be needed. Simultaneous over- or underestimation of both AOD and of AE precludes inferring whether the intensity of the source has been over- or under-estimated. We need to improve the simulation of the dust size distribution in models before we can attempt to quantify adjustments to emissions.

We present in Fig. 14 the results of applying the above considerations to the comparison with the AERONET data. It should be noted that for the judgement on the over- and under-estimation of the emissions based on the AOD and AE other simulated processes might be responsible such as sedimentation, wet deposition, dry deposition, horizontal and vertical transport. These processes have a lesser impact on stations close to the sources than remote ones; impacts due to errors in simulating the above mentioned processes near the sources would be most likely amplified during long-range transport. We therefore focus our present analysis on AERONET data of the year 2000 from the African and Middle East sites and exclude Caribbean-American stations (Figs. S7 and S15 and corresponding figures in the Supplement). According to those figures the AeroCom median and models ECMWF, SPRINTARS, LSCE and ECHAM5-HAM underestimate the dust emissions in Africa while the CAM model overestimate them in this region (Fig. 14a). For the other models, either the AE was not available or the results were not conclusive enough to propose an over/under estimation of the emissions. In the Middle East, the models LSCE and ECHAM5-HAM underestimate the dust emissions while models CAM, MATCH, MOZGN, UMI and SPRINTARS overestimate them (Fig. 14b). Note that the analysis on the Middle East is based only on the station at Solar Village. This station has been documented as affected by dust particles from the deserts in the region (Sabbah and Hasan, 2008).

The regional emissions were computed for each model (Table 5). The regions are illustrated in Fig. 2 and a few models exist that consider desert dust sources outside these regions. The models under/overestimating the emissions are highlighted in blue/red in Table 5. When comparing the emission fluxes between models it is important to consider the simulated size distribution because coarse mode aerosols will dominate the emission (in terms of mass) but will have little impact on the AOD (at 550 nm) and conversely fine mode aerosols will dominate the AOD (at 550 nm) but have smaller impact on the emission mass. Furthermore, factors such as mass extinction efficiency (MEE) and aerosol lifetime should also be considered when comparing emissions between different models. According to the results in Fig. 14 and Table 5, SPRINTARS has larger emissions than CAM in Northern Africa even though CAM overestimates the emissions in this region while SPRINTARS underestimates them. Although this might appear contradictory, it is consistent with the short lifetime of dust particles in SPRINTARS. According to its lifetime (1.6 days), particles are removed shortly after emission and an important fraction probably even before arriving to the measuring site. The apparent underestimation is therefore consistent with the fact that particles are removed too fast from the atmosphere. In the Middle East on the contrary, both models (CAM and SPRINTARS) overestimate the emissions suggesting that dust particles are transported to Solar Village before their removal and thus overestimating the emissions.

Table 5. Yearly emission fluxes [Tg yr^{-1}] for regions illustrated in Fig. 2. Fluxes being overestimated are highlighted in red and those underestimating are highlighted in blue. Models have been grouped according to their size ranges; models GISS to UMI simulate dust aerosols in the size range 0.1–10 μm, models ECMWF and LOA in the size range 0.03–20 μm and UIO_CTM in the range 0.05–25 μm. Models LSCE, TM5, ECHAM5-HAM and MIRAGE describe dust aerosols through 1, 2, 2 and 4 modes respectively. For mass mean radius and standard deviation of each mode see Table 1.

	North Africa	Middle East	Asia	South America	South Africa	Australia	North America
CAM	2271	526	727	13.7	2.9	12.2	286
GISS	1031	125	180	39.9	31.7	87.8	7.3
GOCART	1736	348	873	66.5	25.0	111	13.0
SPRINTARS	2888	531	363	6.9	113	36.8	4.1
MATCH	539	241	100	19.3	24.5	40.9	2.4
MOZGN	1410	376	294	92.8	55.4	89.5	12.7
UMI	933	329	340	47.1	20.6	35.4	6.1
ECMWF	204	68	125	1.0	16.3	57.0	15.3
LOA	772	114	411	0.5	3.5	14.9	4.5
UIO_CTM	1213	206	27	5.0	11.6	9.0	1.8
LSCE	529	39.2	509	0.2	57.2	10.6	7.2
ECHAM5-HAM	401	25.6	54	3.7	40.2	58.4	1.7
MIRAGE	703	292	608	186	25.0	129	70.8
TM5	1091	212	253	30.4	15.3	59.4	8.1
AEROCOM MEDIAN	792	128	137	9.8	11.8	30.7	2.0

We decide to exclude SPRINTARS from the following analysis in view of the uncertainty in the emissions associated to a short lifetime. Based on the above results we suggest that a range of plausible emission for North Africa is 400 to 2200 Tg yr^{-1} while in the Middle East the range of plausible emissions is between 26 and 526 Tg yr^{-1}. We note however that emission fluxes outside these ranges can be possible depending on the definition of parameters such as size distribution, lifetime and MEE.

4.4 General discussion

Because there was no AERONET station affected by the Asian dust sources which met the criteria used in this study, we could not evaluate the performance of the models in simulating the dust cycle in Asia. Months with intense dust activity were masked by anthropogenic emissions which generated AE values above 0.4 and therefore were not recognisable with our definition of dust sites. However, surface concentration measurements in Midway in the Northern Pacific (station 15 in Fig. 2) and Hedo and Cheju (stations 20 and 22 respectively in Fig. 2) in eastern Asia, even though limited, give us some insight in the general model performance in simulating the Asian dust. As described in Sect. 3.2, in general the models reproduce annual cycles in these sites mostly underestimating the observations in Hedo and Cheju while in Midway the models mostly underestimate the observations in spring and overestimate them in months following the spring peak. A few models exist that largely overestimate the surface concentration at these sites. In periods of maximum concentrations the simulated values of a large number of models is within the observed variability (Fig. 15). The differences between models and observations however could be due to the nature of the data; they are considered as climatology in this study even if they do not qualify for it in a strict sense

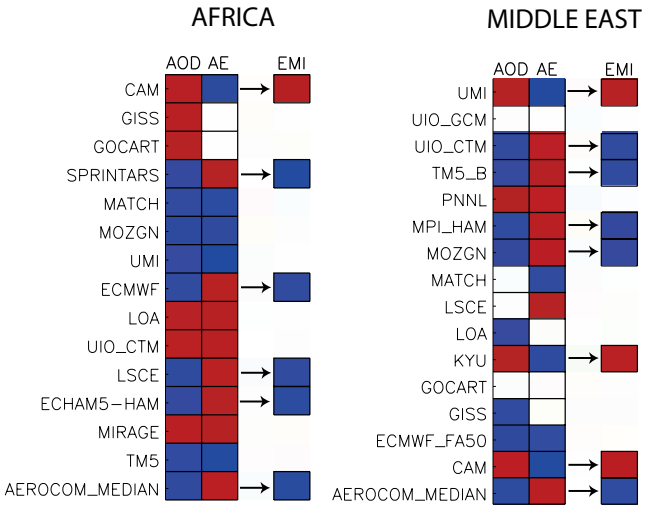

Fig. 14. Suggested resulting over/under estimation (EMI) of the emissions in Africa (left panel) and the Middle East (right panel) based on AERONET Angström Exponent (AE) and aerosol optical depth (AOD). Simultaneous overestimation of the AOD and underestimation of the AE suggests an overestimation of the emissions and vice versa. Overestimations in a given model are illustrated by red color, whereas underestimations are indicated by blue color.

