Global dust model intercomparison in AeroCom phase I

N. Huneeus¹, M. Schulz¹,², Y. Balkanski¹, J. Griesfeller¹,², S. Kinne³, J. Prospero⁴, S. Bauer⁵,⁶, O. Boucher⁷, M. Chin⁸, F. Dentener⁹, T. Diehl¹⁰, R. Easter¹¹, D. Fillmore¹², S. Ghan¹¹, P. Ginoux¹³, A. Grini¹⁴,¹⁵, L. Horowitz¹³, D. Koch⁵,⁶, M. C. Krol¹⁶,¹⁷, W. Landing¹⁸, X. Liu¹⁹,¹¹, N. Mahowald¹⁰, R. Miller⁵,⁶, J.-J. Morcrette²¹, G. Myhre¹⁴,²², J. E. Penner¹⁹, J. Perlwitz⁵,⁶, P. Stier²³, T. Takemura²⁴, and C. Zender²⁵

¹Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France
²Meteorological Institut, Oslo, Norway
³Max-Planck-Institut fur Meteorologie, Hamburg, Germany
⁴Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, USA
⁵Columbia University, New York, USA
⁶NASA GISS, New York, NY, USA
⁷Met Office, Hadley Centre, Exeter, UK
Global dust model intercomparison in AeroCom phase I

N. Huneeus et al.

Received: 5 August 2010 – Accepted: 24 September 2010
– Published: 12 October 2010

Correspondence to: N. Huneeus (nicolas.huneeus@lsce.ipsl.fr)
Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Desert dust plays an important role in the climate system through its impact on Earth’s radiative budget and its role in the biogeochemical cycle as a source of iron in high-nutrient-low-chlorophyll regions. A large degree of diversity exists between the many global models that simulate the dust cycle to estimate its impact on climate. We present the results of a broad intercomparison of a total of 15 global aerosol models within the AeroCom project. Each model is compared to observations focusing on variables responsible for the uncertainties in estimating the direct radiative effect and the dust 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 concentration are included to extend the assessment of model performance. These datasets form a benchmark data set which is proposed for model inspection and future dust model developments. In general, models perform better in simulating climatology of vertically averaged integrated parameters (AOD and AE) in dusty sites than they do with total deposition and surface concentration. Almost all models overestimate deposition fluxes over Europe, the Indian Ocean, the Atlantic Ocean and ice core data. Differences among the models arise when simulating deposition at remote sites with low fluxes over the Pacific and the Southern Atlantic Ocean. This study also highlights important differences in models ability to reproduce the deposition flux over Antarctica. The cause of this discrepancy could not be identified but different dust regimes at each site and issues with data quality should be considered. Models generally simulate better surface concentration at stations downwind of the main sources than at remote ones. Likewise, they simulate better surface concentration at stations affected by Saharan dust than at stations affected by Asian dust. 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 also reproduce the dust transport across the Atlantic in terms of both AOD and AE; they simulate the offshore transport of West Africa throughout
the year and limit the transport across the Atlantic to the summer months, yet overesti-
imating the AOD and transporting too fine particles. However, most of the models do
not reproduce the southward displacement of the dust cloud during the winter respon-
sible of the transport of dust into South America. Based on the dependency of AOD on
aerosol burden and size distribution we use model data bias with respect to AOD and
AE and infer on the over/under estimation of the dust emissions. According to this we
suggest the emissions in the Sahara be between 792 and 2271 Tg/yr and the one in
the Middle East between 212 and 329 Tg/yr.

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
large impact on Earth’s radiative budget due to the absorption and scattering 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 in 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 have
been suggested to human health in the dispersion of vector-borne diseases such as
meningitis (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 to ocean CO₂ uptake in HNLC regions difficult (Textor et al., 2006; Tagliabue et al., 2009).

An exhaustive comparison of different models between 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. 2000 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 that they 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 aerosol models. They studied the differences among state of the art models to simulate 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 when estimating the radiative impact of dust. However, no comparisons against observations
were made in this 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. Previous works comparing two or more global dust models are Tegen (2003), Zender et al. (2004) and Prospero et al. (2010). In the first the modelling of the dust cycle in the climate system is analyzed by comparing two global models to satellite climatology of TOMS aerosol index (AI). In Zender et al. (2004) the emission fluxes and burdens for different models are compared. Finally, Prospero et al. (2010) conducted a more exhaustive intercomparison. In this study the temporal and spatial variability of dust deposition in Florida was analyzed and compared to models within the AeroCom initiative. The comparison revealed among other results that models reproduce the variability but underestimate the intensity of the deposition at sites remote from African source regions.

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 may be a powerful constraint to the overall global dust budget. Total deposition fluxes are most useful when accumulated for a longer time period. 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). Most of the sites are located in the Northern Hemisphere and far away from source regions. The measured values range from 450 [g/m$^2$/yr] in the Taklimakan desert to 0.09 [g/m$^2$/yr] in the equatorial pacific and 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 from sediment traps being part of 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 of 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 since they represent the integrated dust flux to the ocean over a time span of hundreds to possible 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 data of total deposition finally used in this study consist of 84 sites with yearly dust deposition fluxes not coincident 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 ease 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 climatology of dust deposition observation. However, some of these measurements do not cover a long enough period to be climatological in a strict sense. The impact of this assumption on the model evaluation will be considered in the discussion (Sect. 4). Deposition data at the same location were averaged in order to provide one climatological data.

Dust particles have been suggested to be very efficiently removed by wet scavenging, especially over open ocean (Prospero et al., 2010; Hand et al., 2004; Gao et al., 2003). To test the models in terms of wet deposition we compare them against data from the Florida Atmospheric Mercury Study (FAMS) network (Prospero et al., 2010). 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 expand the comparison to the remaining AeroCom models.

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 measurements 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 correspond to measurements in remote regions and that most of the models only simulate dust particles up to 10 µm, we do not consider the cut-off size as a 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 against monthly dust concentrations measurements taken at 19 sites managed by the Rosenstiel School of Marine and Atmospheric Science from the University of Miami (Prospero et al., 1989; Prospero, 1996; Arimoto et al., 1995). The measurements taken in the Pacific Ocean stem 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 among other goals to study the large-scale spatial and temporal variability of aerosols. The measuring sites are located in general remote from dust emission sources and downwind of them (Fig. 2). A list of the stations as well as their location is given in Table S2 in the supplement material. 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 ºC after extracting soluble components with water. This database has been largely used for the evaluation of dust models (Ginoux et al., 2001; Cheng et al., 2008; Tegen et al., 2002). The measurements were
taken in the 1980s and 1990s with different 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 since 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/m³ throughout the year. These stations are located in the Antarctica and in the Ocean Pacific 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. Finally, the third group corresponds to locations downwind of the Sahara and Asian dust source, presented by blue numbers and dots in Fig. 2. In each one of these groups 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.

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 cruise but include also long term measuring stations. The measurements taken during cruise campaigns will be considered as yearly averages even though they represent short-term data. Most of the annually averaged dust arrives on a few days (5% of days) (Mahowald et al., 2009). 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 we 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.
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 concentration captured efficiently particles smaller than 10 µm and therefore the cut-off characteristics of the instrument are not considered to bias the results (Arimoto et al., 1995).

2.3 Aerosol optical depth and Angström Exponent

AOD informs about column integrated aerosol loads, if the mass extinction coefficient is known. The widespread deployment of sun photometers has provided recently global and very reliable information also about dust, unfortunately only 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 miss also 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 clean atmosphere in remote regions and polluted areas (Holben et al., 1998, 2001). We use here total
AOD and coarse mode AOD at 550 nm and Ångström Exponent (AE) from AERONET. The coarse mode AOD relies only on almucantar and azimuth plane measurements whereas the total AOD results from the before mentioned measurements plus direct sun radiances. The AERONET data used in this study combine the measurements from the Sun and Skynet instruments. The AE is calculated from multi-wavelength direct sun observations and delivers useful information on the aerosol size distribution. The model AE when not provided is computed from the AOD at 550 and 865 nm.

