Sensitivity of the Interannual Variability of Mineral Aerosol Simulations to Meteorological Forcing Dataset

Molly B. Smith[1,2], Natalie M. Mahowald[1], Samuel Albani[1,3], Aaron Perry[1], Remi Losno[4], Zihan Qu[4,5], Beatrice Marticorena[5], David A. Ridley[6], Colette L. Heald[6]

5 [1] Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY 14850, USA.
[2] Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York, Albany, NY 12222, USA.
[3] Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, Gif-sur-Yvette, France.
[4] Institut de Physique du Globe de Paris, University of Paris Diderot, USPC, UMR CNRS 7154, Paris France.
10 [5] LISA, Universites Paris Est-Paris Diderot-Paris 7, UMR CNRS 7583, Creteil, France.
[6] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, MA 02139, USA.

Correspondence to: Molly B. Smith (mbs244@cornell.edu)

Abstract. Interannual variability in desert dust is widely observed and simulated, yet the sensitivity of these desert dust simulations to a particular meteorological dataset, as well as a particular model construction, is not well known. Here we use
15 version 4 of the Community Atmospheric Model (CAM4) with the Community Earth System Model (CESM) to simulate dust forced by three different reanalysis meteorological datasets for the period 1990-2005. We then contrast the results of these simulations with dust simulated using online winds dynamically generated from sea surface temperatures, as well as with simulations conducted using other modeling frameworks but the same meteorological forcings, in order to determine the sensitivity of climate model output to the specific reanalysis dataset used. For the seven cases considered in our study,
20 the different model configurations are able to simulate the annual mean of the global dust cycle, seasonality and interannual variability approximately equally well (or poorly) at the limited observational sites available. Overall, aerosol dust source strength has remained fairly constant during the time period from 1990 to 2005, although there is strong seasonal and some interannual variability simulated in the models and seen in the observations over this time period. Model interannual variability comparisons to observations, as well as comparisons between models, suggest that interannual variability in dust
25 is still difficult to simulate accurately, with averaged correlation coefficients of 0.1 to 0.6. Because of the large variability, at least one year of observations at most sites are needed to correctly observe the mean, but in some regions, particularly the remote oceans of the Southern Hemisphere, where interannual variability may be larger than in the Northern Hemisphere, 2-3 years of data are likely to be needed.

1. Introduction

30   Mineral aerosols, or desert dust, are soil particles suspended in the atmosphere, and are intimately connected with

many earth system processes (e.g. Shao et al., 2011). Mineral aerosols both scatter and absorb incoming solar radiation and outgoing long wave radiation (Tegen and Lacis, 1996;Sokolik and Toon, 1996;Miller and Tegen, 1998a;Dufresne et al., 2002) and thus impact directly the radiative budget of the Earth (e.g. Balkanski et al., 2007;Reddy et al., 2005;Myhre et al., 2013). In addition, dust can interact with water and ice clouds, indirectly impacting climate by changing cloud properties or

5   lifetimes (e.g. Rosenfeld et al., 2001;DeMott et al., 2003;Mahowald and Kiehl, 2003;Atkinson et al., 2013;Cziczo et al., 2013). Deposited desert dust can impact snow albedo (e.g. Painter et al., 2007), as well as provide important micronutrients to land and ocean ecosystems (e.g. Swap et al., 1992;Okin et al., 2004;Jickells et al., 2005).

Dust is widely variable in space and time, with 4-fold fluctuations in surface concentration observed across a large region on decadal time scales (Prospero and Lamb, 2003), and globally equally large fluctuations on glacial-interglacial time

10   scales (Kohfeld and Harrison, 2003). The importance of interannual variability of dust in changing modifying precipitation and temperatures (Yoshioka et al., 2007;Evan et al., 2009;Mahowald et al., 2010) and biogeochemistry (Aumont et al., 2008;Doney et al., 2009) have previously been simulated and shown to potentially be large. Previous model intercomparison studies have shown that the models have some skill in simulations of the annual mean and seasonal cycle (Huneeus et al., 2011), and some studies have sought to consider the causes of interannual variability in dust, such as changes in

15   precipitation, winds, surface roughness or land use (e.g. Tegen and Miller, 1998;Mahowald et al., 2002;Miller and Tegen, 1998b;Mahowald et al., 2003;Ginoux et al., 2004;Cowie et al., 2013; Ridley et al., 2014).

It is well established in the dust literature that meteorology and surface conditions play central roles in driving changes in dust emissions, primarily from changes in precipitation, winds, surface roughness or vegetation cover on daily to interannual to geological time scales (e.g. Westphal et al., 1987; Petit et al., 1999; Marticorena and Bergametti, 1996;

20   Mahowald et al., 2002; Prospero and Lamb, 2004; Engelstadter and Washington, 2007; Engelstaedter et al., 2003; Roe, 2008; McGee et al., 2010; Knippertz and Todd, 2012; Cowie et al., 2013; Yu et al., 2015), and we do not seek to repeat or review that work here. Rather, we address the question of how sensitive our simulation of interannual variability is to the meteorology, or alternatively, the modeling framework used. While previous modeling studies have evaluated the annual mean and seasonal cycle coherence across dust models (Huneeus et al., 2011) or contrasted specific models (Luo et al.,

25   2003), prior studies exploring the ability of models to simulate interannual variability and the role of different meteorological mechanisms have used only one model (e.g. Miller and Tegen, 1998b;Mahowald et al., 2002;Mahowald et al., 2003;Ginoux et al., 2004;Ridley et al., 2014). Thus an open question remains of how robust our simulations of inter-annual variability are, and how sensitive they are to the meteorological dataset versus the modeling framework. Characterizing how well IAV is simulated in models allows a better understanding of how much we should trust model

30   output in studies directed at understanding the role of dust IAV in contributing to IAV in total aerosol AOD variability (e.g. Streets et al., 2009) or in dust impacts on ocean biogeochemistry IAV (e.g. Doney et al., 2009). A related question is how much observational data do we need in order to correctly characterize the mean dust amount, based on how much interannual variability we think exists in different locations. Note that the global model simulations used here may miss small scale feature that may be important for IAV, such as dust devils or moist convective events (e.g. Renno et al., 2000;

Marsham et al., 2011).

Here in this study, we use three different reanalysis meteorological datasets, online dynamic winds, and different modeling frameworks to try to understand how robust interannual variability in simulated dust is across 1990-2005. Our emphasis is on conducting sensitivity studies comparing the importance of different model meteorological datasets and

5   frameworks, but we do include some comparison to observations. While we focus on interannual variability, we will contrast that to seasonal variability, which has been commonly evaluated in models and previous intermodel comparisons (Huneeus et al., 2011). The period between 1990-2005 was chosen for this study because it has more available observational and reanalysis data than other years, but it must be noted that this time range does not have as much variability as previous periods (e.g. dry 1980s versus wet 1960s in the Sahel region; Prospero and Lamb, 2003). Model results are compared to

10   limited available in situ concentration, deposition and AOD data, in order to evaluate the models' ability to simulate the spatial and temporal variability observed in the dust cycle. In order to simplify the paper, we will focus on IAV in surface concentrations, and provide information on how deposition and aerosol optical depth contrast with surface concentration variability. Section 2 describes the methods used in the study, including briefly describing the models, data and comparison metrics. Section 3 describes the results of the study, starting with comparison to observations, comparison between different

15   model simulations, and the implication for observational needs.

2. Methodology

2.1 Model Description

Several models are used in this study, all of which include prognostic dust. The atmospheric component of an earth system model, the Community Atmospheric Model (CAM4) of the Community Earth System Model (CESM) (Neale et al., 2013;

20   Hurrell et al., 2013) is capable being forced either by online-calculated dynamical winds or by reanalyses datasets and is the model used for the bulk of the analysis to test sensitivity to which reanalyses winds are used (Section 2.1.1 and 2.1.2). To contrast with this model, results from other models, including the CAM5 version of the same model, and two chemical transport models, driven by reanalyses are also used (Model of Atmospheric Transport and Chemistry or MATCH and GEOS-chem) (Section 2.1.3). More details are described below, along with a summary of the model simulations (Table 1).

25   **2.1.1 CAM4 dust mode**l

For the bulk of the analysis in this study, simulations using the CESM were conducted focusing on the atmospheric model, CAM4 (Neale et al., 2013; Hurrell et al., 2013). This model is capable of simulations based on prognostic dynamic meteorology, as conducted for long climate simulations, or simulations forced to follow specific meteorological events, which allows comparison to specific observational data or field campaigns (Neale et al., 2013). The CESM model includes

30   four main global climate model (GCM) components: atmosphere (in this case, CAM4), land, ocean, and sea ice, all linked by a flux coupler. However, only the land and atmosphere components were prognostic for these simulations, with prescribed ocean and sea ice being used instead.

Dust is entrained into the atmosphere when strong winds occur in dry, unvegetated regions with easily erodible soils (Marticorena and Bergametti, 1995), using the Dust Entrainment and Deposition module (Zender et al., 2003a). There is a dust source when the leaf area index is sufficiently low (<0.3 m^2/m^2) and the soil moisture modifies the threshold wind velocity, as described in more detail in previous studies (Zender et al., 2003a; Mahowald et al., 2006). For most of the

5   simulations used here the CAM4 is used with the bulk aerosol module (BAM), This module includes four size bins: 0.1 to 1.0 μm in diameter, 1.0 to 2.5 μm, 2.5 to 5.0 μm, and 5.0 to 10.0 μm (Zender et al., 2003a; Mahowald et al., 2006), with the source size distributions described in Albani et al. (2014), following Kok (2011). The mass fraction of dust contained in each of the four bins is allowed to change over time as aerosols are transported and deposited out of the atmosphere (Zender et al., 2003a; Mahowald et al., 2006). Transport occurs through the CAM4 tracer advection scheme (Neale et al., 2010). Dry

10   deposition includes both gravitational and turbulent settling (Seinfeld and Pandis, 1998;Zender et al., 2003a), while wet deposition includes both convective and stratiform precipitation, incorporating prescribed solubility and parameterized scavenging coefficients (Mahowald et al., 2006;Albani et al., 2014). Emission over different soil types is parameterized by a geomorphic soil erodibility coefficient (Zender et al., 2003b), following the preferential source ideas of Ginoux et al. (2001). Finally, regional soil erodibility is optimized using the methodology described in Albani et al. (2014), by applying scale

15   factors to existing soil erodibility parameters for macro regions, in order to best match available data (Albani et al., 2014). For example, in most versions of the model the tuning reduced dust source strength over Central Asia and the Atacama Desert, and increased dust source strength over Argentina (Albani et al., 2014), which becomes important when analyzing dust IAV, as discussed in the results sections (Section 3).

20   **2.1.2 Meteorological forcing datasets**

In the CAM model, the reanalysis forcing can be used to nudge the model close to specific weather patterns so that events can be simulated (Lamarque et al., 2011) using a similar methodology to one used elsewhere (e.g. Model of Atmospheric Chemistry and Transport or MATCH; Rasch et al., 1997; Mahowald et al., 1997). In this procedure the

25   horizontal wind components, air temperature, surface temperature, surface pressure, sensible and latent heat flux and wind stress are read into the model simulation from the input meteorological dataset. These fields are subsequently used to internally generate (using the existing CAM4 parameterizations) the variables necessary for (1) calculating subgrid scale transport including boundary layer transport and convective transport; (2) the variables necessary for specifying the hydrological cycle, including cloud and water vapor distributions and rainfall (see Lamarque et al., 2011 for more details;

30   developed in Rasch et al., 1997;Mahowald et al., 1997). While this approach has the advantage of increased versatility with the input reanalysis dataset, there are inconsistencies between the model and reanalyses or observations, which leads to the model being nudged back towards the reanalyses. These inconsistencies mean the model has an anomalous source of energy or water. The approach used in MATCH and CAM was developed to minimize these inconsistencies (e.g. Mahowald et al., 1995; Mahowald et al., 1997). Previous studies have shown that the MATCH/CAM framework can reproduce the reanalysis

precipitation to a very high degree at the sub-daily to monthly to annual time scales (Mahowald et al., 1997;Mahowald, 1996).

Three different reanalysis meteorological datasets were used to simulate dust entrainment, transport and removal in the CAM4 simulations: MERRA (Modern Era-Retrospective Analysis for Research and Applications version 1; Rienecker et al., 2011), NCEP (National Centers for Environmental Prediction)-NCAR (National Center for Atmospheric Research) 50-year reanalysis (Kistler et al., 2001), and ECMWF (European Center for Medium-Range Weather Forecasts) ERA-Interim (Dee et al., 2011). Within the meteorological literature there are many studies contrasting these datasets to available observations, and showing the errors in the reanalyses, especially the moisture transports and precipitation (Trenberth and Guillemot, 1998;Trenberth et al., 2000; Treberth et al., 2011;Trenberth and Fasullo, 2013). Reanalyses should not themselves be considered observations, but are the closest representation we have to observed meteorology, that can drive chemical transport models, and thus represent an important resource. Here in this paper we supplement the meteorological analysis of different reanalysis datasets by contrasting how they impact dust emission, transport and deposition.

A fourth simulation was also conducted using AMIP-type protocol (Atmospheric Model Intercomparison Project; Gates et al., 1999). For AMIP simulations, the monthly mean sea surface temperatures are used to force the model online meteorology, but no atmospheric fields are used to constrain the model, in contrast to the reanalysis-driven simulations described above. Because in this case there is no inconsistency between the atmospheric model and the reanalysis, which can result in sources or sinks of water or energy, it is often considered a more robust way to simulate water vapor and thus chemistry (e.g. Hess and Mahowald, 2008;Trenberth and Guillemot, 1998;Trenberth et al., 2000). However, AMIP simulations cannot simulate exact weather events, but only interannual variability.