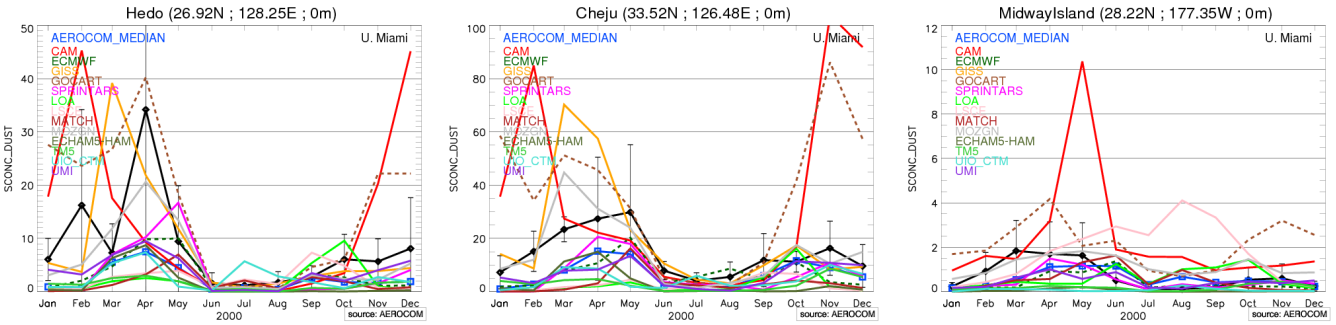

Fig. 15. Measured and simulated surface concentration in the Asian stations of Hedo and Cheju and the Pacifica station in Midway Island. Measurements are presented by the black continuous line. Units are $\mu g\,m^{-3}$.

and the measurement period does not coincide with the year simulated. In addition, the comparisons of surface concentrations at stations affected by Saharan (Barbados, Bermuda and Miami) and Asian dust (Hedo, Cheju and Midway) reveal that in general models have not only smaller biases and errors when reproducing the annual cycle at Asian stations but also present smaller model diversity. This in spite of the fact that dust emissions in global models are generally tuned to fit the observations in a given part of the world and often this tuning is done with observations from North Africa. Because we have no AOD and AE data for the Asian deserts and because of the climatological aspect of the surface concentration data, we cannot assess whether this difference in performance is also observed in other aspect of the dust cycle. A more specific Asian dust data set is needed to address this issue and examine the role of the tuning in the performance of global dust models. We therefore excluded Asia from this study. One way to assess the performance of global dust models over Asia would be to compare measurements of coarse mode AOD against modelled ones.

The models perform better (smaller errors and biases) in simulating the climatology of vertically integrated variables in dusty sites than they do with deposition and surface concentration measurements. The modeled AOD is within a twofold range of the observations at most sites, whereas model under/overestimations of surface concentrations and total deposition are more typically within a range of a factor of 10. Differences in the data can explain this since the AERONET climatology includes the simulated years whereas the deposition and surface concentration climatology do not. The surface measurements were considered as climatology in this study although, in a strict sense of the term, they were not. Furthermore, surface concentration and deposition require that the model correctly simulates the vertical distribution whereas for vertically integrated parameters such as AOD and AE the vertical distribution is less relevant (assuming that they are clear-sky measurements of non-hygroscopic particle such as dust). In addition, this difference in performance might also suggest that AeroCom models (as used in experiment A) are more adequate to assess the radiative impact of dust aerosols than their impact on air quality and/or the biogeochemical cycle.

Throughout the text when comparing the models to each variable and in the consequent analysis, we treated the AeroCom median as any other model even though it is not a real one but a construction from multiple state-of-the-art AeroCom A models. For the integrated variables of AOD (for the year 2000) and AE (year 2000 and climatology) the AeroCom median is among the models with the smallest MNB and NMRS, in some cases even the one with the smallest value. Both MNB and NRMS correspond to the analysis of the averaged values and therefore do not reflect the model performance on the annual cycle. These AeroCom statistics suggest that random error might cancel out when computing the median. By construction, the AeroCom median has the same deficiencies present in most models such as the difficulties to reproduce the fraction of wet deposition when dry deposition dominates and to simulate the transport of Saharan dust to Surinam in northern South America during winter months.

This is the first multi-parameter and multi-model intercomparison of global dust models. Fifteen models from the AeroCom project have been compared to different and multiple datasets. The models were examined in their performance to simulate surface variables such as deposition and dust concentration and the vertically integrated variables of AOD and AE. A recurrent problem when evaluating the performance of a dust model is the data available to do it. A benchmark dataset has been created containing all the information used in this work and available through the AeroCom data server. There are various datasets that have been used for model evaluations (e.g. Prospero et al., 2010; Prospero and Lamb, 2003; Ginoux et al., 2001; Mahowald et al., 2009; and the DIRTMAP data set). These studies concentrated mostly on a single parameter. We have grouped in a single database the data used in these studies to ease future comparison and evaluations.

To further improve this benchmark dataset, additional deposition and surface concentration measurements are needed. Long term measurements of total and wet deposition are required, in particular over remote regions in the Southern Hemisphere where the greatest model diversity is observed and where the role of the atmospheric iron in the ocean biogeochemistry is still under debate (Jickells et al., 2005). With respect to surface concentration, additional surface concentration measurements are needed such as the ones taken during the SEAREX and AEROCE campaigns and those still being measured at Barbados and Miami. Since AOD is dominated by the fine mode due to its higher extinction efficiency and since the coarse mode dominates the surface concentration and deposition, it is important that future measurements as well as model simulations deliver size-resolved information. The absence of this information, in both data and model, prevented us from gaining more insight on the model performance and identifying the possible role of the size distribution in models in the over- and under-estimation of deposition and surface concentration. The AERONET network represents a crucial source of data in validating models. The information of this network should be complemented with satellite products to further evaluate the model performance. The contribution of vertically resolved active measurements from the in-situ Micro-Pulse Lidar Network (MPLNET) and/or from the remote sensor Cloud and Aerosol Lidar with Orthogonal Polarization (CALIOP) onboard the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) would provide valuable information on the vertical distribution of dust aerosols. This information would provide additional constrain to model evaluation and would allow to assess and understand present difficulties to simulate surface variables.