Although AERONET provides daily averaged data of the above mentioned parameters we shall focus on the monthly mean solely. This provides a comprehensive picture of the seasonal dust cycle and excludes the evaluation of the frequency and intensity of dust events. The study of the performance 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. The original data are also aggregated into monthly basis. We use all available stations with measurements for the year 2000 and a climatology constructed considering 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 mixture of these aerosol modes. The authors used the accumulation-mode fraction to discern between these three cases whereas we will use the AE. We consider 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. Considering that the loading of oceanic aerosols does not exceed AOD (at 440 nm) of 0.15 (Dubovik et al., 2002) we filter out the oceanic
aerosol stations from stations dominated by natural aerosol by eliminating those stations with monthly AOD (at 550 nm) smaller than 0.2. It should be noted that remote stations can be dominated by fine mode desert dust and thus have AE larger than 1.2. However, since we cannot separate these stations from those dominated by other fine mode aerosols (sulphate, black carbon, organic matter) 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 (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 the result of a single day of low AE not necessarily linked to desert dust. For comparisons purpose however we consider all month at dusty sites with AE smaller than 1.2. In addition, in view of the model’s coarse resolution (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, Middle East and American 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 Tables S3 and S4 in the supplement.

2.4 AeroCom dust 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 in 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 experiments conducted considered a total of sixteen global models. Each model simulated the year 2000 and defined the conditions of simulation independently. This experiment “A” is documented in Textor et al. (2006) and Kinne et al. (2006). Another experiment “B” was conducted where the emissions were the same for all models (Textor et al., 2007) and where radiative forcing was assessed (Schulz et al., 2006). In this study we will use the model outputs for the year 2000 of the experiment A. For model TM5 where experiment A was not available, we used results from experiment B instead.

The model features which 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) since their configuration was not meant to simulate the dust cycle. Furthermore, some models entering the AeroCom project after the initial publication of experiments A and B were included (CAM for CAnadian Model). 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 varies from variable to variable. Table 2 provides the list with the models used to compute each variable.

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). Aerosol optical depth products from the Moderate resolution Imaging Spectroradiometer (MODIS) are assimilated to better estimate the aerosol fields (Benedetti et al., 2009). We will consider simulations without data assimilation for this study. For evaluation of the impact of data assimilation on the model performance see Benedetti et al. (2009) and Mangold et al. (2010).

We computed the global model dust budgets for each one of the models (Table 3). The range of emissions exceeds the range of 1000–3000 Tg/yr usually attributed to global models (e.g. Zender et al., 2004). The annual emissions of the AeroCom models in the phase I are between 500 and 4400 Tg. The global averaged dust AOD ranges from 0.01 to 0.053 with most 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 of the models.

We highlight 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 performance of each model to reproduce different aspects of the desert dust cycle is evaluated by comparing them against the datasets described above. We conduct the analysis on a station by station basis. We make use of AeroCom tools developed at the Laboratoire du Climat et de l’Environnement (LSCE) to conduct the comparison and evaluation. Additional analysis for each model 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 supplement for information on the stations illustrated. All the models in this study mostly overestimate largely (a factor 10 and up to 2 orders
of magnitude) the deposition fluxes. The level of overestimation varies largely from model to model. All models agree in overestimating the deposition in Europe (light green), the Indian Ocean (dark green), in the Atlantic (orange and black) and the ice core data (pink). The sole underestimation in Europe occurs in Lake Kinneret (49) in Israel. Almost all models (except 2 models) underestimate the deposition at this site. Differences among the models arise in simulating the deposition in remote sites with low flux in the Pacific and the Southern Atlantic; the degree of over/underestimation varies largely from model to model. Most of the models mainly overestimate the deposition data from Mahowald et al. (2009) in the Antarctica whereas most of the models underestimate the deposition in the Weddel Sea (13) in Antarctica (DIRTMAP; Tegen et al., 2002). This difference in performance will be discussed in Sect. 4. The sole data measuring deposition fluxes in the Taklimakan desert in inner Asia comes from station H (purple in Fig. 1). This data is over/underestimated without apparent coherence to the performance in reproducing deposition in other locations close to major sources.

We expand the analysis conducted on 9 AeroCom model in Prospero et al. (2010) to reproduce 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 not only to validate the model performance to reproduce wet and total deposition but also to evaluate models to capture the dust transport across the Atlantic. As in that study we choose three representative stations from the nine stations in the FAMS network to illustrate the model performance. These stations are oriented from south (Little Crawl Key) to north (Lake Barco) and present therefore a latitudinal gradient of deposition in Florida. The results presented in that study are also valid when additional AeroCom models are included. A large amount of models capture the seasonality of the deposition and the dominance of the 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, mostly underestimating it (Fig. 3). The model performance deteriorates from south to north reflecting model difficulties to transport the dust northward.
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 compare the models in term of reproducing the seasonality of the observations (Fig. 6).

Models mainly overestimate the surface concentration from cruises (squares and filled-in circles in Fig. 4) and long term measuring stations (diamonds in Fig. 4). This overestimation exceeds a factor of ten and in most of the cases it even exceeds the two order of magnitudes observed in the deposition. The underestimation is mainly within a factor of 10 with respect of the observation. The cases where the underestimation exceeds this limit correspond to measurements taken in remote regions of the southern Atlantic Ocean. The cruise measurements correspond to short-term measurements. Considering that the dust events occur in a reduced number of days per year (∼5%, Mahowald et al., 2009) the possibility exists that the measurements miss one (or more) of the events or that they actually coincided with an event. The error in the measurements associated to missing a dust event or coinciding with one is represented by the vertical lines in Fig. 4. For each model these errors correspond respectively to 96% and 20% of the model yearly average. When this error is taken into account the model performance depends strongly on the magnitude in the error since the large overestimation can become an underestimation. When comparing the models to yearly averages of the SEAREX and AEROCCE data however, the over and underestimation is mostly within a factor 10 (Fig. 5). The cases where the underestimations exceed this limit correspond mostly to measurements of stations located in the Antarctica (stations 1, 8 and 9). Furthermore, the correspondence between modelled surface concentration and measured one in most of the models improves in stations with higher values; the agreement is much better in stations downwind of mayor dust sources (stations 16 to 21) than in the other two groups (Fig. 5). Likewise, the correspondence is better for stations of the second group (stations 8 to 15) than of the first one (stations 1 to 7). Half of the models present larger differences with the observed surface concentrations at
sites associated to the Asian sources (stations 19 and 21) than at stations measuring the transatlantic dust transport from the Sahara (stations 17 and 18). The remaining models present similar performance to reproduce surface concentrations associated to both deserts. The above reflects perhaps a certain difficulty to simulate simultaneously the magnitude of the dust emissions from Sahara and Asia (Tegen et al., 2002). All models underestimate the surface concentration at Rukomechi (16). This station measures the dust emitted from the Kalahari Desert.

We present the seasonal variability of surface concentration at the different sites as Hovmoller-like diagrams (Fig. 6). Each row corresponds to the monthly surface concentration of a particular station of the network illustrated in Fig. 2. For each station the number attributed in Fig. 2 corresponds to the row in Fig. 6. We continue to group the stations as done above and identify each group of station with a coloured bar on the left side of the left hand figures. The choice of colours is the same as the one applied in Figs. 2 and 5. In Fig. 6 the groups have been labelled Low, Medium and High according to their regime of surface concentration. We note that the stations are not grouped in a geographically meaningful way as usually done in Hovmoller diagram, which is designed to indicate spatial propagation of features with time.