All the CAM4 simulations were conducted using a ~2x2° horizontal resolution. Not all of the input reanalysis data was available in the format required for CAM4 for the entire satellite era, so most of the analysis was conducted over a time period of 15 years when all the input data was accessible. There are parts of the analysis for which data outside this time period was used, but this is always indicated in the results. A summary of the model simulations and time periods of data availability are shown (Table 1). The first year of each simulation was neglected to allow for spinup.

For two of the model simulations used here (CAM4 (MERRA) and CAM4 (NCEP)), additional aerosol species were available for analysis (Sea salts, black carbon (BC), organic carbon (OC), and sulfate (SO_4)), and we use these to contrast the correlations between dust aerosols and other aerosols for a sensitivity study, referenced in Section 3.3.

2.1.3 Other dust models

In this paper, we also include sensitivity studies, where additional models are used, to contrast the importance of meteorological dataset with model construction. We briefly describe here the other models used in the study, with an emphasis on contrasting the differences in the model construction, but refer the interested reader to the specific model descriptions elsewhere.

An AMIP-style simulation was also conducted using the CAM5 model for comparison with the CAM4 AMIP

simulation. While the dust generation module in the CAM5 is identical to the CAM4, there are significant differences in the physics (Hurrell et al., 2013), as well as the aerosol formulation, which has implications for dust (Albani et al., 2014). The CAM5 model includes new planetary boundary layer, radiation, and moist convective parameterizations (Hurrell et al., 2013), in addition to a modal aerosol module (Liu et al., 2011;Ma et al., 2013). Similar to CAM4, the dust simulations in the

5   CAM5 were evaluated and tuned as in Albani et al. (2014). CAM5 assumes all aerosols are internally mixed (all aerosols in the same size are assumed to be mixed together for radiative forcing calculations), in contrast to CAM4, where aerosols are assumed to be externally mixed. CAM5 also allows aerosol indirect cloud effects to be calculated (e.g. Wang et al., 2011), although these are not evaluated here. Like the CAM4 simulations, CAM5 calculations were performed using a ~2x2° horizontal resolution.

10   Dust simulations using the GEOS-chem model (version v9-01-03; http://www.geos-chem.org/) from Ridley et al. (2014) were also considered here (Table 1). The model includes similar processes as the CAM4 model, including externally mixed aerosols, but the GEOS-chem model is forced here only by MERRA-1 winds (referred to as GCHEM(MERRA) here). GCHEM (MERRA) uses the DEAD dust module (Zender et al., 2003a), and employs the dust source function derived from (Ginoux et al., 2001). More details on the dust simulations from the GCHEM (MERRA) model can be found in (Ridley et

15   al., 2013;Ridley et al., 2014). Notice that the version used here for these comparisons does not include source function derived from Koven and Fung (2008) or the vegetation phenology and interannual variability, as included in the study focusing on African emissions (Ridley et al., 2013;Ridley et al., 2014). The horizontal resolutions of the model simulations were all 2x2.5°, and were interpolated onto the CAM grid for analysis in the paper. This model uses a similar dust entrainment scheme as CAM, but has different size distribution, and deposition mechanisms, as well as different boundary

20   layer, and moist convection physics.

It should also be noted that there is a difference between how the GCHEM (MERRA) model and CAM models incorporate reanalysis meteorology. The GCHEM (MERRA) model reads in all meteorological parameters from the reanalysis datasets, including turbulent and moist convective mixing, as well as the hydrology. This has the advantage that the chemical transport model does not have to rederive the hydrological cycle or mixing, which can be difficult (e.g.

25   Mahowald et al., 1995;Rasch et al., 1997). On the other hand, it makes the model less flexible, as it can only be operated using reanalysis datasets with mixing parameters output at the right frequency, in contrast to the MATCH or CAM framework (e.g. Rasch et al., 1997;Mahowald et al., 1997). This means that although both the CAM4 (MERRA) and GCHEM (MERRA) models are forced by MERRA winds, the surface winds and transport may still be slightly different.

Finally, simulations using the Model of Atmospheric Transport and Chemistry (MATCH) (Rasch et al., 1997) with

30   NCEP reanalysis data (Mahowald et al., 1997;Kalnay et al., 1996) are included. These simulations use a similar entrainment and deposition scheme (Zender et al., 2003a), with a simple wet removal scavenging coefficient (Luo et al., 2003), and have been extensively compared against observations (Luo et al., 2003; Luo et al., 2004; Mahowald et al., 2003). In contrast to the CAM models discussed earlier, there is no vegetation phenology included. Instead, the preferential source term from Ginoux et al. (2001) was used. The horizontal resolution of the MATCH (NCEP) model used here is 1.8x1.8°, and the

results were interpolated to the CAM grid before comparison to the other models.

Note that only the GCHEM (MERRA) model simulations are independently developed: the others all come from the same group, which previous studies have suggested that the group developing climate models matters substantially in their behavior (Knutti et al., 2013). There are however, mean differences even between the CAM simulations, such as in dust
5 vertical distributions and transport in relation to vertical mixing (Albani et al., 2014), and the strength of the Sahel source (e.g. Scanza et al., 2015).

2.2 Observational Data Description

For completeness, we compare standard annual means from each simulation to available data (e.g. Ginoux et al.,
10 2001;Huneeus et al., 2011) in order to show that the mean dust cycle is reasonable in our models. In this study, we use the annual mean surface concentration, aerosol optical depth, and deposition compilations from Albani et al. (2014).

For the variability studies, data was chosen for overlap with model runs and availability of longer datasets to evaluate interannual variability. In situ observations from the University of Miami network (Prospero and Nees, 1986;Prospero et al., 1996;Arimoto et al., 1990;Arimoto et al., 1997) were used here, as compiled in (Luo et al., 2003;Mahowald et al., 2003) and
15 updated at some sites (Prospero and Lamb, 2003). In addition, in situ observations from 3 sites in North Africa from the AMMA (African Monsoon Multidisciplinary Analysis) campaign were included (Marticorena et al., 2010). Details on the individual sites and locations are included in Table 2, and are shown in Figures S1, S3 and S5.

Sun photometer-derived aerosol optical depth is included from the AERONET (AErosol RObotic NETwork) database (Holben et al., 2000). Only sites where more than 50% of the modeled aerosol optical depth was from dust are included in
20 this comparison (as filtered in Figure 1 using in Mahowald et al., 2007), and only sites with more than 18 months of data are used here to estimate variability (Table 2). Because both seasalts and dust occur in the coarse mode, we cannot use remote sensing measurements of the coarse versus fine mode to identify dust dominated stations.

Two sites in the Southern Hemisphere are included (bottom of Table 2): surface concentration data from Rio Gallegos (Zihan, 2016) and deposition data from Kerguelen (Heimburger et al., 2012). Because of the limited datasets in the
25 Southern Hemisphere, we will consider these sites separately in Section 3.3. Although there are not AERONET data using these Northern Hemisphere criteria available in the Southern Hemisphere, if we focus on the coarse mode, there is one station which is dominated by dust (i.e. far from the coasts, where sea salts would make up a large percentage of the total aerosol load): Tinga Tingana in southeastern Australia. Located in a region of predominately westerly winds, Tinga Tingana lies downwind of the central Australian desert, allowing desert dust to feature prominently in its aerosol load. Tinga Tingana
30 also has a data record beginning in September of 2002, which overlaps our chosen time period by more than three years, making this the best Southern Hemisphere site to evaluate our modeled AOD variability.

For evaluation of the models' precipitation, we use the CPC Merged Analysis of Precipitation (CMAP) (http://www.esrl.noaa.gov/psd/data/gridded/data.cmap.html) (Xie and Arkin, 1997). This is a combination of in situ and

satellite observations, as well as models, to present the best estimate of precipitation.

2.3 Analysis Methods

For comparison of the variability in the modeled and observational values, we define variability similarly to previous studies (Mahowald et al., 2003):

$$Variability = \frac{\sigma}{\mu} \tag{1}$$

where σ is the standard deviation of the modeled and observed values, and μ is the mean.

Here, we focus on interannual variability (IAV), so that the annual means are used in equation (1) for that calculation. However, in order to contrast with the variability in the seasonal cycle, we will also report the strength of the seasonal cycle in terms of variability, where the values included are the climatological monthly means.

10 Models and observational time series of in situ concentrations and AOD (Table 2) were also correlated, using rank correlations to assess the ability of the models to simulate variability (similar to Mahowald et al., 2003). Rank correlations are used in order to reduce the importance of individual, extremely high data points, which can dominate regular correlations (Wilks, 2006). Similar to the variability, most of the analysis in the paper focuses on of the annual means, which test the ability of the models to simulate interannual variability, but some information about seasonal correlation is also provided, in

15 order to provide context and comparison to previous studies.

We also analyze the observations and model output for trends, by calculating the least squares fit slope, as well as the standard deviation in the slope.

In order to understand how similar model results are, we also calculate the variability at each grid box and compare different models at both the interannual and seasonal time scale. We also correlate the time series of models at individual

20 grid boxes across model simulations, again on the two different time scales.

To show the sensitivity to meteorology, we correlate the 3 CAM4-reanalysis simulations (CAM4-RE) which will give us 3 different correlation coefficients: CAM4-MERRA vs. CAM4-NCEP; CAM4-MERRA vs. CAM4-ERAI; and CAM4-NCEP vs. CAM4-ERAI and then average at each grid point the three different correlation coefficients to find the average correlation. Similar results are conducted using the AMIP simulations (CAM4-AMIP vs. CAM5-AMIP), and for the model

25 simulations using the exact same meteorology (CAM4-MERRA vs. GCHEM-MERRA; and CAM4-NCEP vs. MATCH-NCEP).

Finally, we use the model values to estimate the number of monthly mean observations required to correctly estimate the climatological annual mean value over 1990-2005. To do this, we assume that we would like to have a 95% chance to be within one standard deviation of the climatological mean. 1000 Monte-Carlo simulations were conducted, and each time we

30 chose randomly from the modeled monthly mean values at each grid point, and for each number of observations (between 1 and 50) we calculated the percentage of the time that the mean is within one standard deviation of the climatological mean of the 1990-2005 simulation. At every gridbox, the number of observations that would meet the 95% criteria is then calculated,

providing an estimate of the number of months of observations required.

Note that the modeled monthly mean values are not at all Gaussian distributed, and thus normal methods for determining the number of observations would not work (e.g. Wilks, 2006). Thus, for this analysis, we use we use rank correlations, which works with non-gaussian data. To be consistent with the climate model community (Taylor, 2001; Gleckler et al., 2008), for mean and standard deviation analysis described above, we use these standard metrics, despite the fact that our datasets do have not Gaussian distribution, which will lead to some errors in our results.

3. Results and Discussion

3.1 Comparison of Model and Observational variability

The annual mean distribution of the model simulations included here are evaluated elsewhere in more detail, since many of these model results were previously published (Luo et al., 2003;Huneeus et al., 2011;Albani et al., 2014;Ridley et al., 2013;Ridley et al., 2014), but for completeness we repeat comparisons of annual mean surface concentration, AOD, and deposition between available observations and the model simulations in the online supplement (Table S3; Figure S1-S6). Concentrations vary over several orders of magnitude spatially, and the models are able to simulate these variations (Figure S1, S2). In addition, the models can be shown to be mostly accurate in simulating the observed dust AOD and deposition (Figure S3-S6). Most of the model versions presented here do an equally good job when compared against the observations (Table 3). Recognize that the CAM4 and CAM5 simulations were tuned against these same observations (Albani et al., 2014), while the MATCH (NCEP) and GCHEM (MERRA) models were previously compared to similar observational syntheses (Luo et al., 2003;Huneeus et al., 2011;Ridley et al., 2013;Ridley et al., 2014).

The mean source flux and globally averaged AOD of the three CAM4-RE is 2400 +/- 26% Tg/yr and 0.026 +/- 30%, respectively, while for all the models included here the mean emission flux and AOD is 2400 +/- 26% and 0.025 +/- 40%, respectively. These ranges are similar to previous studies (Huneeus et al., 2011; Reddy et al., 2005). Note that using a similar model (CAM4-RE) and similar methodology to constrain the dust AOD, based on a combination of surface concentration, deposition and AOD in dust regions (Albani et al., 2014) obtains an uncertainty just due to meteorology of 30%. A more recent estimate, based more extensively on remote sensing data with limited information from 4 different models, but without using deposition or surface concentration data, finds a higher AOD of 0.033+/- 0.006 than found here, and a much smaller error estimate (Ridley et al., 2016). Thus there are large differences in the deduced AOD and uncertainties depending on assumptions about how to include different data, as well as the details of the models and methodology used.

Since this study is focused on an inter-comparison of different model simulations, rather than an evaluation of a specific model, we conduct limited comparison to observations. We focus on the highest quality data, coming from in situ concentrations and sun photometry data in dust dominated regions (e.g. Prospero and Lamb, 2003;Holben et al., 2000; Table

2), and ignore satellite based measurements (e.g. Torres et al., 2002;Evan et al., 2006) and dust visibility data (Mahowald et al., 2007) which are more difficult to interpret, both because they are not always only dust aerosols, and because they can have larger errors and thus be more difficult to compare for interannual variability (Torres et al., 2002;Evan et al., 2006;Mahowald et al., 2007). Some previous studies have included comparison of these in situ and sun photometry data in terms of interannual variability for specific model simulation evaluation (e.g. Mahowald et al., 2003;Ridley et al., 2014). The vertical distribution of the aerosol can also vary depending on the meteorology used (e.g. Albani et al., 2014), which may introduce some additional variability and discrepancy for in-situ ground based measurements.