5 Conclusions

Desert dust plays an important role in the climate system through its impact on the earth radiative budget and its role in the biogeochemical cycle as a source of iron in high-nutrient-low-chlorophyll regions. However, there are large differences in the way many global models simulate the dust cycle and the resulting impact of dust on climate. On the one hand, these differences are the product of the various distributions in dust burden and aerosol optical depth which translate into uncertainties in the estimation of the direct radiative effect (Textor et al., 2006; Forster et al., 2007). On the other hand, they result from differences in simulated dust deposition fluxes, which prevents one from properly estimating the impact of dust on ocean CO_2 uptake in HNLC regions (Textor et al., 2006; Tagliabue et al., 2009).

Here we present the results of the first multi-parameter and multi-model intercomparison of a total of 15 global aerosol models within the AeroCom project. Each model is compared to the same set of observations, focusing on variables that have a direct link to the estimation of the direct radiative effect and the dust impact on the biogeochemical cycle, i.e. aerosol optical depth (AOD) and dust deposition. To extend the assessment of model performance we include additional comparisons to Angström exponent (AE), coarse mode AOD and dust surface concentration. Altogether these comprise a new benchmark data set which is available via the AeroCom data server for model inspection and future development of dust models.

Note that the model results used in the present analysis correspond mostly to a coherent set of AeroCom simulations submitted before the year 2005. Many of these models have been changed and are likely improved since submitting their simulations. Therefore the results presented in this study do not necessarily represent the current state of the models.

The models simulate the yearly dust deposition within a factor 10 with respect to the observations. While the deposition is mostly overestimated in Europe, North Atlantic and the Indian Ocean, it is mostly underestimated in the Pacific and South Atlantic Ocean. The limited number of deposition data in HNLC regions and the dependence of the models performance in simulating these data to the location of the data prevent us from concluding on the atmospheric iron contributions in HNLC regions from global dust models. Further measurements and model studies are necessary to address this issue and to assess whether the impact of dust on the ocean biogeochemical cycle in the southern ocean is over- or under-estimated in most models.

In terms of wet and total deposition, models capture the seasonality of the deposition and the dominance of wet over dry deposition in Florida but most underestimate the magnitude. Furthermore, the performance deteriorates from south to north Florida, reflecting difficulties in reproducing the northward dust transport. Data on wet deposition fraction shows that models manage to reproduce the fraction of wet deposition in regions where the wet deposition dominates but fail to do so in sites dominated by dry deposition. Long-term measurement records are needed, ideally on a daily basis and over oceans, to evaluate model ability to reproduce the deposition fluxes. While it is relatively easy to collect and measure wet deposition, there is no easily implemented technique for measuring dust dry deposition to natural surface, in particular the ocean. Thus it is unlikely that we will soon be able to test model dry-deposition simulations in a meaningful way.

The model performance in simulating surface concentration depends on the database used. All models mainly overestimate the surface concentration measured during cruise campaigns mostly by a factor of ten up to a hundred. When using long-term measurements, on the other hand, the overestimations are within a factor of ten with respect to the observations. If the sampling error of missing dust events during short-term cruise measurements is taken into account the large overestimation is reduced and the performance resembles the one observed with long-term measurements. Despite

this large uncertainty, surface observations deliver valuable information in remote regions of the Southern Oceans and South Atlantic Ocean where data are scarce. For both datasets, cruise campaign as well as long-term, all models underestimate the surface concentration within a factor of ten with respect to the observations.

Model performance is better at sites with a large range of measured surface concentrations, reflecting better agreement at stations directly downwind of the main sources than at those in remote regions. The transatlantic dust transport, captured by stations on both sides of the Atlantic, is reproduced by most models. The models coincide in the onset of the period of maximum surface concentration. However they differ in simulating the magnitude of the measurements in this period and its extension in time. For the Pacific stations exposed to Asian dust, most models simulate the general seasonal variations underestimating the observation in months with maximum surface concentration.

A similar conclusion on the regional performance of the models, not contradictory to the above, can be reached based on comparison to the sun photometer data. The models simulate in general the gradient in AOD and AE between the different dusty regions. However the models show less skill in reproducing the magnitude and seasonality in the Middle East of both AOD and AE. Model performance in reproducing Asian dust could not be explored due to the definition of dusty sites used in the study; months with intense dust activities co-incided with AE values above 0.4, influenced by anthropogenic emissions, were masked out. A different selection criteria or approach would be needed to examine the performance of global dust models in this region. Like for surface concentrations, the models reproduce the trans-Atlantic dust transport from the Sahara in terms of AOD and AE. All models reproduce the offshore transport of Saharan dust throughout the year as revealed by data from Capo Verde offshore western Africa. Also they limit the transport across the Atlantic to the Caribbean to the summer months in agreement with measurements at Barbados and Roosevelt Roads, Puerto Rico; however they overestimate the AOD and they transport too fine particles. In contrast, almost no model reproduces the southward displacement of the trans-Atlantic Saharan dust plume during the Winter and Spring as captured by the AOD and AE data at Surinam, which are representative of the dust transport into South America and which has been well documented by various satellite products and by ground-based aerosol measurements.

Models perform better in simulating the climatology of averaged vertically integrated parameters (AOD and AE) in dusty sites than total deposition and surface concentration reflected by smaller MNB and NRMS for AOD and AE than for surface variables. The averaged AOD and AE are within a factor of two of the observations at most sites; in contrast the long-term surface concentrations and total deposition are under- and over-estimated within a factor 10 of the observations. This difference might be explained by the different characteristics of the data climatologies used, as well as the simulated vertical structure important for reproducing dust deposition and surface concentration.

Based on the dependency of AOD and AE on aerosol burden and size distribution we use the simultaneous overestimation or underestimation of AOD and underestimation or overestimation of AE to suggest whether a model is overestimating or underestimating dust emissions. Note, that if AOD and AE bias in a given model is of equal sign then no conclusion with respect to emissions can be made. From this analysis we suggest that the AeroCom median model and models ECMWF, LSCE and ECHAM5-HAM underestimate the emissions in Africa while CAM overestimates them. In the Middle East the models LSCE and ECHAM5-HAM underestimate the emissions, whereas models CAM, MATCH, MOZGN and UMI overestimate them. According to these results we suggest that a range of possible emissions for North Africa is 400 to 2200 $\mathrm{Tg\,yr^{-1}}$ and in the Middle East 26 to 526 $\mathrm{Tg\,yr^{-1}}$. Emission fluxes outside these ranges might be possible however depending on the definition of relevant parameters.