The models present large diversity in simulating the seasonal surface concentration (Fig. 6). A large amount of models (57%) mainly underestimate the surface concentration in stations of the first group; not enough dust is transported to remote sites. Both, over and underestimation exceed mostly 80% of the observed value. Almost all modes (12 out of 14) overestimate the surface concentration in Caledonia (2) at least during the two months of August and September. Models show at this station similar magnitude as in the station of Norfolk (12). The difference in observed surface concentration between these two stations suggest that Caledonia (2) is outside of the southeast dust pathway documented in Mackie et al. (2008). Therefore, the overestimation in Caledonia may reveal difficulties to correctly simulate the south-eastern dust transport outside of Australia. However, since the data are a climatology and considering the episodic nature of the dust emissions in Australia (Mackie et al., 2008) and the closeness of...
these stations it might also be that this overestimation is actually the result of a single event not represented in long term measurements.

Most of the models fail to reproduce, to different degrees, the surface concentration in stations belonging to the second group (Sect. 2.2). A large number of models (86%) mainly underestimate the surface concentration throughout the year at the stations Palmer (8) and King George (9) (Fig. 6). An important number of models have difficulties in reproducing the measurements in Cape Point (11) in terms of both, magnitude and extension. The closeness of the stations to the source suggests an overestimation of the source intensity, although problems in simulating vertical and horizontal transport as well as dust size distribution cannot be excluded. The Australian emissions are captured by the stations Cape Grim (10) and Norfolk Island (12). Most of the models fail to reproduce the seasonality and intensity of the measurements at each one of these stations, however the performance in simulating surface concentration in Norfolk Island (12) is better than in Cape Grim (10). Surface concentration in Hawaii (13) is largely overestimated by a few models suggesting an overestimation of the transported dust out of Asia.

The transatlantic transport of Saharan dust with captured by measurements in Barbados and Miami (stations 17 and 18 respectively), is reproduced by most of the models. The seasonal cycle with maximum surface concentration between the month of May and September is overestimated by five of the models at these stations in periods of maximum concentration. Outside of the summer months the model’s performance varies from station to station. The surface concentration in Bermuda (20) that illustrates the northward transport of the transatlantic dust cloud, is underestimated by nine models (64%) in periods of maximum concentration. The stations of Hedo (19) and Cheju (21) illustrate that the dust transport out of Asia is a year around phenomena with the largest concentrations from February till May and the lowest ones from June to August coherent with the seasonal cycle described in Prospero (1996). Most of the models simulate these seasonal features but fail to reproduce the intensity.
We also compare the models to measurement in Barbados (17) and Miami (18) for the year 2000 (Fig. 7). These two stations illustrate not only the dust transport across the Atlantic but also its northern latitudinal extend. Large model diversity is observed at the station of Barbados. While most of the models coincide on the onset of the period of maximum surface concentration, they differ largely in simulating the magnitude of the measurements during this period and its extension. Furthermore, models also differ in simulating the surface concentration prior and after this period, most of them underestimating the surface concentration from October to May. The model diversity in reproducing the measurements in Miami is smaller. However models exist that reproduce the seasonal cycle at Barbados but fail to do so in Miami. This suggests problems to simulate the process responsible for the northward displacement of the dust transport. The climatology in Fig. 7 (black dashed line) illustrate that it is close to the seasonal cycle of the year 2000. However some differences can be observed, namely the peak in surface concentration in Miami in the year 2000 lags the climatology by one month and the seasonal cycle in Barbados shows more summer variability than the climatology. The climatology presents decreasing values from the maximum in June whereas the measurements from the year 2000 present two maxima, one in June coincident with the climatology and one in August.

### 3.3 Total aerosol optical depth

We compare now 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 month (as defined in Sect. 2.3) and represents therefore 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, Middle East and American stations and stations elsewhere (Figs. 8 and 9 respectively). 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 are used (Sect. 2.3, Fig. 8). Names and location for each one of these sites are given in Table S3 of the supplement. The models present a large diversity in reproducing the averaged AOD at 550 nm on dusty stations but in general the modeled AOD is within a twofold range of the observations at most of the sites (Fig. 10). The data show in general higher AOD in African stations than the ones in the Middle East and these 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. The gradient in AOD between African and Middle East dusty stations is not reproduced by all models which attribute to the Middle East AOD’s the same order of magnitude than in Africa and/or overestimate the AOD in 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 in the Middle East and/or Africa, although the wrong size distribution and thus the extinction could also be responsible. This aspect will be further developed in Sect. 4. Finally, twelve of the models underestimate the AOD in Kanpur station (25), again this suggests that the emissions are underestimated. Nine of the 15 models present a negative bias, the rms error of all models is between 0.1 and 0.2 and except for 4 models the correlation is above 0.65.

The AOD climatology presents a seasonality characterized by high AOD in Africa with maximum values from December to April in the most southern stations shifting progressively to July till September in the most northern African stations (Fig. 11). Most of the models capture these features both over and underestimating the AOD but a few models exist (5) that show dusty season during the same period from south to north mostly overestimating the AOD. The AOD in the Middle East presents a yearly cycle with maximum from May/June to September. Again most of the models reproduce this seasonality while a few models present a too extended dusty season or in the wrong period of the year. With respect to the American stations, no model simulates the AOD in Paddockwood (station 23), most of the models capture the higher AOD in
the summer month of June to September in stations affected by the transatlantic dust transport in this period (stations 18 to 21). At Capo Verde (station 24), even though the higher AOD from June to September is mostly simulated, a large amount of models (10) mainly overestimate the AOD throughout the year. In Kanpur (station 25) on the contrary, the seasonality is reproduced but the magnitude is mainly underestimated by 12 of the 15 models.

In the analysis of data for the year 2000 fewer stations are included since the number of available stations for this particular year is smaller. Only 8 AERONET stations from a total of 446 met the requirements of dusty station described in Sect. 2.3 (Table S4 in the supplement, Fig. 9).

The averaged AOD is again reasonably well simulated by all models; the simulated AOD is within a factor two of the observed one at almost all stations and almost all models (Fig. 12). An AOD gradient similar to the climatology is observed between the dust regions; Africa has the largest AOD followed by Middle East and then stations in America. An important number of models (7) reproduce the AOD gradient between the defined regions. Eight of the models underestimate the AOD at American sites while six of the 15 models overestimate the AOD in all African stations. A few models (4) exist that systematically underestimate the AOD at all or almost all of the dusty stations. The rms error spans now from 0.06 to 0.2 and nine of the 15 models present a negative bias in reproducing the averaged AOD for the year 2000. Except for one model, all models present larger correlation ($R$) when simulating the average of the year 2000 than the one produced with climatology.

No systematic over or underestimation for any of the defined regions is observed in any of the models (Fig. 13). In general a large amount of models reproduce the shifting of maximum AOD in African stations from March in Ouagadougou (1) to June in Dakar (3), yet no model reproduces the second maximum in October in Ouagadougou (station 1) and Banizoumbou (2). They either fail to reproduce the second maximum at all, simulate it delayed by one month or too long in duration. All of the models simulate year-round dust transport off Africa at Capo Verde (8). While a large amount of
models simulate the two maxima present in the observations a few models exist (4) that present only a single maxima. This last may indicate deficiencies in reproducing the mechanism responsible of the dust transport offshore. At the American stations all models reproduce the transatlantic dust transport 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 reproduces the observed seasonal cycle of maximum AOD in winter month in Surinam (5) product of the southward displacement of African dust cloud during the winter (Prospero and Lamb, 2003). No relation is observed between the performance to reproduce the AOD over Africa and America with the performance to reproduce the yearly cycle of AOD in the Middle East. Nine of the models manage to reproduce a yearly cycle with one maximum in the summer months.