Focusing on the amount of IAV, the models tend to simulate values between 0.1 and 0.8 for the variability (standard deviation of annual means divided by climatological annual mean; Section 2.3) at the observing sites, with largest values found at Mace Head in the models, but Bermuda in the observations. The models tend to simulate more IAV in AOD than in concentration (Figure 1 vs. 2; 3a vs. 3b), while the observations suggest similar amounts of variability (Figure 1 vs. 2; 3a vs. 3b). Both the models and the observations suggest much more variability (>2-fold) in the seasonal cycle than in the IAV (contrast Figure 1 vs Figure S7 or Figure 2 vs. Figure S8; Table S1 vs. S2; or Figure 3c and 3d) at many sites (Banizoumbou, Barbados, Bermuda, Cinzana, Miami, Midway Dalanzadgad and Sede Boker), and only slightly larger (1-2-fold) at the other sites.

Next we evaluate the ability of the models to simulate the high and low annual means using rank correlations. Most of the models do not have statistically significant correlation coefficients (Figure 3e and 3f; Table S3). The models do much worse at simulating IAV, than the seasonal cycle (contrast Figures 1 and 2 with Figures S7 and S8; Figure 3e and Figure 3f with Figure 3g and 3h), since at most stations, the models have a statistically significant correlation for the seasonal cycle and simulate a similar amount of variability over the seasonal cycle. For the seasonal cycle, the exceptions are at Izana and Ilorin for all of the models, and the CAM4 (AMIP) simulation, which is not statistically significantly correlated at most of the stations (Table S3; Figure 3e and 3f). For the CAM4, forcing with only SSTs substantially degrades the ability of the dust model to simulate the seasonal cycle, while in CAM5, the seasonal cycle is better simulated.

In model intercomparisons it has been observed that the model mean often does a better job than the individual model simulations (e.g. Flato et al., 2013). We evaluate this in the case of dust using the average of the CAM4 reanalysis models (CAM4 (MERRA), CAM4 (NCEP) and CAM4 (ERAI)). At the observational sites considered here, we do not see a large increase in the correlation coefficients for the model average versus individual models for either the IAV (Table S3) or the seasonal cycle (Table S4).

Overall this section supports previous studies (Prospero, 1996;Mahowald et al., 2003;Ginoux et al., 2004;Marticorena et al., 2010) suggesting that at the limited observational stations (Table 2) seasonal variability is larger than interannual variability (Table S2 and S3; Figure 3c and 3d), and that the model can simulate seasonal variability better than interannual variability in both surface concentrations and AOD (Figure 1 vs. S7; Figure 2 vs. S8; Figure 3e vs. 3g and Figure 3f vs. 3h). Taken overall, the models driven by reanalysis winds (all except CAM4 (AMIP) and CAM5 (AMIP)) compare roughly similarly against the available observations for both seasonal cycle and interannual variability (Figure 3).

New in this section is the evaluation of the relative ability of reanalysis driven models versus sea surface temperature forced models (CAM4 reanalysis versus CAM4 (AMIP)), which suggest a degradation in the ability of the models driven only by sea surface temperature to simulate both seasonal and interannual variability, although this is dependent on which model version (CAM4 vs. CAM5; Table S3 and S4; Figure 3). This correlation potentially provides insight into how much of the variability in dust is driven by sea surface temperatures. There are also significant differences between models driven by the same meteorology, for example (CAM4 (MERRA) versus GCHEM (MERRA) and CAM4(NCEP) and MATCH(NCEP), highlighting the importance of model formulation as well as meteorology. We will return to these points in later sections.

3.2 Comparison to trends in observations in the North Atlantic

Recent studies have highlighted the importance of fluctuations in rainfall in the Sahel for driving interannual variability and decadal scale variability in North Atlantic dust concentrations as seen in Barbados (e.g. Prospero and Lamb, 2003), although the importance of land use, winds, surface roughness and vegetation changes have been noted as well (e.g. Marticorena and Bergametti, 1996; M'bourou et al., 1997; Mahowald et al., 2002; Mahowald et al., 2007; Cowie et al., 2013). Since 2000, it has been noted that the Sahel precipitation no longer anti-correlates with dust at Barbados, suggesting a different mechanism may have become important. (Prospero, 2006; Mahowald et al., 2009). Ridley et al. (2014) proposes the hypothesis that the observed decrease in dust from 1982 to 2008 at Barbados is controlled by source wind strength over source regions in North Africa. For this argument, they use model evaluation with the GCHEM (MERRA) model (a different version of which is included also here), as well as analysis of ERAI and NCEP reanalysis winds and other observations (Ridley et al., 2014). Indeed, station data in the North African region, especially the Sahel, support the idea that winds decreased in this region between the late 1970s and 2003 (Mahowald et al., 2007), and there is also an observed widespread decrease in surface winds across many land regions (McVicar et al., 2012). Of course, there are many issues with the observation of surface winds due to small scale effects of buildings or topography (e.g. discussed in McVicar et al., 2012), so it is unclear how robust trends in observed surface winds are. Note that data correlations between visibility and winds suggest that both precipitation (Prospero and Lamb, 2003) and winds (Engelstaedter and Washington, 2007) are important for changes in dust near the source regions of North Africa (Mahowald et al., 2007). Importantly, Cowie et al. (2013) argue that the trends in surface winds and dust could be from changes in vegetation through the mechanism of surface roughness, which would link these changes also to precipitation. Note that because model calculated surface roughness was not archived in the models, we cannot test the Cowie et al. (2013) hypothesis directly in this study.

Here we can consider whether the hypothesis put forward in Ridley et al. (2014), that the decrease in winds in the surface region is responsible for the observed annually averaged decrease in surface concentration at Barbados, is consistent with the simulated trends in the multiple models included in this analysis. For this part of the paper, we use the full time period of our models, although for some models only some of the 1982-2008 time period is available (Table 1). For simplicity we consider only annual averages. As shown in (Ridley et al., 2014), there is a statistically significant downward

trend in the data at Barbados (Table 4). All the model versions considered here simulate a downward trend in the data at Barbados, although for some models this is not a statistically significant trend (CAM4 (NCEP); CAM5 (AMIP)). Only one model simulates the slope within one standard deviation of the observed value: the CAM4 (MERRA) (Table 4). The GCHEM (MERRA) overpredicts the magnitude of the negative slope as shown by Ridley et al. (2014), while the other

5   model versions underpredict the magnitude of the slope (Table 4).

Only some of the models see a statistically significant decrease in source strength in North Africa (Table 4), and some models predict an increase over this time period. However, a look at the trends in surface concentration across the models show that all the models see a decrease in surface concentrations that extends from the Sahel area of North Africa across the Tropical North Atlantic to Barbados (Figure 4), supporting the idea that the source strength is decreased over the time period

10   1990-2005. While individual models might simulate downward or upward trends elsewhere, this is the only region that sees a consistent model signal across this time period (Figure 4). If we focus on the Sahel (western) area of North Africa, indeed, most of the models simulate a decrease in the source (exceptions are CAM4 (ERAI) and CAM5 (AMIP)) (Table 4). Since the visibility data in North Africa also suggests a decrease across this time period in the western Sahel (Mahowald et al., 2007), these support the idea that the decrease in the source is the cause of the decrease in Barbados surface concentrations.

15   In the CAM4 models, the strongest correlation in IAV of the source occurs with surface winds (Table S5), and indeed in all the models there is a decrease in surface winds over this time period over the source regions, as seen in (Ridley et al., 2014) (Table S6).

There is also a correlation between Sahel precipitation and Barbados concentrations (Prospero and Lamb, 2003), or precipitation and visibility in the Sahel (e.g. Mbourou et al., 1997;Mahowald et al., 2007), so the other driver could be

20   precipitation. In some of the CAM4 model simulations, the IAV of the source strength does feature significant correlations with both LAI and precipitation, but in general those same cases feature even stronger correlations with surface winds (Table S6). Although the quality of the surface wind data precludes us from evaluating the reanalysis for surface winds, as discussed in McVicar et al. (2012), we can evaluate the precipitation, which is commonly done (e.g. Trenberth and Guillemot, 1998). When we do, we see that the reanalysis precipitation datasets are not capturing either interannual

25   variability in precipitation nor the slope in the precipitation, compared to the CMAP precipitation compilation (more details in Section 2.2) (Table 4). Since precipitation and winds, perhaps due to gustiness in moist convection (Engelstaedter and Washington, 2007) or due to surface roughness changes from vegetation (Cowie et al., 2013), are likely related to each other, an error in the IAV of precipitation may be indicative of an error in the IAV of winds.

Overall, the model simulations conducted here support the hypothesis of Ridley et al. (2014), although the quality of

30   the reanalysis data forcing our simulations does not allow us to be conclusive about our results.

3.3 Southern Hemisphere variability

Most of the available dust observations come from the Northern Hemisphere (Table 2). Here we consider three sets of data from the Southern Hemisphere, one of surface concentrations in Rio Gallegos (Argentina/Patagonia, 52S, 69W) (Zihan,

2016) and one of the deposition at Kerguelen Island (49S, 70E), which is likely to be influenced by both South American and South African dust sources (Heimburger et al., 2012; Figure 5). Finally, there is one AERONET station with data in the coarse mode, which is likely to be dominated by dust, since it is far from the coast and downwind of a desert (Tinga Tingana, 29S,140E) (Figure S9). Notice that there is very little data at the first two observational sites, which means these results need to be interpreted carefully, especially for IAV, since 1 and half years of data is measured (Figure 5). The surface concentration data at Rio Gallegos suggests an IAV variability of 0.08 (Table 5), which is at the lower edge, but in the same range as the observations in the Northern Hemisphere (values between 0.06 and 0.4; Figure 3a; Table S1). The models, however, tend to predict too large a variability at this site (between 0.2 and 2). It is unclear why the models overpredict the variability, but it may be due to issues with the actual location of the sources in the model compared to the observations, or the strength of the North-South gradient in concentrations in the models versus observations (e.g. Gaiero, 2007;Gasso et al., 2010). In addition, the way the tuning in Albani et al. (2014) was conducted will increase the interannual variability of the dust cycle in the CAM4 and CAM5 simulations in the Southern Hemisphere (Table 5). These model simulations had too small an emission value from some regions and too large from others, and thus the model source strengths were tuned in order to broadly match observations (Albani et al., 2014). In particular, in Argentina, the dust source strength had to be tuned up very strongly in order to obtain a climatological dust that matched observations. This means that there were very few grid boxes and time periods with active sources, increasing the temporal variability in the model. If we had instead changed the wind threshold in our model formulation in order to tune the source strength (e.g. Tegen and Miller, 1998), we presumably would not have increased our variability as much, highlighting the importance of details in the model formulation for model results. The CAM4-ERAI and CAM4-AMIP simulations especially overpredict variability. Note however, that some of the CAM4 models (CAM4-MERRA and CAM4-NCEP) have similar variability as the non-CAM4 models (GCHEM-MERRA and MATCH-NCEP), especially at Kerguelen and Tinga Tingana, suggesting that some of this variability may also be associated with the meteorological dataset. We will explore the ramifications for variability predictions in Section 3.4 and 3.5. Further downstream of the sources, the amount of variability in deposition observed at Kerguelen is 0.10, which is overpredicted in most of the models (Table 5), while at Tinga Tingana, the IAV variability is 0.42 in the observations, and tends to be smaller in the models.

At both Rio Gallegos and Tinga Tingana, the ratio of the IAV variability to the seasonal variability is 1.4 and 1.5, which are on the lower side of the observations in the Northern Hemisphere (Figure 3c and 3d; Table S1). The models are able to simulate the reduced fraction of variability due to the seasonal cycle at these sites (Table 5). Because of the limited data and length of data, we cannot be sure, but the observations presented here are consistent with a stronger role of interannual variability, compared to seasonal variability, in dust sources in the Southern Hemisphere than in the Northern Hemisphere, as simulated by the models. Additional long term data in the Southern Hemisphere would allow more testing of this model result.

3.4 Spatial analysis of model simulations of variability and correlations

Consistent with previous studies (e.g. Tegen and Miller, 1998;Mahowald et al., 2003; Mahowald et al., 2011), the largest variability in IAV (standard deviation over the mean, described in equation 1, Section 2.3) in modeled dust concentrations occurs not in the main dust source areas or outflow regions of the North Atlantic, but rather adjacent to these regions, where intermittent dust events occur (Figure 6a). For example, in the North Atlantic, north of the main transport pathways. Some of the highest IAV variability occurs over ocean regions, especially in the Southern Hemisphere. There is much more seasonal variability than IAV variability in most locations in the North Hemisphere (Figure 6b), with smaller ratios of seasonal to IAV variability over the southwest US, North Atlantic to Northern Europe, and across large parts of the Southern Hemisphere (Figure 6b). The deposition and AOD IAV variability have similar patterns to the concentration, with a slightly higher variability in deposition and slightly lower in AOD (Figure 6a vs. Figure S10a and S10c), and similar importance of the seasonal cycle (Figure 6b vs. Figure S10b and S10d). As discussed in Section 3.3, for the CAM4 and CAM5 model simulations, some of this enhanced southern hemisphere variability could be due to the tuning of the source areas, because the dust sources were not consistently active enough (Albani et al., 2014). If we consider only the models in which the Southern Hemisphere dust sources did not have to be increased (GCHEM (MERRA) and MATCH (NCEP)), then the monthly mean variability in the South Atlantic is similar to the North Atlantic (Figure 6c), but there still tends to be a smaller proportion of the variability from seasonal variability, and thus a more important role for interannual variability, in the South Atlantic, Indian or Pacific oceans than in the Northern Hemisphere oceans (Figure 6d). Some of the limited observational data supports strong interannual variability in the Southern Hemisphere, although not as strong as some of the model versions (Section 3.3).