The AERONET data and satellite products are important data sources in aerosol model evaluation, but need to be complemented with deposition data in order to properly evaluate the overall dust cycle included in models. Dust deposition measurements are sparse and deliver mostly only total deposition fluxes for a given event or a longer time period not necessarily coincident with the year simulated, thus limiting the model evaluation. The proper testing of models requires the permanent monitoring of dust deposition in a manner similar to that in the network presented in Prospero et al. (2010) and of dust concentrations (Prospero and Lamb (2003).

The new round of experiments conducted within AeroCom Phase II with additional diagnostics, including a multi-year hindcast with observed meteorology, will allow conducting further comparisons to assess the model performance to simulate the dust cycle. Notably, the detailed size distribution information stored in the new experiments will allow addressing issues such as the impact of the simulated size distribution in reproducing the deposition flux and surface concentration. This information was not available from experiments A and B from the Phase I of AeroCom and prevented us from addressing the role of size in explaining the different model performances in reproducing the deposition and surface concentration. In addition to archiving the size-resolved surface concentration and deposition, we recommend also archiving concentrations above the surface at a few locations in order to allow comparisons in elevated mountain stations. In order to further evaluate the model performance, the AERONET data should be complemented with satellite products, notably the vertically resolved information provided by CALIOP and MPLNET.

Supplementary material related to this article is available online at: http://www.atmos-chem-phys.net/11/7781/2011/acp-11-7781-2011-supplement.pdf.

Acknowledgements. The authors would like to thank two reviewers for their useful comments that contributed to improve the manuscript. In addition we thank the AERONET program for establishing and maintaining the used sites. This study was co-funded by the European Commission under the EU Seventh Research Framework Program (grant agreement No 218793, MACC). O. Boucher was supported by the Joint DECC and Defra Integrated Climate Programme, DECC/Defra (GA01101). S. Ghan and R. Easter were funded by the US Department of Energy, Office of Science, Scientific Discovery through Advanced Computing (SciDAC) program and by the NASA Interdisciplinary Science Program under grant NNX07AI56G. The Pacific Northwest National Laboratory is operated for DOE by Battelle Memorial Institute under contract DE-AC06-76RLO 1830.

Edited by: M. Kanakidou

The publication of this article is financed by CNRS-INSU.

References

Arimoto, R., Duce, R. A., Ray, B. J., Ellis, W. G., Cullen, J. D., and Merrill, J. T.: Trace-Elements in the Atmosphere over the North-Atlantic, J. Geophys. Res.-Atmos., 100(D1), 1199–1213, 1995.

Aumont, O., Bopp, L., and Schulz, M.: What does temporal variability in aeolian dust deposition contribute to sea-surface iron and chlorophyll distributions?, Geophys. Res. Lett., 35(7), L07607, doi:10.1029/2007GL031131, 2008.

Balkanski, Y., Schulz, M., Moulin, C. and Ginoux, P.: The formulation of dust emissions on global scale: formulation and validation using satellite retrievals, in: Emissions of Atmospheric Trace Compounds, edited by: Granier, C., Artaxo, P., and Reeves, C., 239–267, Kluwer Academic Publishers, Dordrecht, 2004.

Balkanski, Y., Schulz, M., Claquin, T., and Guibert, S.: Reevaluation of Mineral aerosol radiative forcings suggests a better agreement with satellite and AERONET data, Atmos. Chem. Phys., 7, 81–95, doi:10.5194/acp-7-81-2007, 2007.

Bauer, S. E. and Koch, D.: Impact of heterogeneous sulfate formation at mineral dust surfaces on aerosol loads and radiative forcing in the Goddard Institute for Space Studies general circulation model, J. Geophys. Res.-Atmos., 110(D17), D17202, doi:10.1029/2005JD005870, 2005.

Bauer, S. E., Balkanski, Y., Schulz, M., Hauglustaine, D. A., and Dentener, F.: Global modeling of heterogeneous chemistry on mineral aerosol surfaces: Influence on tropospheric ozone chemistry and comparison to observations, J. Geophys. Res.-Atmos., 109(D2), D02304, doi:10.1029/2003JD003868, 2004.

Bellouin, N., Boucher, O., Haywood, J., and Reddy, M. S.: Global estimate of aerosol direct radiative forcing from satellite measurements, Nature, 438(7071), 1138–1141, 2005.

Benedetti, A., Morcrette, J. J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J., and Suttie, M.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, J. Geophys. Res.-Atmos., 114, D13205, doi:10.1029/2008JD011115, 2009.

Berglen, T. F., Berntsen, T. K., Isaksen, I. S. A., and Sundet, J. K.: A global model of the coupled sulfur/oxidant chemistry in the troposphere: The sulfur cycle, J. Geophys. Res.-Atmos., 109(D19), D19310, doi:10.1029/2003JD003948, 2004.

Bian, H. S. and Zender, C. S.: Mineral dust and global tropospheric chemistry: Relative roles of photolysis and heterogeneous uptake, J. Geophys. Res.-Atmos., 108(D21), 4672, doi:10.1029/2002JD003143, 2003.

Cakmur, R. V., Miller, R. L., Perlwitz, J., Geogdzhayev, I. V., Ginoux, P., Koch, D., Kohfeld, K. E., Tegen, I., and Zender, C. S.: Constraining the magnitude of the global dust cycle by minimizing the difference between a model and observations, J. Geophys. Res.-Atmos., 111(D6), D06207, doi:10.1029/2005JD005791, 2006.

Cheng, T., Peng, Y., Feichter, J., and Tegen, I.: An improvement on the dust emission scheme in the global aerosol-climate model ECHAM5-HAM, Atmos. Chem. Phys., 8, 1105–1117, doi:10.5194/acp-8-1105-2008, 2008.

Chiapello, I., Moulin, C., and Prospero, J. M.: Understanding the long-term variability of African dust transport across the Atlantic as recorded in both Barbados surface concentrations and large-scale Total Ozone Mapping Spectrometer (TOMS) optical thickness, J. Geophys. Res.-Atmos., 110(D18), D18S10, doi:10.1029/2004JD005132, 2005.

Chin, M., Rood, R. B., Lin, S. J., Muller, J. F., and Thompson, A. M.: Atmospheric sulfur cycle simulated in the global model GOCART: Model description and global properties, J. Geophys. Res.-Atmos., 105(D20), 24671–24687, 2000.

de Meij, A., Krol, M., Dentener, F., Vignati, E., Cuvelier, C., and Thunis, P.: The sensitivity of aerosol in Europe to two different emission inventories and temporal distribution of emissions, Atmos. Chem. Phys., 6, 4287–4309, doi:10.5194/acp-6-4287-2006, 2006.