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 multiple-angle sky observations (almucantar and azimuth plane measurements) and is therefore less available than the total AOD which is also retrieved with direct sun measurements. As a consequence of this difference in number of available measurements the coarse mode AOD can show larger values than the total AOD.

The coarse mode AOD climatology presents a seasonal cycle similar to the total AOD with maxima in coarse AOD coinciding with the ones in total AOD (Figs. 14 and 11 respectively). We highlight that stations Bandoukoui (3), Bidi Bahn (7) and Cape San Juan (21) do not have coarse mode AOD measurements for the selected period. The coarse mode AOD in this period represents more than half of the total AOD illustrating the dominance of coarse dust particles. The models in general successfully reproduce the dominance of coarse dust particles in periods of maximum AOD and present seasonality similar to the total AOD. Yet models exist that underestimate the coarse mode AOD in periods of maximum AOD suggesting a larger influence of fine particle.
The observed coarse mode AOD for the year 2000 shows the same features as the climatology; similar seasonality than total AOD and dominance of coarse mode dust particles in month of maximum total AOD (Fig. 15). Most of the models reproduce this seasonality and simulate a dominance of coarse mode particles in periods of maximum AOD. Yet again models exist that simulate a larger influence of fine mode particles in periods of maximum AOD.

### 3.5 Angström Exponent

We analyse now the climatology of the AE for dusty sites. Again, we start by analyzing the averaged AE (Fig. 16) and then the seasonal cycle at the 25 stations selected with climatological data (Fig. 17). Next we reproduce this analysis with the 8 stations selected using the data of the year 2000 (Figs. 18 and 19).

In general the over/under estimation of the models does not exceed twofold the observations (Fig. 16). Nine of the 13 models delivering AE underestimate it in the Middle East simulating larger particles than observed while nearly all models overestimate the AE in a good number of American stations. Larger diversity is observed when simulating the AE in African stations. Except for stations Ilorin (1) and Djougou (2) the averaged observed AE shows larger values in the Middle East than in Africa thus indicating smaller dust particles in the former. The AE in the American stations spans the range of values observed in Africa and the Middle East contrary to what could be expected for dust particles that have crossed the Atlantic Ocean and that should be dominated by small particles. We recall the reader that only month dominated by coarse dust aerosols or the mixture of coarse and fine aerosols are analyzed. Only half of the models reproduce this difference in AE between the Middle East and Africa while ten of the models simulate the wide range of AE in American stations. The rms error spans between 0.17 and 0.77 with 10 models having their rms between 0.2 and 0.36. The bias is positive for 9 of the 13 models indicating that these models simulate smaller than observed particles in months either presenting a mixture of fine and coarse particles or dominated by dust.
A few models exist (3) that fail to reproduce the AE variability at all stations and present rather homogenous yearly cycle (Fig. 17). In Africa the AE presents periods with coarse aerosol (AE < 0.4) in spring in most southern stations extending to almost throughout the year towards the north, going beyond the period of maximum AOD. This feature is captured by a large number of models (8 out of 13 models) in most of the cases reducing the duration of period with large AE. In the Middle East only some models (6) manage to reproduce large particles observed preceding the period of high AOD and the mixture of fine and coarse particles during the month of high AOD. Most of the models simulate the yearly cycle in the American stations 18 to 21 consistent with a dust contribution of large aerosols in the summer month. Yet models have more difficulties in reproducing the presence of large particles (smaller AE) from February to May in Surinam (17). The station of 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 with the observations mainly in the onset and duration of periods characterized by large particles. Most of the models have difficulties to simulate the yearly AE cycle with large particles present from May to July in the station of Kanpur in Northern India (25).

The averaged AE for the year 2000 shows that the smallest particles (largest AE) are observed in Solar Village (station 4) and Surinam (5) while African stations present values smaller than in the Middle East but larger than in the remaining American stations (Fig. 18). At Capo Verde (8) the observed averaged AE is comparable to values observed in Barbados (6) and Roosvelt Roads (7). The lower AE in American stations than in African ones, suggesting larger particles across the Atlantic than in the source region, is explained by the fewer data used for computing the average in the former. Eleven of the 13 models reproduce the observed AE for the year 2000 with differences not exceeding observations by a factor of two. While most models (10 out of 13) underestimate the AE in the Middle East they present larger diversity when simulating the AE in Africa and America. However, a large amount of models (9) reproduce the AE in Barbados (station 6) and Roosvelt Roads (station 7) better than in African Stations. The
same four models presenting a negative bias with the climatology present a negative bias with respect to the average of the year 2000. The range of rms error between the models is now 0.1 and 0.74 with eleven models having a rms error between 0.13 and 0.35. Except for two models, all models present a larger correlation when comparing the year 2000 than the climatology.

The seasonal cycle of AE at dusty stations of the year 2000 presents similar features as seen in the climatology. In Africa coarse particles (AE < 0.4) dominate periods with maximum AOD which most models (7 models out of 13) reproduce mostly overestimating the AE (Fig. 19). In Capo Verde (station 8) coarse particles dominate throughout the year, which again is captured by 7 models. A large amount of models (7) also 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 southward displacement of the dust transport in winter months described in Ginoux et al. (2001). Finally, in the Middle East, contrary to the African stations, only a few models (4) manage to simulate the seasonal cycle of AE, the median model being one of them. A large amount of models (6) suggest larger particles (i.e. underestimate) throughout the year.

4 Discussion

4.1 Surface variables

Most of the models (13 out of 15) agree in mainly overestimating the deposition fluxes throughout the world by a factor of ten. This overestimation, in particular in the southern ocean, might imply an overestimation of the role of dust iron in governing CO₂ variability when coupled to an ocean biogeochemistry model as suggested already by Tagliabue et al. (2009). Varying model performances are found however when it comes to reproducing deposition fluxes at certain stations in the Pacific and the South Atlantic Ocean. Different dust regimes influence each of these sites as indicated by the magnitude of the measured deposition. Difficulties to simulate 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 Antartica dust deposition fluxes used in Mahowald et al. (2009) stem from dissolved iron measurement in snow samples known to be too low (Edwards and Sedwick, 2001). The overestimation in the Northern Hemisphere may suggest a problem in representing the intensity of emissions, size distribution of the transported dust, transport itself and/or representation of deposition flux. The available data do not allow yet attributing the differences with the observations to any process in particular. When comparing the models against long-term measurements of total and wet deposition taken in Florida, models capture the seasonality of the deposition and the dominance of wet deposition but underestimate mainly the magnitude. Furthermore the performance deteriorates from south to north reflecting difficulties to reproduce the dust transport northward. 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). In the meantime of having this long term measurements, alternative techniques to evaluate deposition may be necessary. One such method is the inversion of measured dissolved aluminum in the surface ocean, and use of assumed Al concentration and solubility to infer an annual mean dust flux (Han et al., 2008).

The model performance to simulate surface concentration depends on the data sets used. When using measurements from cruises all models agree in mainly overestimating the surface concentration and this mostly from a factor of ten up to
two orders of magnitude. When using long-term measurements 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 error of the most likely case of missing dust events during the measure-
ments is taken into account the large overestimation is reduced (up to 96%) and the
performance resembles the one observed with long-term measurements. Yet for both
datasets and all models, when the surface concentration is underestimated it is within
a factor of ten with respect to the observations. The cases where this limit is exceeded
corresponds mostly to measurements in remote regions of the Southern Hemisphere.