Next we consider how similar the temporal variability is in the model simulations covering the same time period. If two model simulations are temporally correlated, it implies the timing of the monthly mean variability in the models is similar. Of course, to obtain the fraction of the variability that is similar, the correlation coefficient needs to be squared (if we assume a Gaussian distribution in the model output), which means that even a statistically significant (at 95%) high correlation of 0.8 (for 16 different years) only implies that 60% of the variability is similar. However, the correlation is a useful way to consider how similar the simulations are in their variability. The correlations between the model simulations for the surface concentration suggest that the models simulate similar IAV variability over some of the globe, but over most of the globe, there is no statistically significant correlation (Figure 7). The strongest correlations occur in the model simulations with reanalysis driven simulations (Figure 7a, b, d, e), and the simulations with time varying-SSTs (AMIP) were more different (Figure 7c and f). This suggests that sea surface temperature forcings are not the only important driver for the dust cycle. Notice that simulations with the same model but different winds (CAM4 (MERRA) vs. CAM4 (NCEP); Figure 7a) had similar correlation coefficients as using different winds and model framework (e.g. CAM4 (MERRA) vs. MATCH (NCEP); Figure 7e) or different models with the same winds (e.g. CAM4 (MERRA) vs. GCHEM (MERRA); Figure 7d). This is made more clear when we average the correlations across CAM4 reanalysis models and compare to simulations using

different model frameworks driven with the same meteorological data (Figure 8a vs. 8b), which show similar patterns of correlations This suggests that both model framework and winds contribute to variability, and perhaps in a similar, but not additive magnitude.

We can explore the drivers of interannual variability by focusing on the average correlation between the CAM4 simulations driven by reanalysis (CAM4-RE average) and the North Atlantic Oscillation (NAO) and the El Nino/Southern Oscillation (ENSO) climate indices (Figure 8c and 8d). These show correlations similar to previous studies (e.g. Moulin et al., 1997;Mahowald et al., 2003), and suggest some correlation between NAO and ENSO. Because we are using a relatively short time period (16 years), these signals do not show as statistically significantly as if we use a longer time period, but the same models (e.g. MATCH in Mahowald et al., 2003 included 22 years) have shown a similar pattern in previous studies. There is much more coherence in the simulated variability from the reanalysis winds, than seen in the NAO or ENSO (Figure 8a vs. Figure 8c and Figure 8d). This is also consistent with the lower correlation between the SST-driven model simulations (AMIP-style) (Figure 8e). There are much higher correlations between model results when we use reanalysis winds compared to forcing with only sea-surface temperatures, indicating the value of using reanalyses datasets to obtain more robust results. Notice that there is a much stronger correlation between models if we consider the seasonal cycle (Figure 8f), indicating the difficulty the models have with simulating IAV compared with the seasonal cycle. This is consistent with previous studies showing similar simulation of seasonal cycle across many models (e.g. Huneeus et al., 2011).

Surprisingly in some remote ocean regions, the models are simulating similar interannual variability (notably, parts of the Atlantic, Pacific and Indian Oceans), although this could be spurious due to the short time period considered (Figure 8). There are some statistically significant trends in the model simulations over the 1990-2005 time period, as seen in Barbados (Section 3.2) and the nearby North Atlantic and some parts of North Africa (Figure 4), which may be responsible for some of this coherence.   The patterns and magnitude of correlation coefficients is similar for deposition and AOD in the models (Figure S11).

A comparison of other aerosols in two of the CAM4-reanalysis based simulations available here (Figure S12), is consistent with the idea that dust is more highly variable. For other aerosol types, especially BC, OC and SO_4, which include in this study no IAV in the sources, there is some correlation between the models driven by different meteorology far from the sources as well as close (Figure S12). We will next discuss regional averages to understand how similarly ocean basin averages are simulated, to see if these IAV correlations in some regions are large enough to provide coherent basin estimates.

3.5 Modelled temporal trends in different regions

As discussed in the cases of the Tropical North Atlantic and South America (Sections 3.2 and 3.3), as well as in previous studies (Tegen and Miller, 1998), some of the variability in dust comes from the source regions. Looking across the model simulations considered here, we see strong IAV in many of the source regions, with the strongest IAV in the smallest

source regions, as seen previously (Tegen and Miller, 1998;Mahowald et al., 2003) (Figure 9). Only the western Sahel source (Section 3.2) is simulated to have a statistically significant trend in all the model simulations (Table 4, also seen in Figure 4 and 9). There are strong increases in some of the model simulations of the Australian source, consistent with the observed increase in drought over this time period (Cai et al., 2014), although we do not know of dust observations verifying this. For example, visibility data extending through 2003 does not support an increase in dust source strength in Australia (Mahowald et al., 2007). Previous studies have shown that there are decreases in dust from South America over 1990-2005 period in some models (Doney et al., 2009), but these are not shown in all the model versions or supported by robust observational data (Doney et al., 2009). Yu et al. (2016) show evidence for strong decadal variability in the Saudi Arabian source over a different time period than considered here, with an increase between 2000 and 2015, arguing that a drought in the Fertile Crescent is responsible for the increase after 2000. The model simulations included here show a lower dust source during 2000-2005 (Figure 9), consistent with their results, and correlation between precipitation and dust source in these simulations (Table 7). Considering the 1990-2005 period, on a global average, some models simulate an increase in the global dust source, other models simulate a decrease, suggesting no clear trends from the modeling of the global dust cycle over this time period (Figure 9).

If we consider regional averages, there are moderate correlations in time (0.4-0.8) in the annual mean source strength, except for the Sahel region (0.13), suggesting they simulate similar interannual variability (Table 6; Figure 9). Although the emphasis of this paper is on the temporal variability, there are significant differences in the climatological mean source strengths in the models used here (Table 6), highlighting the uncertainties in dust sources. The larger sources (such as North Africa and the Sahel) have mean source strengths varying by 25% while the smaller sources (such as North and South America) vary up to 160%, just due to differences in the meteorology, using the same CAM4 model and observational constraints (Albani et al., 2014).

Focusing on the drivers of the CAM4 modeled variation in sources suggests that LAI has the strongest correlation with IAV in sources for several source regions (Australia, South Africa and South America (Argentina), while surface winds have the highest correlations for East Asia, Middle East, North Africa and the western Sahel (Table 7). It is reassuring that the model winds have high correlations with IAV source strength in the model for East Asia, for there is a strong correlation in the current climate observations between winds and sources in this region (e.g. Sun et al., 2001), as well as speculation that past climate variability in winds from East Asia is sensitive to synoptic scale wind events (e.g. Roe, 2008;McGee et al., 2010). Soil moisture has the highest correlation for North America. Notice that IAV in LAI is likely to be a strong function of soil moisture, and precipitation (Table 8), and LAI may cause changes in surface roughness and therefore winds (Cowie et al., 2013), so these variables are all likely to be related. There are trends across this relatively short time period of a few of the source regions and variables, averaged across all the models (Table S7), but longer records tend to suggest oscillation in dust related variables: for example, the downward trend in dust in the Sahel from the 1980s followed a strong upward trend between the 1960s and the 1980s (Prospero and Lamb, 2003), and indeed there may have been even longer trends in dust from North Africa (Mulitza et al., 2010). Thus, trends in the short time period considered here (1982-2008) may not

necessarily be representative of the longer term trends.

The current generation of earth system models have a great deal of difficulty in simulating not only precipitation (as discussed in Section 3.2) but also LAI (e.g. Sitch et al., 2015;Mahowald et al., 2016), and thus this strong dependence on difficult-to-simulate variables may decrease our ability to simulate interannual variability. Note that using satellite retrieved

5   vegetation (e.g. Zhu et al., 2013), as done in some models (e.g. Ridley et al., 2014), may remove model-derived uncertainties, but the satellite retrieved vegetation has large uncertainties, especially in regions with low LAI, and does not do a good job of detecting brown vegetation, which is very important in resisting dust entrainment (e.g. Okin et al., 2001). This suggests that simulation of the surface conditions (e.g. vegetation, soil moisture, surface winds) in the source regions is likely to limit our ability to accurately simulate IAV in dust.

10   Downwind of the source regions, there is some coherence in the simulated variability in the surface concentrations as well (Table 8), especially in the North Atlantic, North Pacific, Equatorial Atlantic, South Pacific and South Indian Ocean (with correlation coefficients between different CAM4-RE simulations averaging between 0.46 to 075), but other regions there is less coherence. The sources which dominate the surface concentration, deposition and AOD in different downwind regions were not diagnosed in this study, but were previously shown for the mean in related model simulations (see Albani et

15   al., 2014; Figure S1), and suggest source-receptor type relationships consistent with previous studies (e.g. Tanaka and Chiba, 2006;Mahowald, 2007). Those studies suggest that, as expected, North African source dominant North Atlantic sources, and East Asian and Central Asian sources dominate the North Pacific. The Central Asian sources are important for the North Indian Ocean. The Southern Hemisphere sources tend to dominate the regions just downwind of the sources (Figure S1, Albani et al., 2014). These results are consistent with the available source provenance data, which were used to 'tune' the

20   CAM4 simulations (see Albani et al., 2014 for more details).

For the downwind regions, we can consider the importance of the climate indices in the time series (Figure 10; Table 8). If we examine the regional correlation coefficients of dust surface concentration to the NAO, there is very little correlation across this 16-year period for the North Atlantic (Table 8 and Figure 8c). Previously studies showed a larger contribution, which might be due to the short time series used here (Moulin et al., 1997) (Mahowald et al., 2003; Ginoux et

25   al., 2002). The largest magnitude correlations are seen in regions remote from the NAO (Antarctic, for example, with an anti-correlation of -0.42), which may be due to spurious correlations (Table 8). If we consider ENSO (Table 8), we see that the strongest correlations occur in the ocean basins where this oscillation dominates the physics (also seen in Figure 8d). The North Pacific had the highest correlation coefficient (0.62), followed by the North Atlantic (0.45) and Equatorial Pacific (0.42). These results, for the shorter term oscillation of ENSO in contrast to the decadal oscillations from NAO, are more

30   similar to previous results (e.g. Mahowald et al., 2003).

For some applications (e.g. IAV in iron fluxes on biogeochemistry or IAV in dust contributions to AOD), deposition and AOD are more important (Streets et al., 2009;Doney et al., 2009), and thus we briefly consider the time series of these model fields (Figures S13 and S14). The mean deposition from the different model simulations can vary widely (Table S9), as seen in the source strengths (Table 6) and previous studies (e.g. Huneeus et al., 2011). Downwind of the large source

regions (e.g. North Atlantic, Central North Atlantic, Equatorial Atlantic and Northern and Equatorial Indian Ocean) the standard deviations between CAM4-RE climatology mean are the lowest (7-25%), and they are largest in the remote ocean regions, like the Southern Ocean (100%) (Table S9). The standard deviation between models is slightly larger if all models are included (Table S9). Similarly, for the AOD in the ocean basins, the difference in model simulations is smallest close to the source regions (20%) and largest in the remote ocean regions (100%) (Table S10).

Here we emphasize the temporal variability, and the comparison of the CAM4-RE model simulations suggest in many basins there are consistent signals, with correlations in net deposition above 0.3 in most basins, with the exception of North Indian and South Atlantic (Table S9; Figures S13 and S14). Interestingly, AOD is consistent (r>0.3) in different basins, with North Central Pacific, North Atlantic Equatorial Pacific having the lowest correlations. This highlights the disconnect between AOD and deposition, which can make diagnosing deposition variability from AOD difficult, as noted previously (e.g. Mahowald et al., 2003). Note even in the basin with the most consistent simulations (0.68 in the North Atlantic), the variability in deposition simulated similarly in the models represents only 50% of the variability (if we assume for simplicity Gaussian distributions, which is not true of these values, suggesting even less of the variability is similarly simulated).

3.6 Implication of modelled variability for sampling

The large variability in dust implies that it may be difficult to constrain the dust cycle from limited observations. While we have satellite data, we can only use that data to constrain dust, when it is the dominant aerosol, which occurs only in limited regions just downwind of major sources like North Africa (Mahowald et al., 2007). Over much of the ocean, we only have individual daily averaged values from cruise data (Baker et al., 2006;Buck et al., 2006;Sholkovitz et al., 2012). Previous model studies have shown that over much of the ocean, modeled daily averaged dust concentrations will tend to underestimate the annual average, and only be within a factor of 10 to 2 of the true modeled average (Mahowald et al., 2008). Here we consider how many monthly means are required to obtain an estimate of the annual mean that is within one standard deviation of the true model annual mean value (Figure 11). Over most of the globe, the number of monthly mean observations is 8-12, or almost a full seasonal cycle, as expected. But in regions with large variability (Figure 6), longer time periods of more than 2 full years are required to obtain a mean and standard deviation including the true mean (Figure 11). Note that here we assume that there is no trend or significant change in the dust cycle. Characterizations of changes in dust show that there are interesting trends over the longer term, which requires longer records (Prospero and Lamb, 2003;Ridley et al., 2014).

4. Conclusions

Simulations of annual mean variability in 7 different model simulations are compared to better understand how robust the

variability is in models for the period 1990-2005. Although the emphasis of this paper was not on evaluation of specific models, the models were compared with in situ concentration (Prospero and Lamb, 2003;Marticorena et al., 2010) and AERONET AOD (Holben et al., 2000) observations. The models considered here were 4 versions of the CAM4, the GCHEM (MERRA), MATCH (NCEP) and CAM5-AMIP (Table 1) (Albani et al., 2014;Luo et al., 2003;Ridley et al., 2014).

5   Here we ignore the possible dust sources from land use and land cover change, which is hypothesized to represent 25% of dust sources currently (Ginoux et al., 2012).