Denman, K. L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P. M., Dickinson, R. E., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., da Silva Dias, P. L., Wofsy, S. C., and Zhang, X.: Couplings between changes in the clymate system and biogeochemistry, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averty, K. B., Tignor, M., and Miller, H. L.: Cambridge University Press, Cambridge, UK, and New York, NY, USA, 2007

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.

Dubovik, O., Holben, B., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanre, D., and Slutsker, I.: Variability of absorption and optical properties of key aerosol types observed in worldwide locations, J. Atmos. Sci., 59(3), 590–608, 2002.

Duce, R. A., Liss, P. S., Merrill, J. T., Atlas, E. L., Buat-Menard, P., Hicks, B. B., Miller, J. M., Prospero, J. M., Arimoto, R., Church, T. M., Ellis, W., Galloway, J. N., Hansen, L., Jickells, T. D., Knap, A. H., Reinhardt, K. H., Schneider, B., Soudine, A., Tokos, J. J., Tsunogai, S., Wollast, R., and Zhou, M.: The atmospheric input of trace species to the world ocean, Global Biogeochem. Cy., 5(3), 193–259, 1991.

Eck, T. F., Holben, B. N., Reid, J. S., Sinyuk, A., Dubovik, O., Smirnov, A., Giles, D., O'Neill, N. T., Tsay, S. C., Ji, Q., Al Mandoos, A., Khan, M. R., Reid, E. A., Schafer, J. S., Sorokine, M., Newcomb, W., and Slutsker, I.: Spatial and temporal variability of column-integrated aerosol optical properties in the southern Arabian Gulf and United Arab Emirates in summer, J. Geophys. Res.-Atmos., 113(D1), D01204, doi:10.1029/2007JD008944, 2008.

Edwards, R. and Sedwick, P.: Iron in East Antarctic snow: Implications for atmospheric iron deposition and algal production in Antarctic waters, Geophys. Res. Lett., 28(20), 3907–3910, 2001.

Formenti, P., Schuetz, L., Balkanski, Y., Desboeufs, K., Ebert, M., Kandler, K., Petzold, A., Scheuvens, D., Weinbruch, S., and Zhang, D.: Recent progress in understanding physical and chemical properties of mineral dust, Atmos. Chem. Phys. Discuss., 10, 31187–31251, doi:10.5194/acpd-10-31187-2010, 2010.

Forster, P., Ramaswmamy, V., Artaxo, P., Bernsten, T., Betts, R., Fahey, D. W., Haywood, J., Lean, J., Lowe, D. C., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schulz, M., and van Dorland, R.: Changes in Atmospheric Constituents and in Radiative Forcing, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averty, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, UK, and New York, NY, USA, 2007.

Gao, Y., Fan, S. M., and Sarmiento, J. L.: Aeolian iron input to the ocean through precipitation scavenging: A modeling perspective and its implication for natural iron fertilization in the ocean, J. Geophys. Res.-Atmos., 108(D7), 4221, doi:10.1029/2002JD002420, 2003.

Ghan, S. J. and Easter, R. C.: Impact of cloud-borne aerosol representation on aerosol direct and indirect effects, Atmos. Chem. Phys., 6, 4163–4174, doi:10.5194/acp-6-4163-2006, 2006.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., and Lin, S. J.: Sources and distributions of dust aerosols simulated with the GOCART model, J. Geophys. Res.-Atmos., 106(D17), 20255–20273, 2001.

Goossens, D. and Rajot, J. L.: Techniques to measure the dry aeolian deposition of dust in arid and semi-arid landscapes: a comparative study in West Niger, Earth Surface Processes and Landforms, 33(2), 178–195, 2008.

Grini, A., Myhre, G., Zender, C. S., and Isaksen, I. S. A.: Model simulations of dust sources and transport in the global atmosphere: Effects of soil erodibility and wind speed variability, J. Geophys. Res.-Atmos., 110(D2), D02205, doi:10.1029/2004JD005037, 2005.

Guelle, W., Balkanski, Y. J., Schulz, M., Marticorena, B., Bergametti, G., Moulin, C., Arimoto, R., and Perry, K. D.: Modeling the atmospheric distribution of mineral aerosol: Comparison with ground measurements and satellite observations for yearly and synoptic timescales over the North Atlantic, J. Geophys. Res.-Atmos., 105(D2), 1997–2012, 2000.

Han, Q., Moore, J. K., Zender, C., Measures, C., and Hydes, D.: Constraining oceanic dust deposition using surface ocean dissolved Al, Global Biogeochem. Cy., 22(2), GB2003, doi:10.1029/2007GB002975, 2008.

Hand, J. L., Mahowald, N. M., Chen, Y., Siefert, R. L., Luo, C., Subramaniam, A., and Fung, I.: Estimates of atmospheric-processed soluble iron from observations and a global mineral aerosol model: Biogeochemical implications, J. Geophys. Res.-Atmos., 109(D17), D17205, doi:10.1029/2004JD004574, 2004.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – A federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66(1), 1–16, 1998.

Holben, B. N., Tanre, D., Smirnov, A., Eck, T. F., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Castle, J. V., Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K., and Zibordi, G.: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET, J. Geophys. Res.-Atmos., 106(D11), 12067–12097, 2001.

Hollingsworth, A., Engelen, R. J., Textor, C., Benedetti, A., Boucher, O., Chevallier, F., Dethof, A., Elbern, H., Eskes, H., Flemming, J., Granier, C., Kaiser, J. W., Morcrette, J. J., Rayner, P., Peuch, V. H., Rouil, L., Schultz, M. G., and Simmons, A. J.: Toward a monitoring and forecasting system for atmospheric composition: The GEMS project, B. Am. Meteorol. Soc., 89(8), 1147–1164, doi:10.1175/2008BAMS2355.1, 2008.

Horowitz, L. W., Walters, S., Mauzerall, D. L., Emmons, L. K., Rasch, P. J., Granier, C., Tie, X. X., Lamarque, J. F., Schultz, M. G., Tyndall, G. S., Orlando, J. J., and Brasseur, G. P.: A global simulation of tropospheric ozone and related tracers: Description and evaluation of MOZART, version 2, J. Geophys. Res.-Atmos., 108(D24), 4784, doi:10.1029/2002JD002853, 2003.

Husar, R. B., Prospero, J. M., and Stowe, L. L.: Characterization of tropospheric aerosols over the oceans with the NOAA advanced very high resolution radiometer optical thickness operational product, J. Geophys. Res.-Atmos., 102(D14), 16889–16909, 1997.