We recall that for both surface concentration and deposition the period the data
were taken is not coincident with the simulated year, which could explain part of the
model-observation differences. However, the large over/under estimation by most of
the models points to other issues. Considering the episodic nature of dust events and
the few days in which they occur (Prospero et al., 2010; Mahowald et al., 2009), the
use of measurement of short duration for model evaluation is delicate considering that
important events may not be represented (more likely if measurements were taken
without aiming at dust episodes) or weighing in too much in the measurements and
the model might appear with a positive/negative bias. Particle size is also an important
factor and source of misfit when comparing deposition and/or surface concentration
to model outputs. If properly known on the instrument and model side, it would allow
correcting for instrument size cut-off and would give more insight into model perfor-
mance. We therefore suggest that size-resolved surface concentration and deposition
be archived in future model experiments in contrast to size-integrated variables, as they
were available for this study.

4.2 Vertically integrated variables

The models reproduce the AOD and AE within a factor of two with respect to the ob-
servations. In general, models manage to reproduce the seasonal cycle of AOD in
African dust stations and the dominance of coarse particles in periods of maximum
AOD. In the Middle East however models show more difficulties in reproducing both
AOD and AE; while nine models reproduce the AOD seasonal cycle only four models manage to simulate the yearly cycle of AE. The models reproduce the transport off Western Africa throughout the year as captured by the station of Capo Verde and limit the transatlantic dust transport to summer month as illustrated by the station of Barbados. While all models reproduce the AOD seasonality in this station most of the models transport too fine particles across the Atlantic Ocean (i.e. overestimate the AE). The latitudinal displacement of the dust cloud is linked to the movements of the Intertropical Convergence Zone (ITCZ). During the summer the ITCZ reaches its most northern position allowing the dust to be transported to the Caribbean whereas during the winter the ITCZ reaches its southernmost position pushing the dust to reach South America (Ginoux et al., 2001). The seasonal cycle of AOD in Surinam with minimum AOD in the summer and the maximum in AOD observed in Barbados during the summer illustrates this seasonal shift of the dust cloud. While most of the models successfully simulate the AOD seasonal cycle in Barbados they do not reproduce the minimum AOD confined to the summer month in Surinam. This might indicate problems in simulating the processes in general circulation in the tropics and/or wet deposition responsible for this southward shift of the transatlantic dust cloud.

### 4.3 Emissions

No dataset specifically measuring the emission of dust particles could be used in this study. Still, the combination of model performance to simulate AOD and AE allows us to infer on their performance to simulate 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 smaller particles will have larger AOD. Based on the latter, the combination of AE and AOD measurements can be used to infer whether the emissions are over or underestimated. To illustrate this lets 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 are necessary. This can be achieved by either emitting more fine mode parti-
cles, 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 emissions, especially of the coarse dust particles, if interference from other aerosol components can be excluded. In the same way the opposite case, 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 adjustments might also be needed in addition. Simultaneous over- or underestimation of AOD and of AE does not allow inferring whether the intensity of the source has been over or underestimated. Nevertheless an improvement in the simulation of the dust size distribution is recommended before exploring the magnitude of needed adjustment to the emissions.

We present in Fig. 20 the results of applying the above considerations to the comparison with the AERONET data. The far away distance to the source region of the American stations prevents us from applying the same reasoning to interpret the AERONET data in that region. It should be noted that for the judgement on the over/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 influence less so stations close to the source regions. For the present analysis we use AERONET data of the year 2000 from the African and Middle East sites (Fig. 13). According to this the AeroCom median and models ECMWF, LSCE and ECHAM5-HAM underestimate the dust emissions in Africa while model CAM overestimate them in this region (Fig. 20). For the other models either the AE was not available or the results were not conclusive to propose an over/under estimation of the emissions. In the Middle East, the AeroCom median and models ECMWF, LOA, LSCE, ECHAM5-HAM and TM5 underestimate the dust emissions whereas models CAM and UMI overestimate them (Fig. 20). We highlight that the analysis on the Middle East is based only on the station of Solar Village (Fig. 9). Yet 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 one of the models (Table 4). The regions are illustrated in Fig. 2. Some of the models simulate emissions outside of these regions. The models underestimating the emissions (highlighted in blue in Table 4) according to our analysis above correspond to most of the lower fluxes whereas model CAM, the one overestimating the emission has the highest flux after SPRINTARS (2888). In the Middle East, the fluxes identified as overestimating and underestimating the emissions cover a large range of fluxes. The ones overestimating range from 329 (UMI) to 526 (CAM) Tg/yr while those underestimating range from 25.6 (ECHAM5-HAM) to 212 (TM5) Tg/yr. In both regions models with fluxes in the range of those underestimating/overestimating the emissions were not identified as having too weak/strong emissions since the combination of AOD and AE did not allow concluding on the emission intensity. Assuming that the emission flux is unique for each region and based on our above results we suggest the emissions in the Sahara should be between 792 and 2271 Tg/yr and the ones in the Middle East between 212 and 329 Tg/yr.

The above is only an enveloping analysis. A more exhaustive study to estimate dust emissions was conducted by Cakmur et al. (2006) through matching model outputs to a wide variety of observations at a worldwide array of stations. According to their results the African emissions range between 964 and 1803 Tg/yr whereas the emissions in Arabia are 23 and 132 Tg/yr. The range of African emissions proposed in this work is consistent with the results from Cakmur et al. (2006), but large differences exist between the emissions in Arabia or the Middle East. While our Middle East emissions are based only on comparison with data at one station and one year of simulation, the results of Cakmur et al. (2006) are based on different data sets and five year of simulation.

### 4.4 General discussion

No AERONET station affected by the Asian dust source was selected according to the criteria used in this study preventing an evaluation of the performance of the models in simulating the dust cycle in Asia. Months with intense dust activities were masked out
by anthropogenic emissions generating AE values above 0.4 and therefore not recognisable with our definition of dust sites. However surface concentration measurements in Hedo and Cheju (stations 17 and 20 respectively in Figs. 2, 4 and 5), even though limited, give us some insight in the general model performance to simulate the Asian dust. All in all, the models reproduce the seasonal features of maxima and minima but fail to replicate the magnitude of the data. Furthermore, models in general simulate better the surface concentration in stations affected by Saharan dust than in stations affected by Asian dust. This last might indicate that models have difficulties in simulating simultaneously the Asian and Saharan dust cycle as already suggested by Tegen et al. (2002). A more specific dust data set would need to be found to verify this and examine the performance of global dust models in the Asian region. We therefore left it outside of this study. One way to assess the performance of global dust models over Asia would be comparing measurements of coarse mode AOD against modeled ones.

The models perform better in simulating the climatology of vertically integrated variables in dusty sites than they do with surface measurements, i.e. deposition and surface concentration. The modeled AOD is within a twofold range of the observations at most of the sites whereas for surface concentrations and total deposition the under/over estimation is mostly a factor of 10 with respect to the observations and even more. Data quality can explain this difference since AERONET’s climatology includes the simulated years whereas the deposition and surface concentration climatology do not. The surface measurements were considered as climatology in this study without being one in a strict sense. Furthermore, surface concentration and deposition simulations requires having the vertical distribution correct whereas for vertically integrated parameters such as AOD and AE the vertical distribution is almost not relevant (considering they are clear sky measurements of non-hygroscopic particle such as dust). In addition, this difference in performance might also reveal that more weight is giving in models to simulate correctly AOD and AE than surface variables.