The model simulations do roughly similarly well (or poorly) compared to observations when driven by reanalysis meteorology, but less well when driven by sea surface temperatures with meteorology being prognostically calculated, implying that using reanalyzed meteorology does improve dust simulations (Figure 3). The models' ability to simulate the

10   observations is strongest for the seasonal cycle, and the models are less able to simulate the interannual variability, similar to previous studies (Mahowald et al., 2003;Ginoux et al., 2004) (Figure 3). Surface concentration and deposition have similar distributions of variability, while AOD tends to have less variability (Figure 3). There is more variability, especially interannual variability, in parts of the Southern Hemisphere (Figure 6). Some of this was artificially (potentially) enhanced in the simulations considered here because of the way that the CAM4 and CAM5 were tuned in the Southern Hemisphere. But

15   the very limited observations at some stations in the Southern Hemisphere suggest there could a larger fraction of interannual variability than seasonal variability compared with the Northern Hemisphere. This should be tested with more long term stations in the Southern Hemisphere.

Model simulations of interannual variability is sensitive to both meteorology as well as model construction, and thus drawing firm conclusions about how best to capture observed variability based on only one model is likely to be difficult.

20   Our results that model construction as well as meteorology is important is consistent with general circulation model studies which suggest that modeling group tend to have models which behave similarly (Knutti et al., 2013). These studies also complement reanalysis-based studies of the energy and water cycle, showing that there remain issues with the reanalysis datasets (Treberth et al., 2011; Trenberth and Fasullo, 2013).

Here we considered the hypothesis from (Ridley et al., 2014) that Barbados surface concentrations were decreasing

25   over the period 1980-2008 due to decrease in winds in North Africa and thus a decreasing source, which follows previous studies in highlighting the importance of winds on short (Engelstaedter and Washington, 2007; Sun et al., 2001) and long time scales in some source regions (Roe, 2008;McGee et al., 2010). Consistent with that hypothesis, most of the model versions considered here can simulate a decreasing concentration at Barbados, and the models trace this back to a decrease in dust source in the western Sahel region of North Africa, linked to a decrease in surface winds. This is consistent with the

30   meteorological station data in this region, which show both a correlation between dust sources and wind, as well as a decrease in winds over this time period (Mahowald et al., 2007). Basic meteorological principles and previous studies (e.g. Engelstadter et al., 2007) suggest associations between winds and precipitation, and the reanalysis models still cannot do a good job simulating mean or variability in precipitation (Treberth et al., 2011; Trenberth and Fasullo, 2013), suggesting that it will be difficult to improve the dust source, transport and deposition variability without improvements in the reanalyses. In

addition, other studies have noted the relationship between surface roughness changes due to vegetation, driving wind changes, and thus changes in source strength (Cowie et al., 2013), which could not be tested here. Of course, the time period considered here is relatively short, so it is unclear whether other drivers might be more important on longer time scales (e.g. Prospero and Lamb, 2003;Mulitza et al., 2010;Mahowald et al., 2010).

5   Because of the strong variability in dust, model simulations suggest that observations need to be made for around 1 year in many regions, but in remote regions, especially in the Southern Hemisphere, observations need to be made for more than 2 years in order to sample the modeled variability and correctly capture the annual mean concentrations (Figure 11). Of course, this assumes there is no long-term variability. Long time records of dust concentrations represent some of our most important data records to characterize variability in desert dust (e.g. Prospero and Lamb, 2003), and more long term datasets

10 should be collected.

Acknowledgements

We acknowledge the support of NSF 0932946 and 1003509, DOE DE-SC0006735 and DE-SC0016362, as well as the assistance of Computational and Information Systems Laboratory of the National Center for Atmospheric Research, whose resources were used for these simulations. DAR and CLH acknowledge support from NASA (NN14AP38G). We thank the

15 PIs of the AERONET stations for access to the AERONET data (Table 2). CMAP Precipitation data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/

References

Albani, S., Mahowald, N. M., Perry, A. T., Scanza, R. A., Heavens, N. G., Zender, C. S., Maggi, V., Kok, J. F., and Otto-Bliesner, B. L.: Improved dust representation in the Community Atmosphere Model. J. Adv. Model. Earth Syst., 6,

20   541-570, doi: 10.1002/2013MS000279, 2014.

Arimoto, R., Ray, B. J., Duce, R. A., Hewitt, A. D., Boldi, R., and Hudson, A.: Concentrations, sources and fluxes of trace elements in the remote marine atmosphere of New Zealand, Journal of Geophysical Research, 93, 22,389-322,405, 1990.

Arimoto, R., Duce, R. A., Ray, B. J., Ellis Jr., W. G., Cullen, J. T., and Merrill, J. T.: Trace elements in the atmosphere over

25   the North Atlantic, Journal of Geophysical Research, 100, 1199-1213, 1995.

Arimoto, R., Ray, B. J., Lewis, N. F., and Tomza, U.: Mass-particle size distributions of atmospheric dust and the dry deposition of dust to the remote ocean, Journal of Geophysical Research, 102, 15867-15874, 1997.

Atkinson, J., Murray, B., Woodhouse, M., Whale, T., Baustian, K., Carslaw, K. S., Doobie, S., O'Sullivan, D., and Malkin, T.: The importance of feldspar for ice nucleation by mineral dust in mixed-phase clouds, Nature, 498, 355-358

30   doi:310.1038/nature12278, 2013.

Aumont, O., Bopp, L., and Schulz, M.: What does temporal variability in aeolian dust deposition contribute to sea-surface iron and chlorophyll distributions?, Geophysical Research Letters, 35, doi:10.1029/2007GL031131, 2008.

Baker, A., French, M., and Linge, K.: Trends in aerosol nutrient solubility along a west-east transect of the Saharan dust plume, Geophysical Research Letters, 33, doi:10.1029/2005GL024764, 2006.

5 Balkanski, Y., Schultz, M., Claquin, T., and Guibert, S.: Reevaluation of mineral aerosol radiative forcings suggests a better agreement with satellite and AERONET data, Atmospheric Chemistry and Physics, 7, 81-95, 2007.

Buck, C., Landing, W. M., Resing, J. A., and Lebon, G.: Aerosol iron and alumninum solubility in the northwest Pacific Ocean: Results from the 2002 IOC Cruise, Geochemistry, Geophysics and Geosyst., 7, doi:10.0129/2005GC000977, 2006.

10 Cai, W., Purich, A., Cowan, T., van Rensch, P., and Weller, E.: Did climate change-induced rainfall trends contribute to the Australian Millenium drought?, Journal of Climate, 27, 3145-3168, 2014.

Cowie, S., Knippertz, P., and Marsham, J.: Are the vegetation-related roughness changes the cause of the recent decrease in dust emission from the Sahel?, Geophysical Research Letters, 40, 1868-1872; doi:1810.1002/grl.50273, 2013.

Cziczo, D., Froyd, K., Hoose, C., Jensen, E., Diao, M., Zondio, M., Smith, J., Twohy, C. H., and Murphy, D. M.: Clarifying
15   the dominant sources and mechansism of cirrus cloud formation, Science express, doi:10.10126/science.1234145, 2013.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Holm, E., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M.,
20   McNally, A., Monge-Sanz, B., Morccrette, J.-J., Park, B.-K., Peubey, C., Rosnay, P., Tavolato, C., Thepaut, J.-N., and Vitarts, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterlly Journal of the Royal Meteorological Society, 137, 553-597, 2011.

DeMott, P., Sassen, K., Poellot, M., Baumgardner, D., Rogers, D., Brooks, S., Prenni, A., and Kreidenweis, S.: African dust aerosols as atmospheric ice nuclei, Geophysical Research Letters, 30, 1732, doi:1710.1029/2003GL017410, 2003.

25 Doney, S., Lima, I., Feeley, R., Glover, D., Lindsay, K., Mahowald, N., Moore, J. K., and Wanninkhof, R.: Mechanisms governing interannual variability in upper-ocean inorganic carbon system and air-sea CO2 fluxes: physical climate and atmospheric dust, Deep Sea Research Part II: Topical Studies in Oceanography, 56, 640-655, 2009.

Dufresne, J.-L., Gauier, C., Ricchiazzi, P., and Rouquart, Y.: Longwave Scattering Effects of Mineral Aerosols, American Meteorlogical Society, 59, 1959-1966, 2002.

30 Engelstaedter, S., Kohfeld, K. E., Tegen, I., and Harrison, S. P.: Controls of dust emissions by vegetation and topographic depressions: An evaluation using dust storm frequency data, Geophysical Research Letters, 30, 1294, doi:1210.1029/2002GL016471, 012003, 2003.

Engelstaedter, S., and Washington, R.: Temporal controls on global dust emissions: the role of surface gustiness, Geophysical Research Letters, 34, doi:10.1029/2007GL029971, 2007.

Evan, A., Dunion, J., Foley, J., Heidinger, A., and Velden, C.: New evidence for a relationship between North Atlantic cyclone activity and African dust outbreaks, Geophysical Research Letters, 33, doi:10.1029/2006GL026408, 2006.

Evan, A., Vimont, D., Bennartz, R., Kossin, J. P., and Heidinger, A.: The role of aerosols in the evolution of tropical Atlantic Ocean temperature anomalies, Science, doi:10.1126/science.1167404, 2009.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, Reason, C., and Rummukainen, M.: Chapter 9: Evaluation of climate models, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, 2013.

Gaiero, D.: Dust provenance in Antarctic ice during glacial periods: From where in southern South America?, Geophysical Research Letters, 34, doi:10.1029/2007GL03520, 2007.

Gasso, S., Stein, A., Marino, F., Castellano, E., Udisti, R., and Ceratto, J.: A combined observational and modeling approach to study modern dust transport from the Patagonia desert to East Antarctica, Atmospheric Chemistry and Physics, 10, 8287-8303; doi:8210.5194/acp-8210-8287-2010, 2010.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B. N., Dubovik, O., and Lin, S.-J.: Sources and distribution of dust aerosols with the GOCART model, Journal of Geophysical Research, 106, 20255-20273, 2001.

Ginoux, P., Prospero, J. M., Torres, O., and Chin, M.: Long-term simulation of global dust distribution with the GOCART model: correlation with the North Atlantic Oscillation, Environmental Modeling and Software, 19, 113-128, 2004.

Ginoux, P., Prospero, J., Gill, T. E., Hsu, N. C., and Zhao, M.: Global scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS deep blue aerosol products, Reviews of Geophysics, 50, , DOI:10.1029/2012RG000388, 2012.

Gleckler, P., Taylor, K. E., and Doutriaux, C.: Performance metrics for climate models, Journal of Geophysical Research, 113, D06104, doi:06110.01029/02007JD008972, 2008.

Gregg, W., Conkright, M., Ginoux, P., O'Reilly, J., and Casey, N.: Ocean primary productiona nd climate: Global decadal changes, Geophysical Research Letters, 30, 1809, doi:1810.1029/2003GL016889, 2003.

Heimburger, A., Losno, R., Triquet, S., Dulac, F., and Mahowald, N.: Direct measurement of atmospheric iron, cobalt and aluminum-derived dust deposition at Kerguelen Islands, Global Biogeochemical Cycles, 26, doi:10.1029/2012GB004301, 2012.

Hess, P., and Mahowald, N.: Interannual variability in hindcasts of atmospheric chemistry: the role of meteorology, Atmospheric Chemistry and Physics, 9, 5261-5280, 2008.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Cimel, J. P. B., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET-A Federated Instrument Network and Data Archeive for Aerosol Characterization, 2000.

Huneeus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Kinne, S., Prospero, J., Bauer, S., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Fillmore, D., Ghan, S., Ginoux, P., Grini, A., Horowitz, L., Koch, D., Krol, M. C., Landing, W., Liu, X., Mahowald, N., Miller, R., Morccrette, J.-J., Myhre, G., Penner, J., Perlwitz, J., Siter, P., Takemura, T., and Zender, C.: Global dust model intercomparison in AEROCOM, Atmospheric Chemistry and Physics, 11, 1-36, 2011.

Hurrell, J. W., M. M. Holland, S. Ghan, J. -F. Lamarque, D. Lawrence, W. H. Lipscomb, N. Mahowald, D. Marsh, P. Rasch, D. Bader, W. D. Collins, P. R. Gent, J. J. Hack, J. Kiehl, P. Kushner, W. G. Large, S. Marshall, S. Vavrus, and Vertenstein, M.: The Community Earth System Model: A Framework for Collaborative Research, Bulletin of the American Meteorological Society, 0.1175/BAMS-D-1112-00121, 2013.

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

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., .Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bulletin American Meteorological Society, 77, 437-471, 1996.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., Dool, H. v. e., Jenne, R., and Fiorino, M.: The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation, Bulletin American Meteorological Society, 82, 247-267, 2001.

Knippertz, P., and Todd, M.: Mineral dust aerosols over the Sahara: Meteorological controls on emission and transport and implications for modeling, Reviews of Geophysics, 50, doi:10.1029/2011RG000362, 2012.

Knutti, R., Masson, D., and Gettelman, A.: Climate model genealogy: Generation CMIP6 and how we got there, Geophysical Research Letters, 40, 1194-1199; doi:1110.1002/grl.50256, 2013.

Kohfeld, K., and Harrison, S.: Glacial-interglacial changes in dust deposition on the Chinese Loess Plateau, Quarternary Science Reviews, 22, 1859-1878, 2003.

Kok, J.: A scaling theory for the size distribution of emitted dust aerosols suggests climate models underestimate the size of the global dust cycle, Proceedings of the National Academy of Science USA, 108, 1016-1021, 2011.

Lamarque, J. F., Emmons, L., Hess, P., Kinnison, D., Tilmes, S., Vitt, F., Heald, C., Holland, E., Lauritzen, P., Neu, J., Orlando, J., Rasch, P., and Tyndall, G.: CAM-chem: description and evaluation of interactive amtopsheric chemistry in CESM, Geoscientific Model Development, 4, 2199-2278, 2011.

Liu, X., Easter, R., Ghan, S., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Eckman, A., Hess, P., Mahowald, N., Collins, W., Iacono, M., Bretherton, C., Flanner, M., and Mitchell, D.: Toward a minimal representation of aerosol direct and indirect ef fects: model description and evaluation, Geoscientific Model Development Discussions, 4, 3485-3598, 2011.