Jickells, T. D., An, Z. S., Andersen, K. K., Baker, A. R., Bergametti, G., Brooks, N., Cao, J. J., Boyd, P. W., Duce, R. A., Hunter, K. A., Kawahata, H., Kubilay, N., laRoche, J., Liss, P. S., Mahowald, N., Prospero, J. M., Ridgwell, A. J., Tegen, I., and Torres, R.: Global iron connections between desert dust, ocean biogeochemistry, and climate, Science, 308(5718), 67–71, 2005.

Jimenez-Guerrero, P., Perez, C., Jorba, O., and Baldasano, J. M.: Contribution of Saharan dust in an integrated air quality system and its on-line assessment, Geophys. Res. Lett., 35(3), L03814,

doi:10.1029/2007GL031580, 2008.

Kim, K. W., Kim, Y. J., and Oh, S. J.: Visibility impairment during Yellow Sand periods in the urban atmosphere of Kwangju, Korea, Atmos. Environ., 35(30), 5157–5167, 2001.

Kinne, S., Lohmann, U., Feichter, J., Schulz, M., Timmreck, C., Ghan, S., Easter, R., Chin, M., Ginoux, P., Takemura, T., Tegen, I., Koch, D., Herzog, M., Penner, J., Pitari, G., Holben, B., Eck, T., Smirnov, A., Dubovik, O., Slutsker, I., Tanre, D., Torres, O., Mishchenko, M., Geogdzhayev, I., Chu, D. A., and Kaufman, Y.: Monthly averages of aerosol properties: A global comparison among models, satellite data, and AERONET ground data, J. Geophys. Res.-Atmos., 108(D20), 4634, doi:10.1029/2001JD001253, 2003.

Kinne, S., Schulz, M., Textor, C., Guibert, S., Balkanski, Y., Bauer, S. E., Berntsen, T., Berglen, T. F., Boucher, O., Chin, M., Collins, W., Dentener, F., Diehl, T., Easter, R., Feichter, J., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Herzog, M., Horowitz, L., Isaksen, I., Iversen, T., Kirkevåg, A., Kloster, S., Koch, D., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Lesins, G., Liu, X., Lohmann, U., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, O., Stier, P., Takemura, T., and Tie, X.: An AeroCom initial assessment optical properties in aerosol component modules of global models, Atmos. Chem. Phys., 6, 1815–1834, doi:10.5194/acp-6-1815-2006, 2006.

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.

Kohfeld, K. E. and Harrison, S. P.: DIRTMAP: the geological record of dust, Earth-Sci. Rev., 54(1–3), 81–114, 2001.

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.

Liao, H., Adams, P. J., Chung, S. H., Seinfeld, J. H., Mickley, L. J., and Jacob, D. J.: Interactions between tropospheric chemistry and aerosols in a unified general circulation model, J. Geophys. Res.-Atmos., 108(D1), 4001, doi:10.1029/2001JD001260, 2003.

Li-Jones, X. and Prospero, J. M.: Variations in the size distribution of non-sea-salt sulfate aerosol in the marine boundary layer at Barbados: Impact of African dust, J. Geophys. Res.-Atmos., 103(D13), 16073–16084, 1998.

Liu, X. and Penner, J. E.: Effect of Mount Pinatubo H_2SO_4/H_2O aerosol on ice nucleation in the upper troposphere using a global chemistry and transport model, J. Geophys. Res.-Atmos., 107(D12), 4141, doi:10.1029/2001JD000455, 2002.

Liu, X., Penner, J. E., Das, B. Y., Bergmann, D., Rodriguez, J. M., Strahan, S., Wang, M. H., and Feng, Y.: Uncertainties in global aerosol simulations: Assessment using three meteorological data sets, J. Geophys. Res.-Atmos., 112(D11), D11212, doi:10.1029/2006JD008216, 2007.

Mackie, D. S., Boyd, P. W., McTainsh, G. H., Tindale, N. W., Westberry, T. K., and Hunter, K. A.: Biogeochemistry of iron in Australian dust: From eolian uplift to marine uptake, Geochem. Geophy. Geosy., 9, Q03Q08, doi:10.1029/2007GC001813, 2008.

Maenhaut, W., Fernandez-Jimenez, M.-T., Rajta, I., Dubtsov, S., Meixner, F. X., Andreae, M. O., Torr, S., Hargrove, J. W., Chimanga, P., and Mlambo, J.: Long-term aerosol composition measruments and source apportionment at Rukomechi, Zimbabwe, J. Aerosol Sci., 31(Suppl. 1), S469–S470, 2000a.

Maenhaut, W., Fernandez-Jimenez, M.-T., Vanderzalm, J. L., Hooper, B. M., Hooper, M. A., and Tapper, N. J.: Aerosol composition at Jabiru, Australia and impact of biomass burning, J. Aerosol Sci., 31(Suppl. 1), S745–S746, 2000b.

Mahowald, N., Kohfeld, K., Hansson, M., Balkanski, Y., Harrison, S. P., Prentice, I. C., Schulz, M., and Rodhe, H.: Dust sources and deposition during the last glacial maximum and current climate: A comparison of model results with paleodata from ice cores and marine sediments, J. Geophys. Res.-Atmos., 104(D13), 15895–15916, 1999.

Mahowald, N. M., Baker, A. R., Bergametti, G., Brooks, N., Duce, R. A., Jickells, T. D., Kubilay, N., Prospero, J. M., and Tegen, I.: Atmospheric global dust cycle and iron inputs to the ocean, Global Biogeochem. Cy., 19(4), GB4025, doi:10.1029/2004GB002402, 2005.

Mahowald, N. M., Muhs, D. R., Levis, S., Rasch, P. J., Yoshioka, M., Zender, C. S., and Luo, C.: Change in atmospheric mineral aerosols in response to climate: Last glacial period, preindustrial, modern, and doubled carbon dioxide climates, J. Geophys. Res.-Atmos., 111(D10), D10202, doi:10.1029/2005JD006653, 2006.

Mahowald, N., Jickells, T. D., Baker, A. R., Artaxo, P., Benitez-Nelson, C. R., Bergametti, G., Bond, T. C., Chen, Y., Cohen, D. D., Herut, B., Kubilay, N., Losno, R., Luo, C., Maenhaut, W., McGee, K. A., Okin, G. S., Siefert, R. L., and Tsukuda, S.: Global distribution of atmospheric phosphorus sources, concentrations and deposition rates, and anthropogenic impacts, Global Biogeochem. Cy., 22(4), GB4026, doi:10.1029/2008GB003240, 2008.