This is the first multi-parameter and multi-model intercomparison conducted on global dust models. Fifteen models from the AeroCom intercomparison 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 verti-
cally 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. Previous efforts to create a dataset for model evaluations
are the works of Prospero et al. (2010); Prospero and Lamb (2003); Ginoux et al.
(2001); Mahowald et al. (2009) and the DIRTMAP data set. These works concentrated
mostly on a single parameter. We have grouped in a single database the data used
in these studies as to ease future comparison or 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 particu-
lar over remote regions in the Southern Hemisphere where the largest 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, addi-
tional surface concentration measurements are needed such as the ones taken during
the SEAREX and AEROCE campaigns and still measured in the Barbados and Mi-
ami. Since AOD is dominated by the fine mode due to its higher extinction efficiency
and that the coarse mode dominates the surface concentration and deposition it is im-
portant 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 current size distribution in models in the overestimation of deposition and surface
concentration. The AERONET network represents a crucial source of data in validat-
ing models. The information of this network should be complemented with satellite
products to further evaluate the model performance.
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, large diversities exist between the many global models that simulate the dust cycle and the impact of dust on climate. On one hand, these model diversities are the product of the scatter 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, that prevents one to 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 with 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. Additional comparisons to Angström Exponent (AE), coarse mode AOD and dust surface concentration are included to extend the assessment of model performance. Altogether this forms a new benchmark data set, which is available via the AeroCom data server for model inspection and future development of dust models. This data set is composed of four different dust deposition compilations of total yearly flux, a data set of monthly averages of wet and total deposition in Florida (USA), yearly and monthly averaged dust surface concentration across the world and total AOD, coarse mode AOD as well as AE at dusty sites from the AERONET network. The dusty sites were identified as stations with at least two month with averaged AE smaller than 0.4 and AOD larger than 0.2. We remind that for surface concentration and deposition the period the data were taken is not coincident with the simulated year but are still of great value when evaluating the model performance.
We recall the reader 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 supposedly improved since submitting their simulations. Therefore the results presented in this study do not necessarily represent the current state of the models.

The model performance to simulate surface concentration depends on the database used. All models agree in mainly overestimating the surface concentration taken during cruise campaigns and this mostly from a factor of ten up to two orders of magnitude. When using long-term measurements on the other hand the overestimation is within a factor of ten with respect to the observations. If the 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. In cases where models underestimate the surface concentration, they do so within a factor of ten. When this limit is exceeded then mostly with measurement in remote regions of the Southern Hemisphere. 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 underestimate the magnitude. Furthermore the performance deteriorates from south to north Florida reflecting difficulties to reproduce the dust transport northward. Equivalent long term measurement records are needed, ideally on a daily basis and over oceans, to evaluate models to reproduce the deposition fluxes. From the current comparison one may conclude with very large uncertainty, that the impact of dust on the ocean biogeochemical cycle is overestimated in most models.
The models mostly underestimate the surface concentrations at the sites under inspection here. Model performance is better at sites with a large range of measured surface concentrations, reflecting a 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 of the 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 West Pacific stations, which capture the Asian dust, most of the models simulate the general seasonal features of maxima and minima but fail to reproduce the intensity. Our models thus seem to simulate better the transatlantic dust transport from the Sahara than the one from Asian dust sources across the Pacific.

A similar conclusion on regional model performance, 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 more difficulties in reproducing the magnitude and seasonality in the Middle East of both, AOD and AE. Model performance to reproduce Asian dust could not be explored due to the definition of dusty sites used; months with intense dust activities were masked out by anthropogenic emissions generating AE values above 0.4. A different selection criteria or approach would need to be found to verify this and examine the performance of global dust models in this region. However, again the models reproduce the transatlantic transport of dust from the Sahara in terms of AOD and AE. They agree in reproducing the offshore transport of Saharan dust throughout the year as revealed by data from Capo Verde and limit the transport across the Atlantic to the summer months as revealed by the stations in Barbados and Roosevelt Roads, yet overestimating the AOD and transporting too fine particles. In contrast, almost no model reproduces the most southward displacement of the Saharan dust cloud during the winter captured by the AOD and AE data at Surinam, representative of the transport of dust into South America.
Models perform better in simulating the climatology of averaged vertically integrated parameters (AOD and AE) in dusty sites than total deposition and surface concentration. The modeled monthly average AOD and AE is within a factor of two of the observations at most of the sites whereas for surface concentrations and total deposition the under/over estimation is within a factor 10 of the observations and up to two orders of magnitude according to the dataset. The different characteristics of the data climatologies used, but also the differences in model vertical structure in simulating the dust deposition and surface concentration might explain this difference.

Based on the dependency of AOD and AE on aerosol burden and size distribution we use the simultaneous overestimation/underestimation of AOD and underestimation/overestimation of AE to suggest whether a model is overestimating/underestimating dust emissions. Note, that if AOD and AE bias in a given model is of equal sign 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 ones to underestimate the emissions are AeroCom median and models ECMWF, LOA, LSCE, ECHAM5-HAM, TM5 whereas models CAM and UMI overestimate them. According to these results we estimate the emissions in the Sahara to be between 792 and 2271 Tg/yr and the ones in the Middle East between 212 and 329 Tg/yr.

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 dust models. Dust deposition measurements are sparse and deliver mostly only total deposition fluxes for a given event or longer time period not necessarily coincident with the year simulated limiting the model evaluation. Permanent monitoring of dust deposition equivalent to the network presented in Prospero et al. (2010) are therefore needed.

The new round of experiments conducted within AeroCom phase II with additional diagnostics 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 its role 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 surface concentration above the surface at a few locations in order to allow comparisons in elevated mountain stations.

Supplementary material related to this article is available online at: http://www.atmos-chem-phys-discuss.net/10/23781/2010/acpd-10-23781-2010-supplement.pdf.

Acknowledgements. The authors would like to 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.

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


Han, Q., Moore, J. K., Zender, C., Measures, C., and Hydes, D.: Constraining oceanic dust


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.

<table>
<thead>
<tr>
<th>N</th>
<th>Model</th>
<th>Resolution</th>
<th>Characteristics of size distribution</th>
<th>Reference of emission scheme</th>
<th>Model Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CAM</td>
<td>2.8° × 2.8° × 26 levels</td>
<td>4 bins 0.1-1.0-2.5-5-10 µm</td>
<td>Zender et al., 2003; Mahowald et al., 2006</td>
<td>Mahowald et al., 2006</td>
</tr>
<tr>
<td>2</td>
<td>ECMWF</td>
<td>0.7° × 0.7° × 60 levels</td>
<td>3 bins 0.03-0.55-0.9-20 µm</td>
<td>Morcrette et al., 2009</td>
<td>Morcrette et al., 2009</td>
</tr>
<tr>
<td>3</td>
<td>GISS</td>
<td>5° × 4° × 20 layers</td>
<td>4 bins 0.1-1-2-4-8 µm</td>
<td>Cakmur et al., 2006</td>
<td>Schmidt et al., 2006; Bauer and Koch, 2005; Miller et al., 2006</td>
</tr>
<tr>
<td>4</td>
<td>GOCART</td>
<td>2° × 2.5° × 30 layers</td>
<td>5 bins 0.1-1.0-1.8-3.0-6.0-10.0</td>
<td>Ginoux et al., 2001</td>
<td>Chin et al., 2000</td>
</tr>
<tr>
<td>5</td>
<td>SPRINTARS</td>
<td>1.125° × 1.125° × 20 layers</td>
<td>6 bins 0.1-0.22-0.46-1.0-2.15-4.64-10.0</td>
<td>Takemura et al., 2009</td>
<td>Takemura et al., 2005</td>
</tr>
<tr>
<td>6</td>
<td>LOA</td>
<td>3.75° × 2.5° × 19 layers</td>
<td>2 bins 0.03-0.5-20 µm</td>
<td>Balkanski et al., 2004</td>
<td>Reddy et al., 2005a</td>
</tr>
<tr>
<td>7</td>
<td>LSCE</td>
<td>3.75° × 2.5° × 19 layers</td>
<td>1 mode mmr = 1.25 µm σ0 = 2.0</td>
<td>Balkanski et al., 2004</td>
<td>Schulz, 2007</td>
</tr>
<tr>
<td>8</td>
<td>MATCH</td>
<td>1.9° × 1.9° × 28 s layer</td>
<td>4 bins 0.1-1.0-2.5-5.0-10</td>
<td>Zender et al., 2003</td>
<td>Zender et al., 2003</td>
</tr>
<tr>
<td>9</td>
<td>MOZGN</td>
<td>1.9° × 1.9° × 28 layers</td>
<td>5 bins 0.1-1.0-1.8-3.0-6.0-10.0</td>
<td>Ginoux et al., 2001</td>
<td>Horowitz et al., 2003; Tie et al., 2005</td>
</tr>
</tbody>
</table>
### Table 1. Continued.