Luo, C., Mahowald, N., and Corral, J. d.: Sensitivity study of meteorological parameters on mineral aerosol mobilization, transport and idstribution, Journal of Geophysical Research, 108, 4447, 4410.1029/2003JD0003483, 2003.

Luo, C., Mahowald, N., and Jones, C.: Temporal variability of dust mobilization and concentration in source regions, Journal of Geophysical Research, 109, doi:10.1029/2004JD004861, 2004.

5   Ma, P.-L., Rasch, P., Wang, H., Zhang, K., Easter, R., Tilmes, S., Fast, J., Liu, X., Yoon, J.-H., and Lamarque, J. F.: The role of circulation features on black carbon transport into the Arctic in the Community Atmosphere Model version 5 (CAM5), Journal of Geophysical Research, 118, 4657-4669; doi:4610.1002/jgrd.50411, 2013.

Mahowald, N., Rasch, P. J., and Prinn, R. G.: Cumulus parameterizations in chemical transport models, Journal of Geophysical Research, 100, 26173-26189, 1995.

10   Mahowald, N.: Development of a 3-Dimensional Chemical Transport Model Based on Observed Winds and Use in Inverse Modeling of the Sources of CCl3F, PhD, Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, 1996.

Mahowald, N., Rasch, P., Eaton, B., Whittlestone, S., and Prinn, R.: Transport of 222radon to the remote troposphere using the Model of Atmospheric Transport and Chemistry and assimilated winds from ECMWF and the National Center

15   for Environmental Prediction/NCAR, Journal of Geophysical Research, 102, 28,139-128,151, 1997.

Mahowald, N., Zender, C., Luo, C., del Corral, J., Savoie, D., and Torres, O.: Understanding the 30-year Barbados desert dust record, Journal of Geophysical Research, 10.129/2002JD002097, 2002.

Mahowald, N., and Kiehl, L.: Mineral aerosol and cloud interactiocns, Geophysical Research Letters, 30, 10.109/2002GL016762, 2003.

20   Mahowald, N., Luo, C., Corral, J. d., and Zender, C.: Interannual variability in atmospheric mineral aerosols from a 22-year model simulation and observational data, Journal of Geophysical Research, 108, 4352, 4310.1029/2002JD002821, 2003.

Mahowald, N., D. Muhs, Levis, S., Rasch, P., Yoshioka, M., and Zender, C.: Change in atmospheric mineral aerosols in response to climate: last glacial period, pre-industrial, modern and doubled-carbon dioxide climates Journal of

25   Geophysical Research, 111, D10202, doi:10210.11029/12005JD006653, 2006.

Mahowald, N., Ballentine, J.-A., Feddema, J., and N. Ramankutty: Global trends in visibility: implications for dust sources, ACP, 3309-3339, 2007.

Mahowald, N., Jickells, T. D., and Alex R. Baker, P. A., Claudia R.Benitez-Nelson, Gilles Bergametti,Tami C. Bond, Ying Chen, David D. Cohen, Barak Herut, Nilgun Kubilay, Remi Losno, Chao Luo, Willy Maenhaut, Kenneth A.

30   McGee, Gregory S. Okin, Ronald L. Siefert, Seigen Tsukuda: The global distribution of atmospheric phosphorus deposition and anthropogenic impacts, Global Biogeochemical Cycles, 22, doi:10.1029/2008GB003240, 2008.

Mahowald, N., Kloster, S., Engelstaedter, S., Moore, J. K., Mukhopadhyay, S., McConnell, Albani, S., Doney, S., Bhattacharya, A., Curran, M., Flanner, M., Hoffman, F., Lawrence, D., Lindsay, K., Mayewski, P., Neff, J.,

Rothenberg, D., Thomas, E., Thornton, P., and Zender, C.: Observed 20th century desert dust variability: impact on climate and biogeochemistry, Atmospheric Chemistry and Physics, 10, 10875-10893, 2010.

Mahowald, N., Lo, F., Y., Z., Harrison, L., Funk, C., Lombardozzi, D., and Goodale, C.: Leaf Area Index in earth system models: evaluation and projections, Earth System Dynamics, 7, 211-229, 2016.

5   Marsham, J., Knippertz, P., Dixon, N., Parker, D., and Lister, G.: The importance of representation of deep convection for modeled dust-generated winds over West Africa during the summer, Geophysical Research Letters, 38, doi:10.1029/2001GL048368, 2011.

Marticorena, B., and Bergametti, G.: Modeling the atmospheric dust cycle: 1. Design of a soil-derived dust emission scheme, Journal of Geophysical Research, 100, 16,415-416,430, 1995.

10   Marticorena, B., and Bergametti, G.: Two-year simulations of seasonal and interannual changes of the Saharan dust emissions, Geophysical Research Letters, 23, 1921-1924, 1996.

Marticorena, B., Chatenet, B., Rajot, J., Traore, S., Diallo, A., Kone, I., Maman, A., NDiaye, T., and Zakou, A.: Temporal variability of mineral dust concentrations over West Africa: analyses of a pluriannual monitoring from the AMMA Sahelian Dust Transect, Atmos. Chem. Phys., 10, 8899-8915; doi:8810.5194/acp-8810-8899-2010, 2010.

15   Mbourou, G. N. T., Bertrand, J. J., and Nicholson, S. E.: The Diurnal and Seasonal Cycles of Wind-Borne Dust over Africa North of the Equator, Journal of Applied Meteorology, 36, 868-882, 1997.

McGee, D., Broecker, W., and Winckler, G.: Gustiness: the driver of glacial dustiness?, Quarternary Science Reviews, doi:10.1026/jquascirev.2010.1006.1009, 2010.

McVicar, T., Roderick, M., Donohue, R., Li, L., Van Niel, T., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald,
20   N., Mescherskaya, A., Kruger, A., Rehman, S., and Dinpashoh, Y.: Global review and synthesis of trends in observed terrestrial near-surface wind speeds: implications for evaporation, Journal of Hydrology, 416-7, 182-205, 2012.

Miller, R., and Tegen, I.: Climate Response to Soil Dust Aerosols, American Meteorlogical Society, 11, 3247-3267, 1998a.

Miller, R. L., and Tegen, I.: Climate Response to Soil Dust Aerosols, Journal of Climate, 11, 3247-3267, 1998b.

25   Moulin, C., Lambert, C. E., Dulac, F., and Dayan, U.: Control of atmospheric export of dust from North Africa by the North Atlantic Oscillation, Nature, 387, 691-694, 1997.Moulin, C., and Chiapello, I.: Evidence of the control of summer atmospheric transport of African dust over the Atlantic by Sahel sources from TMS satellites (1979-2000), Geophysical Research Letters, 31, doi:10.1029/2003GL018931, 2004.

Mulitza, S., Heslop, D., Pittauerova, D., Fischer, H., Meyer, I., Stuut, J.-B., Zabel, M., Mollenhauer, G., Collins, J., Kuhnert,
30   H., and Schulz, M.: Increase in African dust flux at the onset of commercial agriculture in the Sahel region., Nature, 466, doi:10.1038/nature09213, 09226-09228, 2010.

Myhre, G., Shindel, D., Breon, G.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J. F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Nature Radiative Forcing, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth

Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T., Qin, D., Plattner, G. K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., Cambridge University Press, Cambridge, UK, 659-740, 2013.

Neale, R., Richter, J., Park, S., Lauritzen, P., Vavrus, S., Rasch, P., and Zhang, M.: The mean climate of the Community Atmosphere Model (CAM4) in forced SST and fully copled experiments, Journal of Climate, 26, 5150-5168; doi:5110.1175/jcli-d-5112-00236.00231., 2013.

Okin, G., Mahowald, N., Chadwick, O., and Artaxo, P.: The impact of desert dust on the biogeochemistry of phosphorus in terrestrial ecosystems, Global Biogeochemical Cycles, 18, GB2005, doi:2010.1029/2003GB002145, 2004.

Okin, G. S., Murray, B., and Schlesinger, W. H.: Degradation of sandy arid shrubland environments: observations, process modelling, and management implications, Journal of Arid Environments, 47, 123-144, 2001.

Painter, T., Barrett, A., Landry, C. C., Neff, J. C., Cassidy, M. P., Lawrence, C. R., McBride, K. E., and Farmer, G. L.: Impact of disturbed desert soils on duration of mountain snow cover, Geophysical Research Letters, 34, doi:10.1029/2007GL030284, 2007.

Petit, J. R., Jouzel, J., Raynaud, D., and al., e.: Climate and atmospheric history of the past 420,000 years from the Vostok Ice core, Antarctica, Nature, 399, 429-436, 1999.

Prospero, J.: Long term records of dust transport over the North Atlantic Ocean: an overview, Fall American Geophysical Union, San Francisco, 2006.

Prospero, J., and Lamb, P.: African droughts and dust transport to the Caribbean: climate change implications, Science, 302, 1024-1027, 2003.

Prospero, J. M., and Nees, R. T.: Impact of the North African drought and El Niño on mineral dust in the Barbados trade winds, Nature, 320, 735-738, 1986.

Prospero, J. M., and Savoie, D. L.: Effect of continental sources on nitrate concentrations over the Pacific Ocean, Nature, 339, 687-689, 1989.

Prospero, J. M.: Saharan Dust Transport over the North Atlantic Ocean and Mediterranean: An Overview, in: The Impact of Desert Dust Across the Mediterranean, edited by: Guerzoni, S., and Chester, R., Kluwer Academic Publishers, Netherlands, 133-151, 1996.

Prospero, J. M., Barrett, K., Church, T., Dentener, F., Duce, R. A., Galloway, J. N., II, H. L., Moody, J., and Quinn, P.: Atmospheric deposition of nutrients to the North Atlantic Basin, Biogeochemistry, 35, 27-73, 1996.

Prospero, J. M.: Long-range transport of mineral dust in the global atmosphere: Impact of African dust on the environment of the southeastern United States, Proc. Natl. Academy Science, 96, 3396-3403, 1999.

Rasch, P. J., Mahowald, N. M., and Eaton, B. E.: Representations of transport, convection and the hydrologic cycle in chemical transport models: Implications for the modeling of short-lived and soluble species, J Geophysical Res, 102, 18127-28138, 1997.

Reddy, M., Boucher, O., Belloiun, N., Schulz, M., Balkanski, Y., Dufresne, J., and Pham, M.: Estimates of global multicomponent aerosol optical depth and direct radiative perturbation in the Laboratoire de Meteorologie Dynamique general circulation model, J. Geophysical Research, 110, D10S16; doi:10.1029/2004JD004757, 2005.

Renno, N., Nash, A. A., Lunine, J., and Murphy, J.: Martian and terrestrial dust devils: test of a scaling theory using Pathfinder data, Journal of Geophysical Research, 105, 1859-1865, 2000.

Ridley, D., Heald, C., Pierce, J., and Evans, M. J.: Toward resolution-independent dust emissions in global models: impacts on the seasonal and spatial distribution of dust, Geophysical Research Letters, 40, 3873-2877; doi:3810.1002/grl.50409, 2013.

Ridley, D., Heald, C., and Prospero, J.: What controls the recent changes in African mineral dust aerosols across the Atlantic?, Atmospheric Chemistry and Physics, 14, 5735-5747; doi:5710.5194/acp-5714-5735-2014, 2014.

Ridley, D., Heald, C., Kok, J., and Zhao, C.: An observationally-constrained estimate of global dust aerosol optical depth, Atmos. Chem. Phys. Discussions, doi: 10.5194/acp-2016-5385, 2016.

Rienecker, M., Suarez, M., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M., Schubert, S., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C., Reichle, R., Robertson, F., Ruddick, A., Sienkiewicz, M., and Woolen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, Journal of Climate, 24, 3624-3648; doi: 3610.1175/JCLI-D-3611-00015.00011, 2011.

Roe, G.: On the interpretation of Chinese loess as a paleoclimate indicator, Quaternary Research, 71, 150-161, 2008.

Rosenfeld, D., Rudich, Y., and Lahav, R.: Desert dust suppressing precipitation: a possible desertification feedback loop, Proceedings of the National Academy of sciences of the United States of America, 98, 5975-5980, 2001.

Scanza, R., Mahowald, N., Ghan, S., Zender, C., Kok, J., Liu, X., and Zhang, Y.: Dependence of dust radiative forcing on mienralogy in the Community Atmosphere Model, Atmospheric Chemistry and Physics, 15, 537-561, 2015.

Seinfeld, J., and Pandis, S.: Atmospheric Chemistry and Physics, John Wiley and Sons, Inc, New York, 1326 pp., 1998.

Shao, Y., Wyrwoll, K.-H., Chappel, A., Huang, J., Lin, Z., McTainsh, G., Mikami, M., Tanaka, T., Wang, X., and Yoon, S.: Dust cycle: an emerging core theme in earth system science, Aeolian Research, 2, 181-204, 2011.

Sholkovitz, E., Sedwick, P., Church, T., Baker, A., and Powell, C.: Fractional solubility of aerosol iron: synthesis of a global-scale data set, Geochimica et Cosmochimica Acta, 89, 173-189, 2012.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S., Murray-Tortarolo, G., Ahlstrom, A., Doney, S., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Piao, S., LeQuere, C., Smith, B., Zhu, Z., and Myneni, R.: Recent trends and drivers of regional sources and sinks of carbon dioxide, Biogeosciences, 12, 653-679; doi: 610.5194/bg-5112-5653-2015, 2015.

Sokolik, I. N., and Toon, O. B.: Direct radiative forcing by anthropogenic airborne mineral aerosols, Nature, 381, 681-683, 1996.