Mahowald, N. M., Engelstaedter, S., Luo, C., Sealy, A., Artaxo, P., Benitez-Nelson, C., Bonnet, S., Chen, Y., Chuang, P. Y., Cohen, D. D., Dulac, F., Herut, B., Johansen, A. M., Kubilay, N., Losno, R., Maenhaut, W., Paytan, A., Prospero, J. A., Shank, L. M., and Siefert, R. L.: Atmospheric Iron Deposition: Global Distribution, Variability, and Human Perturbations, Annu. Rev. Mar. Sci., 1, 245–278, 2009.

Mahowald, N., Albani, S., Engelstaedter, S., Winckler, G., and Goman, M.: Model insight into glacial-interglacial paleodust records, Quaternary Science Reviews, 30(7–8), 832–854, 2011.

Mangold, A., De Backer, H., De Paepe, B., Dewitte, S., Chiapello, I., Derimian, Y., Kacenelenbogen, M., Léon, J.-F., Huneeus, N., Schulz, M., Ceburnis, D., O'Dowd, C. D., Flentje, H., Kinne, S., Benedetti, A., Morcrette, J. J., and Boucher, O.: Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part IV: Evaluation by means of case studies, J. Geophys. Res, 116, D03302, doi:10.1029/2010JD014864, 2011.

Miller, R. L., Cakmur, R. V., Perlwitz, J., Geogdzhayev, I. V., Ginoux, P., Koch, D., Kohfeld, K. E., Prigent, C., Ruedy, R.,

Schmidt, G. A., and Tegen, I.: Mineral dust aerosols in the NASA goddard institute for Space Sciences ModelE atmospheric general circulation model, J. Geophys. Res.-Atmos., 111(D6), D06208, doi:10.1029/2005JD005796, 2006.

Morcrette, J. J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J. W., Razinger, M., Schulz, M., Serrar, S., Simmons, A. J., Sofiev, M., Suttie, M., Tompkins, A. M., and Untch, A.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling, J. Geophys. Res.-Atmos., 114, D06206, doi:10.1029/2008JD011235, 2009.

Myhre, G., Bellouin, N., Berglen, T. F., Berntsen, T. K., Boucher, O., Grini, A., Isaksen, I. S. A., Johnsrud, M., Mishchenko, M. I., Stordal, F., and Tanre, D.: Comparison of the radiative properties and direct radiative effect of aerosols from a global aerosol model and remote sensing data over ocean, Tellus B, 59(1), 115–129, 2007.

Nyanganyura, D., Maenhaut, W., Mathuthua, M., Makarau, A., and Meixner, F. X.: The chemical composition of tropospheric aerosols and their contributing sources to a continental background site in northern Zimbabwe from 1994 to 2000, Atmos. Environ., 41(12), 2644–2659, 2007.

Ozer, P., Laghdaf, M., Lemine, S. O. M., and Gassani, J.: Estimation of air quality degradation due to Saharan dust at Nouakchott, Mauritania, from horizontal visibility data, Water Air Soil Pollut., 178(1–4), 79–87, 2007.

Prospero, J. M.: The Atmospheric Transport of Particles to the Ocean, in: Particle Flux in the Ocean, edited by: Ittekkot, V., Schäfer, P., Honjo, S., and Depetris, P. J., John Wiley & Sons Ltd., New York, 1996.

Prospero, J. M.: Long-term measurements of the transport of African mineral dust to the southeastern United States: Implications for regional air quality, J. Geophys. Res.-Atmos., 104(D13), 15917–15927, 1999.

Prospero, J. M. and Lamb, P. J.: African droughts and dust transport to the Caribbean: Climate change implications, Science, 302(5647), 1024–1027, 2003.

Prospero, J. M. and Nees, R. T.: Impact of the North African Drought and El-Nino on Mineral Dust in the Barbados Trade Winds, Nature, 320(6064), 735–738, 1986.

Prospero, J. M., Glaccum, R. A., and Nees, R. T.: Atmospheric Transport of Soil Dust from Africa to South-America, Nature, 289(5798), 570–572, 1981.

Prospero, J. M., Uematsu, M., and Savoie, D. L.: Mineral aerosol transport to the Pacific Ocean, edited by: Riley, J. P., 187–218, Academic Press, New York, 1989.

Prospero, J. M., Savoie, D. L., and Arimoto, R.: Long-term record of nss-sulfate and nitrate in aerosols on Midway Island, 1981–2000: Evidence of increased (now decreasing?) anthropogenic emissions from Asia, J. Geophys. Res.-Atmos., 108(D1), 4019, doi:10.1029/2001JD001524, 2003.

Prospero, J. M., Landing, W. M., and Schulz, M.: African dust deposition to Florida: temporal and spatial variability and comparison to models, J. Geophys. Res., 115, D13304, doi:10.1029/2009JD012773, 2010.

Reddy, M. S., Boucher, O., Balkanski, Y., and Schulz, M.: Aerosol optical depths and direct radiative perturbations by species and source type, Geophys. Res. Lett., 32(12), L12803, doi:10.1029/2004GL021743, 2005a.

Reddy, M. S., Boucher, O., Bellouin, N., Schulz, M., Balkanski, Y., Dufresne, J. L., and Pham, M.: Estimates of global multicomponent aerosol optical depth and direct radiative perturbation in the Laboratoire de Meteorologie Dynamique general circulation model, J. Geophys. Res.-Atmos., 110(D10), D10S16, doi:10.1029/2004JD004757, 2005b.

Reid, J. S., Jonsson, H. H., Maring, H. B., Smirnov, A., Savoie, D. L., Cliff, S. S., Reid, E. A., Livingston, J. M., Meier, M. M., Dubovik, O., and Tsay, S. C.: Comparison of size and morphological measurements of coarse mode dust particles from Africa, J. Geophys. Res.-Atmos., 108(D19), 8593, doi:10.1029/2002JD002485, 2003.

Sabbah, I. and Hasan, F. A.: Remote sensing of aerosols over the Solar Village, Saudi Arabia, Atmos. Res., 90(2–4), 170–179, 2008.

Schmidt, G. A., Ruedy, R., Hansen, J. E., Aleinov, I., Bell, N., Bauer, M., Bauer, S., Cairns, B., Canuto, V., Cheng, Y., Del Genio, A., Faluvegi, G., Friend, A. D., Hall, T. M., Hu, Y. Y., Kelley, M., Kiang, N. Y., Koch, D., Lacis, A. A., Lerner, J., Lo, K. K., Miller, R. L., Nazarenko, L., Oinas, V., Perlwitz, J., Rind, D., Romanou, A., Russell, G. L., Sato, M., Shindell, D. T., Stone, P. H., Sun, S., Tausnev, N., Thresher, D., and Yao, M. S.: Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data, J. Climate, 19(2), 153–192, 2006.

Schulz, M.: Constraining model estimates of the aerosol radiative forcing, Université Pierre et Marie Curie, Paris VI, Paris, 2007.