<table>
<thead>
<tr>
<th>N</th>
<th>Model</th>
<th>Resolution</th>
<th>Characteristics of size distribution</th>
<th>Reference of emission scheme</th>
<th>Model Reference</th>
</tr>
</thead>
</table>
| 10 | ECHAM5-HAM    | 1.8° × 1.8° × 31 layers | 2 modes
mmr = 0.37, 1.75 µm
σ0 = 1.5, 2.0 | Tegen et al., 2002 | Stier et al., 2005 |
| 11 | MIRAGE        | 2.5° × 2.0° × 24 layers | 4 modes
mmr = 0.03, 0.16, 2.1, 2.5 µm
σ0 = 1.6, 1.8, 1.8, 2.0 | Ginoux et al., 2001 | Ghan and Easter, 2006 |
| 12 | TM5           | 6° × 4° × global 1° × 1° × North America and Europe 25 layers | 2 modes
mmr = µm
σ0 = 1.59, 2.0 | Dentener et al., 2006; de Meij et al., 2006 | Krol et al., 2005 |
| 13 | 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 | Grini et al., 2005 | Berglen et al., 2004; Myhre et al., 2007 |
| 14 | UMI           | 2.5° × 2° × 30 layers | 4 bins
0.05-0.63-1.25-2.5-10 um radius | Ginoux et al., 2001 | Liu and Penner, 2002; Liu et al., 2007 |

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 σ0 = 2.0.
3 The mmr values are global annual averages. They vary spatially and temporally with the mode volume and number mixing ratios.
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).

<table>
<thead>
<tr>
<th>Model</th>
<th>AOD</th>
<th>AE</th>
<th>SCONC</th>
<th>DEPO</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECMWF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GISS</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GOCART</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>SPRINTARS</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>LOA</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>LSCE</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>MATCH</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>MOZGN</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>ECHAM5-HAM</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>MIRAGE</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>TM5</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>UIO_CTM</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>UMI</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>
Table 3. Mass balance for each one of the models. NaN represents variables not provided by the model. MEE corresponds to the mass extinction efficiency.

<table>
<thead>
<tr>
<th>N</th>
<th>Mode</th>
<th>Emission [Tg/yr]</th>
<th>Load [Tg]</th>
<th>Deposition [Tg/yr]</th>
<th>Wet Depo [Tg/yr]</th>
<th>Dry Depo [Tg/yr]</th>
<th>Sedim [Tg/yr]</th>
<th>OD550</th>
<th>Dust [m²/g]</th>
<th>MEE</th>
<th>Life Time [days]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AEROCOM_MEDIAN</td>
<td>1123</td>
<td>15.8</td>
<td>1257</td>
<td>357</td>
<td>396</td>
<td>314</td>
<td>0.023</td>
<td>0.72</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>CAM</td>
<td>4313</td>
<td>25.7</td>
<td>2058</td>
<td>1382</td>
<td>675</td>
<td>NaN</td>
<td>0.035</td>
<td>0.69</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ECMWF</td>
<td>514</td>
<td>54.7</td>
<td>5999</td>
<td>3248</td>
<td>2582</td>
<td>179</td>
<td>0.027</td>
<td>0.25</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>GISS</td>
<td>1507</td>
<td>29.0</td>
<td>1488</td>
<td>456</td>
<td>352</td>
<td>680</td>
<td>0.034</td>
<td>0.60</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>GOCART</td>
<td>3157</td>
<td>29.5</td>
<td>3178</td>
<td>583</td>
<td>120</td>
<td>2475</td>
<td>0.035</td>
<td>0.60</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>SPRINTARS</td>
<td>3995</td>
<td>17.2</td>
<td>3984</td>
<td>628</td>
<td>2791</td>
<td>565</td>
<td>0.024</td>
<td>0.72</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>LOA</td>
<td>1276</td>
<td>13.7</td>
<td>1275</td>
<td>417</td>
<td>521</td>
<td>336</td>
<td>0.034</td>
<td>1.28</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>LSCE</td>
<td>1158</td>
<td>20.3</td>
<td>1156</td>
<td>616</td>
<td>310</td>
<td>231</td>
<td>0.031</td>
<td>0.77</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>MATCH</td>
<td>981</td>
<td>17.3</td>
<td>1070</td>
<td>517</td>
<td>431</td>
<td>122</td>
<td>0.033</td>
<td>0.96</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>MOZGN</td>
<td>2371</td>
<td>21.1</td>
<td>2368</td>
<td>425</td>
<td>1943</td>
<td>NaN</td>
<td>0.022</td>
<td>0.52</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ECHAM5-HAM</td>
<td>664</td>
<td>8.2</td>
<td>676</td>
<td>374</td>
<td>37</td>
<td>265</td>
<td>0.010</td>
<td>0.60</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>MIRAGE</td>
<td>2066</td>
<td>22.0</td>
<td>2048</td>
<td>1361</td>
<td>687</td>
<td>NaN</td>
<td>0.053</td>
<td>1.22</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>TM5</td>
<td>1683</td>
<td>9.3</td>
<td>1682</td>
<td>295</td>
<td>592</td>
<td>794</td>
<td>0.013</td>
<td>0.68</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>UIO_CTM</td>
<td>1572</td>
<td>21.7</td>
<td>1571</td>
<td>681</td>
<td>890</td>
<td>NaN</td>
<td>0.026</td>
<td>0.61</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>UMI</td>
<td>1688</td>
<td>19.3</td>
<td>1691</td>
<td>619</td>
<td>1073</td>
<td>NaN</td>
<td>0.021</td>
<td>0.56</td>
<td>4.2</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Yearly emission fluxes [Tg/yr] for regions illustrated in Fig. 2. Fluxes being overestimated are highlighted in red and those underestimating are highlighted in blue.