Streets, D., Yan, F., Chin, M., Diehl, T., Mahowald, N., Schultz, M., Wild, M., Wu, Y., and Yu, C.: Anthropogenic and natural contributions to regional trends in aerosol optical depth, 1980-2006, Journal of Geophysical Research, 114, doi:10.2029/2008JD011624, 2009.

Swap, R., Garstang, M., Greco, S., Talbot, R., and Kallberg, P.: Saharan dust in the Amazon Basin, Tellus, 44B, 133-149, 1992.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research, 106, 7183-7192, 2001.

Tegen, I., and Lacis, A.: Modeling of particle size distribution and its influence on the radiative properties of mineral dust aerosol, Journal of Geophysical Research, 101, 19,237-219,244, 1996.

Tegen, I., and Miller, R.: A general circulation model study on the interannual variability of soil dust aerosol, Journal of Geophysical Research, 103, 25,975-925,995, 1998.

Tegen, I., Harrison, S. P., Kohfeld, K., Prentice, C., Coe, M., and Heimann, M.: The impact of vegetation and preferential source areas on global dust aerosol: Results from a model study, Journal of Geophysical Research, 107, 4576, doi:4510.1029/2001JD000963, 2002.

Torres, O., Bhartia, P., Herman, J., Sinyuk, A., Ginoux, P., and Holben, B.: A long-term record of aerosol optical depth from TOMS observations andn comparison to AERONET measurements, J. ATmospheric Sciences, 59, 398-431, 2002.

Treberth, K., Fasullo, J., and Mackaro, J.: Atmospheric moisture transports from ocean to land and global energy flows in reanalyses, Journal of Climate, 24, 4907-4924; doi: 4910.1175/2011JCL14171.14171, 2011.

Trenberth, K., and Fasullo, J.: Regional energy and water cycles: Transports from ocean to land, Journal of Climate, 26, 7827-7851, 2013.

Trenberth, K. E., and Guillemot, C. J.: Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses, Climate Dynamics, 14, 213-231, 1998.

Trenberth, K. E., Stepaniak, D. P., and Caron, J. M.: The Global monsson as seen through the divergent atmospheric circulation, Journal of Climate, 3969-3993, 2000.

Wang, M., Ghan, S., Ovchinnikov, M., Liu, X., Easter, R., Kasianov, E., Qian, Y., and Morrison, H.: Aerosol indirect effects in a multi-scale aerosol-cliamte model PNNL-MMF, Atmospheric Chemistry and Physics, 11, 5431-5455; doi:5410.5194/acp-5411-5431-2011, 2011.

Wilks, D.: Statistical Methods in the Atmospheric Sciences, International Geophysics Series, Academic Press, Burlington, MA, 627 pp., 2006.

Xie, P., and Arkin, P.: Global precipitation: A 17-year monthy anslysis based on gauge observations, satellite estimates and numerical model outputs, Bulletin of the American Meteorological Society, 78, 2539-2558, 1997.

Yoshioka, M., Mahowald, N., Conley, A., Collins, W., Fillmore, D., and Coleman, D.: Impact of desert dust radiative forcing on Sahel precipitation: relative importance of dust compared to sea surface temperature variations, vegetation changes and greenhouse gas warming, Journal of Climate, 20, 1445-1467, 2007.

Yu, Y., Notaro, M., Liu, Z., Wang, F., Alkolibi, F., Fadda, E., and Bakhrjy, F.: Climatic controls on the interannual to decadal variaiblity inthe Saudi Arabian dust activity: Toward the development of a seasonal dust prediction model, Journal of Geophysical Research, 120, 1739-1758, doi:1710.1002/2014JD022611, 2015.

Zender, C., Bian, H., and Newman, D.: Mineral Dust Entrainment and Deposition (DEAD) model: Description and 1990s dust climatology, Journal of Geophysical Research, 108, 4416, doi:4410.1029/2002JD002775, 2003a.

Zender, C., Newman, D., and Torres, O.: Spatial Heterogeneity in Aerolian Erodibility: Uniform, Topographic, Geomorphic and Hydrologic Hypotheses., Journal of Geophysical Research, 108, 4543, doi:4510.1029/2002JD003039, 2003b.

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R., and Myneni, R.: Global data sets of vegetation leaf area index (LAI)3g and fraction of photosynthetically active radiation (FPAR)3g derived from global inventory modeling and mapping studies (GIMMS) Normalized difference vegetation index (NDVI3g) for the period 1981 to 2011, Remote Sensing, 5, 927-948, 2013.

Zihan, Q.: Chemical properties of continental aerosol transported over the Southern Ocean: Patagonian and Namibian sources, PhD, IPGP, Université Pierre et Marie Curie (UPMC), Paris, France, 215 pp., 2016.

Table 1: Description of model simulations considered here.

Case name	Model base	Meteorology	Retuned sources	Years available	Citation
CAM4 (MERRA)	CAM4	MERRA	Y	1980-2008	(Albani et al., 2014)
CAM4 (NCEP)	CAM4	NCEP	Y	1989-2006	(Albani et al., 2014)
CAM4 (ERAI)	CAM4	ERA-Interim	Y	1989-2008	(Albani et al., 2014)
CAM4 (AMIP)	CAM4	Online/AMIP	Y	1980-2006	(Albani et al., 2014)
GCHEM (MERRA)	GEOS-CHEM	MERRA	N	1982-2008	(Ridley et al., 2014)
MATCH (NCEP)	MATCH	NCEP	N	1982-2008	(Luo et al., 2003)
CAM5 (AMIP)	CAM5	Online/AMIP	Y	1990-2008	(Albani et al., 2014)

Table 2: Description of observational sites included here. Locations are plotted in the maps (Figure S1, S3 and S5).

Site	Latitude °N	Longitude °E	Years	Citation or PI for Aeronet
Surface Concentration				
Banizoumbou	14	3	2006-2013	(Marticorena et al., 2010)
Barbados	13	-59	1979-2008	(Prospero and Lamb, 2003)
Bermuda	32	-65	1989-1997	(Arimoto et al., 1995)
Cinzana	13	-6	2006-2013	(Marticorena et al., 2010)
Izana	28	-16	1989-1998	(Arimoto et al., 1995)
Macehead	53	-10	1989-1994	(Arimoto et al., 1995)
Mbour	14	-17	2006-2013	(Marticorena et al., 2010)
Miami	26	-80	1974-1999	(Prospero, 1999)
Midway	28	-177	1982-2000	(Prospero and Savoie, 1989)
AOD				
Bahrain	26	50	1998-2006	B. Holben
Dalanzadgad	43	104	1997-2012	B. Holben
Ilorin	8	4	1998-2009	R. Pinker
Sede Boker	30	34	1998-2010	A. Karnieli
Southern Hemisphere observations				
Rio Gallegos Surface concentrations	-52	-69	2011-2014	(Zihan, 2016)
Kerguelen deposition	-49	70	2008-2010	(Heimburger et al., 2012)
Tinga Tingana AOD	-29	140	2002-2012	R. Mitchell

Table 3: Annual average spatial comparison to observations for different cases (described in Table 1 and Methods). Correlations which are statistical significant at the 95 percentile are in bold.

Case	Surface Concentration Correlation (log space)	Deposition (log space)	AOD
CAM4 (MERRA)	**0.73 (0.89)**	**0.94 (0.84)**	**0.73**
CAM4 (NCEP)	**0.67 (0.86)**	**0.79 (0.84)**	**0.67**
CAM4 (ERAI)	**0.48 (0.81)**	**0.58 (0.76)**	**0.87**
CAM4 (AMIP)	**0.79 (0.78)**	**0.73 (0.84)**	**0.41**
GCHEM (MERRA)	**0.73 (0.90)**	**0.63 (0.84)**	**0.43**
MATCH (NCEP)	**0.83 (0.84)**	**0.43 (0.81)**	0.32
CAM5 (AMIP)	**0.79 (0.89)**	**0.59 (0.85)**	**0.70**

Table 4: Slope of the normalized annual mean values from 1982 to 2008 (or time period available, shown in Table 1) for the Western Sahel (13 to 22°N and -20 to 13°E) and North Africa (0 to 35°N and -20 to 40°E) and model (Figure 4) (statistically significant values are in bold, standard deviation of slope in parenthesis for Barbados surf. conc. slope) in the first four columns. Values are normalized by dividing by the mean, so that slopes represent relative change per year. The last column is the correlation of interannual variability in precipitation in each model compared to observations.

	Slope Barbados surf. conc.	Slope Sahel source	Slope North African source	Slope Sahel Precip.	Correl. With obs. Sahel precip.
CAM4 (MERRA)	**-0.014 (0.0054)**	**-0.0065**	-0.0005	0.0035	**0.43**
CAM4 (NCEP)	-0.0017 (0.016)	**-0.0169**	**-0.0074**	**0.0425**	0.29
CAM4 (ERAI)	**-0.0058 (0.0051)**	-0.0006	0.002	0.0008	**0.81**
CAM4 (AMIP)	**-0.0079 (0.0035)**	**-0.0061**	**-0.0037**	**0.0186**	**0.42**
GCHEM (MERRA)	**-0.025 (0.0047)**	**-0.021**	**-0.0072**	0.0035	**0.43**
MATCH (NCEP)	**-0.0087 (0.0059)**	**-0.0047**	0.027	**0.0425**	0.29
CAM5 (AMIP)	-0.01 (0.01)	**0.0027**	0.029		
Obs	**-0.016 (0.006)**			**0.0089**	

Table 5: Variability in Southern Hemisphere
Values for the IAV variability (annual average standard deviation divided by mean) and the ratio of the variability from the seasonal cycle over the IAV (for the surface concentration in the model cases and data from Rio Gallegos; deposition data from Kerguelen; and coarse mode AOD from Tinga Tingana (locations listed in Table 1).

Model/ Observations	Rio Gallegos Surface concentrations		Kerguelen deposition		Tinga Tingana AOD	
	IAV variability	Ratio variability from seasonal cycle over IAV	IAV variability	Ratio variability from seasonal cycle over IAV	IAV variability	Ratio variability from seasonal cycle over IAV
CAM4 (MERRA)	1.78	1.06	0.22	1.21	0.11	3.61
CAM4 (NCEP)	2.13	0.79	0.27	1.02	0.23	2.21
CAM4 (ERAI)	2.66	0.84	1.42	1.32	0.21	1.81
CAM4 (AMIP)	3.86	0.55	2.06	0.48	0.17	2.68
GCHEM (MERRA)	0.67	1.26	0.32	1.03	0.25	1.87
MATCH (NCEP)	0.68	1.37	0.17	2.16	0.21	2.55
CAM5 (AMIP)	0.42	3.42	0.46	2.37	0.14	3.25
Obs	0.10	1.39	0.08	4.01	0.42	1.48

Table 6: Regional sources of dust. For each source region, the averaged correlation across time between annual mean source strengths for the CAM-RE cases is shown in the second column. The following columns show the climatological mean source strength (Tg/year) for the mean of the 3 CAM4-RE simulations and the mean of the 7 simulations included in this study. The +/- % standard deviation is also shown, and represents the standard deviations across the models included in the averaging. The regions are defined as follows: Australia: 35 to 25°S, 130 to 150°E; East Asia: 35 to 50°N, 70 to 112°E; Middle East: 10 to 45°N, 40 to 70°E; North Africa: 10 to 35°N, 40°W to 40°E; Sahel (western): 13 to 22°N, 40°W to 13°E; South Africa: 35 to 20°S, 15 to 40°E; South America (Argentina): 55 to 35°S, 285 to 310°E.

	Avg. IAV Temporal Correlation	Mean CAM4-RE	Mean all
Australia	0.73	25 +/-70%	38 +/- 90%
East Asia	0.58	230+/- 70%	230 +/-70%
Middle East	0.40	570+/-40%	510 +/- 40%
North Africa	0.47	1370 +/-23%	1490 +/- 40%
North America	0.78	72 +/-160%	70 +/-130%
Sahel (western)	0.13	460 +/-27%	520 +/- 40%
South Africa	0.46	9 +/- 90%	8 +/- 70%
South America	0.68	14 +/-160%	34 +/- 130%
Globe	0.49	2400 +/-26%	2500 +/- 40%

Table 7: Correlations in meteorological variables and mobilization in different regions for IAV. Time series are correlated for the annual average over 1990-2005 in each region (only including gridboxes which are active at any time in that model simulation). Values shown are the averages of the correlations across the CAM4-Reanalysis models (CAM4 (MERRA), CAM4 (NCEP) and CAM4 (ERAI)). The regions are defined as in Table 6.