Schulz, M., Textor, C., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Dentener, F., Guibert, S., Isaksen, I. S. A., Iversen, T., Koch, D., Kirkevåg, A., Liu, X., Montanaro, V., Myhre, G., Penner, J. E., Pitari, G., Reddy, S., Seland, Ø., Stier, P., and Takemura, T.: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations, Atmos. Chem. Phys., 6, 5225–5246, doi:10.5194/acp-6-5225-2006, 2006.

Sokolik, I. N., Winker, D. M., Bergametti, G., Gillette, D. A., Carmichael, G., Kaufman, Y. J., Gomes, L., Schuetz, L., and Penner, J. E.: Introduction to special section: Outstanding problems in quantifying the radiative impacts of mineral dust, J. Geophys. Res.-Atmos., 106(D16), 18015–18027, 2001.

Stier, P., Feichter, J., Kinne, S., Kloster, S., Vignati, E., Wilson, J., Ganzeveld, L., Tegen, I., Werner, M., Balkanski, Y., Schulz, M., Boucher, O., Minikin, A., and Petzold, A.: The aerosol-climate model ECHAM5-HAM, Atmos. Chem. Phys., 5, 1125–1156, doi:10.5194/acp-5-1125-2005, 2005.

Su, L. and Toon, O. B.: Saharan and Asian dust: similarities and differences determined by CALIPSO, AERONET, and a coupled climate-aerosol microphysical model, Atmos. Chem. Phys., 11, 3263–3280, doi:10.5194/acp-11-3263-2011, 2011.

Swap, R., Garstang, M., Greco, S., Talbot, R., and Kallberg, P.: Saharan Dust in the Amazon Basin, Tellus B, 44(2), 133–149, 1992.

Tagliabue, A., Bopp, L., and Aumont, O.: Evaluating the importance of atmospheric and sedimentary iron sources to Southern Ocean biogeochemistry, Geophys. Res. Lett., 36, L13601, doi:10.1029/2009GL038914, 2009.

Takemura, T., Nozawa, T., Emori, S., Nakajima, T. Y., and Nakajima, T.: Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model, J. Geophys.

Res.-Atmos., 110(D2), D02202, doi:10.1029/2004JD005029, 2005.

Takemura, T., Egashira, M., Matsuzawa, K., Ichijo, H., O'ishi, R., and Abe-Ouchi, A.: A simulation of the global distribution and radiative forcing of soil dust aerosols at the Last Glacial Maximum, Atmos. Chem. Phys., 9, 3061–3073, doi:10.5194/acp-9-3061-2009, 2009.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res.-Atmos., 106(D7), 7183–7192, 2001.

Tegen, I.: Modeling the mineral dust aerosol cycle in the climate system, Quaternary Sci. Rev., 22(18–19), 1821–1834, 2003.

Tegen, I., Harrison, S. P., Kohfeld, K., Prentice, I. C., Coe, M., and Heimann, M.: Impact of vegetation and preferential source areas on global dust aerosol: Results from a model study, J. Geophys. Res.-Atmos., 107(D21), 4576, doi:10.1029/2001JD000963, 2002.

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.

Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O., Chin, M., Dentener, F., Diehl, T., Feichter, J., Fillmore, D., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Horowitz, L., Huang, P., Isaksen, I. S. A., Iversen, T., Kloster, S., Koch, D., Kirkevåg, A., Kristjansson, J. E., Krol, M., Lauer, A., Lamarque, J. F., Liu, X., Montanaro, V., Myhre, G., Penner, J. E., Pitari, G., Reddy, M. S., Seland, Ø., Stier, P., Takemura, T., and Tie, X.: The effect of harmonized emissions on aerosol properties in global models – an AeroCom experiment, Atmos. Chem. Phys., 7, 4489–4501, doi:10.5194/acp-7-4489-2007, 2007.

Thomson, M. C., Molesworth, A. M., Djingarey, M. H., Yameogo, K. R., Belanger, F., and Cuevas, L. E.: Potential of environmental models to predict meningitis epidemics in Africa, Trop. Med. Int. Health, 11(6), 781–788, 2006.

Tie, X. X., Madronich, S., Walters, S., Edwards, D. P., Ginoux, P., Mahowald, N., Zhang, R. Y., Lou, C., and Brasseur, G.: Assessment of the global impact of aerosols on tropospheric oxidants, J. Geophys. Res.-Atmos., 110(D3), D03204, doi:10.1029/2004JD005359, 2005.

Todd, M. C., Karam, D. B., Cavazos, C., Bouet, C., Heinold, B., Baldasano, J. M., Cautenet, G., Koren, I., Perez, C., Solmon, F., Tegen, I., Tulet, P., Washington, R., and Zakey, A.: Quantifying uncertainty in estimates of mineral dust flux: An intercomparison of model performance over the Bodele Depression, northern Chad, J. Geophys. Res.-Atmos., 113, D24107, doi:10.1029/2008JD010476, 2008.

Uno, I., Wang, Z., Chiba, M., Chun, Y. S., Gong, S. L., Hara, Y., Jung, E., Lee, S. S., Liu, M., Mikami, M., Music, S., Nickovic, S., Satake, S., Shao, Y., Song, Z., Sugimoto, N., Tanaka, T., and Westphal, D. L.: Dust model intercomparison (DMIP) study over Asia: Overview, J. Geophys. Res.-Atmos., 111(D12), D12213, doi:10.1029/2005JD006575, 2006.

Vanderzalm, J. L., Hooper, M. A., Ryan, B., Maenhaut, W., Martin, P., Rayment, P. R., and Hooper, B. M.: Impact of seasonal biomass burning on air quality in the "Top End" of regional Northern Australia, Clean Air and Environmental Quality, 37(3), 28–34, 2003.

Wagener, T., Guieu, C., Losno, R., Bonnet, S., and Mahowald, N.: Revisiting atmospheric dust export to the Southern Hemisphere ocean: Biogeochemical implications, Global Biogeochem. Cy., 22(2), GB2006, doi:10.1029/2007GB002984, 2008.

Woodage, M. J., Slingo, A., Woodward, S., and Comer, R. E.: UK HiGEM: Simulations of Desert Dust and Biomass Burning Aerosols with a High-Resolution Atmospheric GCM, J. Climate, 23(7), 1636–1659, 2010.

Zender, C. S., Bian, H. S., and Newman, D.: Mineral Dust Entrainment and Deposition (DEAD) model: Description and 1990s dust climatology, J. Geophys. Res.-Atmos., 108(D14), 4416, doi:10.1029/2002JD002775, 2003.

Zender, C. S., Miller, R. L., and Tegen, I.: Quantifying mineral dust mass budgets: Systematic terminology, constraints, and current estimates, in EOS, 509–512, 2004.