<table>
<thead>
<tr>
<th></th>
<th>North Africa</th>
<th>Middle East</th>
<th>Asia</th>
<th>South America</th>
<th>South Africa</th>
<th>Australia</th>
<th>North America</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEROCOM MEDIAN</td>
<td>792</td>
<td>128</td>
<td>137</td>
<td>9.8</td>
<td>11.8</td>
<td>30.7</td>
<td>2.0</td>
</tr>
<tr>
<td>CAM</td>
<td>2271</td>
<td>526</td>
<td>727</td>
<td>13.7</td>
<td>2.9</td>
<td>12.2</td>
<td>286</td>
</tr>
<tr>
<td>ECMWF</td>
<td>204</td>
<td>68</td>
<td>125</td>
<td>1.0</td>
<td>16.3</td>
<td>57.0</td>
<td>15.3</td>
</tr>
<tr>
<td>GISS</td>
<td>1031</td>
<td>125</td>
<td>180</td>
<td>39.9</td>
<td>31.7</td>
<td>87.8</td>
<td>7.3</td>
</tr>
<tr>
<td>GOCART</td>
<td>1736</td>
<td>348</td>
<td>873</td>
<td>66.5</td>
<td>25.0</td>
<td>111</td>
<td>13.0</td>
</tr>
<tr>
<td>SPRINTARS</td>
<td>2888</td>
<td>531</td>
<td>363</td>
<td>6.9</td>
<td>113</td>
<td>36.8</td>
<td>4.1</td>
</tr>
<tr>
<td>LOA</td>
<td>772</td>
<td>114</td>
<td>411</td>
<td>0.5</td>
<td>3.5</td>
<td>14.9</td>
<td>4.5</td>
</tr>
<tr>
<td>LSCE</td>
<td>529</td>
<td>39.2</td>
<td>509</td>
<td>0.2</td>
<td>57.2</td>
<td>10.6</td>
<td>7.2</td>
</tr>
<tr>
<td>MATCH</td>
<td>539</td>
<td>241</td>
<td>100</td>
<td>19.3</td>
<td>24.5</td>
<td>40.9</td>
<td>2.4</td>
</tr>
<tr>
<td>MOZGN</td>
<td>1410</td>
<td>376</td>
<td>294</td>
<td>92.8</td>
<td>55.4</td>
<td>89.5</td>
<td>12.7</td>
</tr>
<tr>
<td>ECHAM5-HAM</td>
<td>401</td>
<td>25.6</td>
<td>54</td>
<td>3.7</td>
<td>40.2</td>
<td>58.4</td>
<td>1.7</td>
</tr>
<tr>
<td>MIRAGE</td>
<td>703</td>
<td>292</td>
<td>608</td>
<td>186</td>
<td>25.0</td>
<td>129</td>
<td>70.8</td>
</tr>
<tr>
<td>TM5</td>
<td>1091</td>
<td>212</td>
<td>253</td>
<td>30.4</td>
<td>15.3</td>
<td>59.4</td>
<td>8.1</td>
</tr>
<tr>
<td>UIO.CTM</td>
<td>1213</td>
<td>206</td>
<td>27</td>
<td>5.0</td>
<td>11.6</td>
<td>9.0</td>
<td>1.8</td>
</tr>
<tr>
<td>UMI</td>
<td>933</td>
<td>329</td>
<td>340</td>
<td>47.1</td>
<td>20.6</td>
<td>35.4</td>
<td>6.1</td>
</tr>
</tbody>
</table>
Fig. 1. Measured yearly deposition fluxes versus modeled ones; units are g/m$^2$/yr. 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. (2009) 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 scatter plot. Black continues line is the 1:1 line whereas the black dotted lines correspond to the 10:1 and 1:10 lines.
Fig. 1. Continued.
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 the Saharan 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 material. Rectangles illustrate regions defined to compute the emissions presented in Table 4.
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 year average. Units are g/m²/month.
Fig. 4. Scatterplot of yearly averaged surface concentration from Mahowald et al. (2009) versus modeled one. Measured Fe (and converted to dust by assuming a 3.5% Fe in dust) during cruise is represented by filled-in circle. Data corresponding to term measurement 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).
Fig. 4. Continued.
Fig. 5. Caption on next page.
Fig. 5. Yearly averaged measured surface concentration from the network operated by the university of Miami vs. modeled one at each station, units are µg/m³. 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 the Saharan and Asian dust source (blue). The location of each station is illustrated in Fig. 2. 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. Black continues line is the 1:1 line whereas the black dotted lines correspond to the 10:1 and 1:10 lines.
Fig. 6. Caption on next page.
Fig. 6. Monthly averages of measured and simulated surface concentration. Units are µg/m³. Each row corresponds to the seasonal cycle at one of the stations. 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. The row for each station corresponds to the number presented in Fig. 2. White color corresponds to month without measurements. Measured data are shown in the figure on the upper left side. For relative differences (in %) of model with respect to observations see Fig. S1 in supplement material.
Fig. 7. Measured and simulated surface concentration in Barbados and Miami. Measurements of the year 2000 are presented by the black continues line and the climatology (Fig. 6) is presented by the black dashed line. Units are µg/m³.
Fig. 8. Location of selected AERONET dusty sites based on the climatology build from the multi-annual database 1996–2006. Dusty stations are grouped regionally; Africa (orange), 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 material.
Fig. 9. Same as Fig. 8 but for stations selected based on data of the year 2000. Names and locations for each selected station are given in Tables S4 in the supplement material.
Fig. 10. Caption on next page.
Fig. 10. Averaged AOD at 550 nm vs. 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 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 material. 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. Black continues line is the 1:1 line whereas the black dotted lines correspond to the 2:1 and 1:2 lines.
Fig. 11. Caption on next page.
Fig. 11. AERONET AOD at 550 nm and model output at dusty stations. Each row corresponds to the seasonal cycle at one of the stations. They have been grouped into African (AF, orange), Middle East (ME, violet) and American (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 figures. 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 material. White color corresponds to month without measurements or month not complying with the selection criteria (Sect. 2.3). AERONET data are shown in the first figure on the upper left side and correspond to the climatology based on the multi-annual database 1996–2006. For relative differences (in %) of model with respect to observations see Fig. S2 in supplement material.
Fig. 12. Caption on next page.
**Fig. 12.** Same as Fig. 10 but when data of the year 2000 are used. Location of each station is illustrated in Fig. 9. Name and location of each station are given in Table S4 of the supplement material. 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. Black continues line is the 1:1 line whereas the black dotted lines correspond to the 2:1 and 1:2 lines.
Fig. 13. Caption on next page.
Fig. 13. Same as Fig. 11 but when data of the year 2000 are used. Location of each station is illustrated in Fig. 9. Name and location of each station are given in Table S4 in the supplement material. White color corresponds to month without measurements or month not complying with the selection criteria (Sect. 2.3). AERONET data are shown in the first figure on the upper left side. For relative differences (in %) of model with respect to observations see Fig. S3 in supplement material.
Fig. 14. Caption on next page.
Fig. 14. Same as Fig. 11 but for coarse mode AOD. Same stations as the ones used for total AOD (Fig. 8) are considered. Stations Bandoukouï (3), Bidi Bahn (7) and Cape San Juan (21) do not have coarse mode AOD for the selected period. For relative differences (in %) of model with respect to observations see Fig. S4 in supplement material.
Fig. 15. Caption on next page.
**Fig. 15.** Same as Fig. 13 but for coarse mode AOD. Same stations as the ones used for total AOD (Fig. 9) are considered. The Station of Roosevelt Roads (7) did not have coarse mode AOD data for the selected period. For relative differences (in %) of model with respect to observations see Fig. S5 in supplement material.
Fig. 16. Same as Fig. 10 but for Angström Exponent.
Fig. 17. Same as Fig. 11 but for Angström Exponent. For relative differences (in %) of model with respect to observations see Fig. S6 in supplement material.
Fig. 18. Same as Fig. 12 but for Angström Exponent.
Fig. 19. Same as Fig. 13 but for Angström Exponent. For relative differences (in %) of model with respect to observations see Fig. S7 in supplement material.
Fig. 20. Suggested over/under estimation of the emissions in Africa (left) and the Middle East (right) based on AERONET AE and AOD. Simultaneous overestimation of the AOD and underestimation of the AE suggests a overestimation of the emissions and vice versa. Overestimations are illustrated by red whereas underestimations are indicated by blue.