	Precipitation .	Soil moisture	Leaf Area Index	Sfc. Wind
Australia	-0.59	-0.61	-0.72	0.10
East Asia	0.06	-0.08	-0.32	0.67
Middle East	-0.32	-0.33	-0.28	0.36
North Africa	-0.27	-0.26	-0.20	0.51
North America	-0.57	-0.64	-0.53	0.32
Sahel (western)	-0.28	-0.26	-0.42	0.81
South Africa	-0.37	-0.38	-0.55	0.22
Globe	-0.36	-0.29	-0.46	0.33

Table 8: Surface concentration over ocean basins. For each ocean region, the averaged correlation across time between annual mean deposition fluxes for the CAM-RE cases is shown in the second column. The third column shows the annual mean correlation with NAO, while the third column shows the annual mean correlation with the El Nino/Southern Oscillation climate
5 **index. Regions are defined as the ocean gridboxes (not including sea ice or land boxes) in the following latitude and longitude areas as from (Gregg et al., 2003): North Atlantic (>30°N; 270 to 30°E); North Pacific (>30°N; 120 to 270°E); North Central Atlantic (10 to 30°N, 270 to 30°E); North Central Pacific (10 to 30N; 120 to 270°E); North Indian (10 to 30°N; 30 to 120°E); Equatorial Atlantic (-10 to 10°N; 300 to 30°E); Equatorial Pacific (-10 to 10°N; 120 to 285°E);**
10 **Equatorial Indian (-10 to 10°N; 30-120°E); South Atlantic (-30 to -10°N; 30 to 300°E); South Pacific (-30 to -10°N; 120 to 295°E); South Indian (-30 to -30°N, 30 to 120°E); Antarctic (<-30°N).**

	CAM4-RE across model Correlation	NAO correlation	El Nino Correlation
North Atlantic	0.66	0.10	0.45
North Pacific	0.51	0.19	0.62
North Central Atlantic	0.75	0.04	-0.10
North Central Pacific	0.46	-0.19	0.01
North Indian	0.30	0.13	0.38
Equatorial Atlantic	0.59	-0.02	-0.31
Equatorial Pacific	0.19	-0.12	0.42
Equatorial Indian	0.31	-0.18	-0.15
South Atlantic	0.11	-0.22	-0.42
South Pacific	0.65	0.03	0.03
South Indian	0.46	0.29	0.16
Antarctic	0.28	-0.42	-0.63
Global	0.42	0.01	-0.03
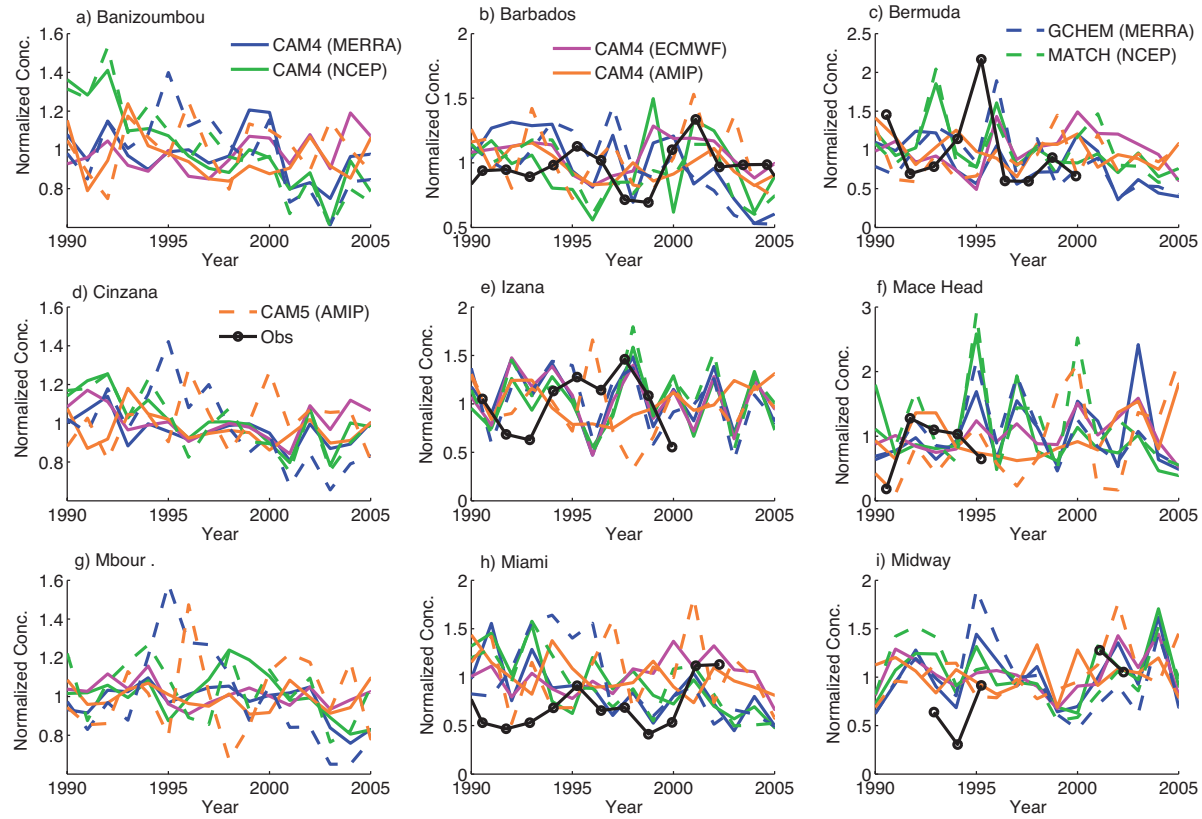
Figure 1: Time series of annual mean dust concentration divided by the long term mean for each time series (unitless) at stations in Table 2: a) Banizoumbou, b) Barbados, c) Bermuda, d) Cinzana, e) Izana, f) Macehead, g) Mbour, h) Miami, i) Midway. Observations are in black: if no observations shown, observations are from a different time period, and only used for variability and seasonal cycle calculations. Different colors and line styles indicate the different model versions: CAM4 (MERRA) (blue solid), CAM4 (NCEP) (green solid), CAM4 (ERAI) (pink solid), CAM4 (AMIP) (orange solid), GCHEM (MERRA) (blue dashed), MATCH (NCEP) (green dashed), CAM5 (AMIP) (orange dashed). The annual means are divided by the long term mean to allow comparison with seasonal variability, since they are similarly normalized (Figure S7).
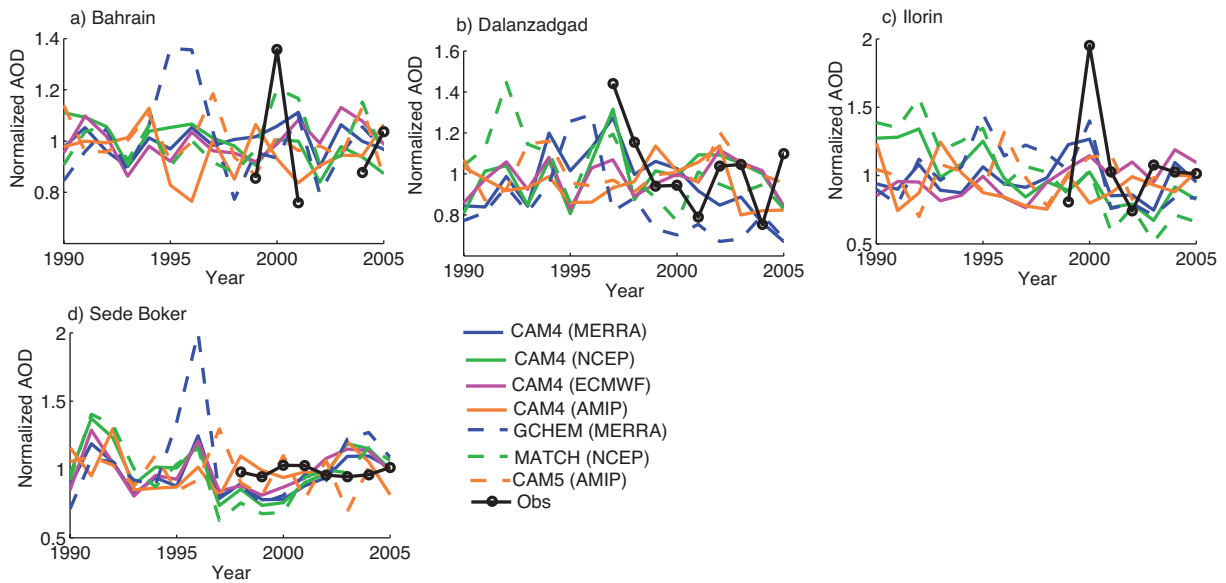
Figure 2: Timeseries of annual average AOD for model simulations (based on dust only), compared with AERONET observations in Table 2 for: a) Bahrain b) Dalanzadgad, c) Ilorin and d) Sede Boker for each of the different model versions (Colors are the same as in Figure 1).
5   **Observational data from AERONET stations (citations listed in Table 2). The annual means are divided by the long term mean to allow comparison with seasonal variability, since they are similarly normalized (Figure S8).**
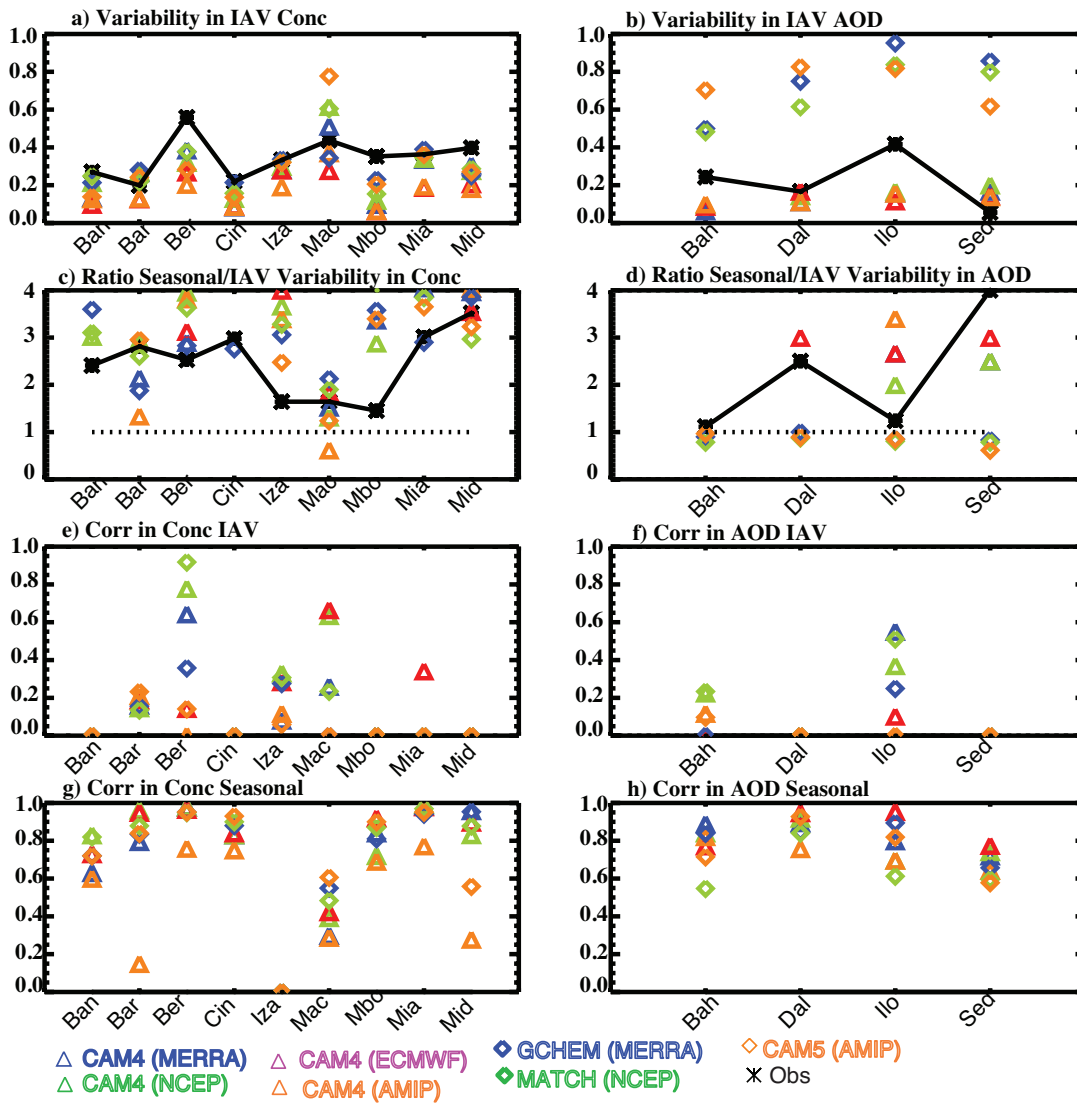
Figure 3: Model and observed IAV variability for (a) the concentration and (b) AOD from the observation sites listed in Table 2 and shown in Figures 1-2. IAV Variability is defined as the standard deviation over the mean using annual means. The ratio of the IAV variability to the seasonal cycle variability (12-month climatology) is shown for (c) concentration and (d) AOD. Model results are in color, while observations are in black. Correlation coefficients for IAV annual mean time series for (e) concentration and (f) AOD between the models and observations at the stations. The correlation coefficients between model and observed values for the seasonal cycle (g and h) for concentration and AOD. Observations are described in Table 2. Concentration stations are abbreviated: Ban: Banizoumbou; Bar: Barbados; Ber: Bermuda; Cin: Cinzana; Iza: Izana; Mac: Macehead; Mbo: Mbour; Mia: Miami; Mid: Midway. AOD stations abbreviated: Bah: Bahrain; Dal: Dalanzadgad; Ilo: Ilorin and Sed: Sede Boker.
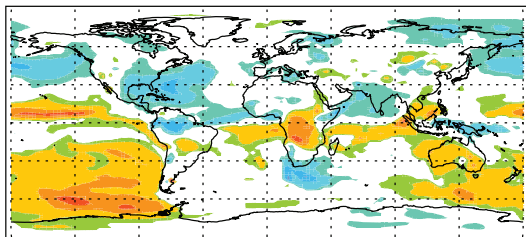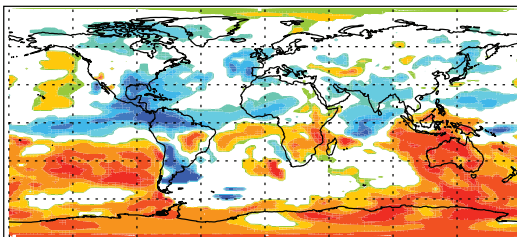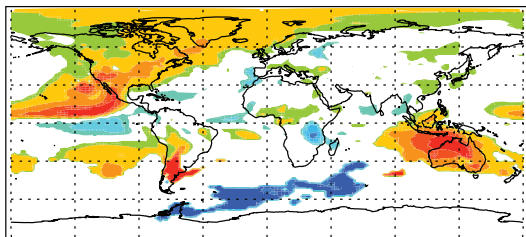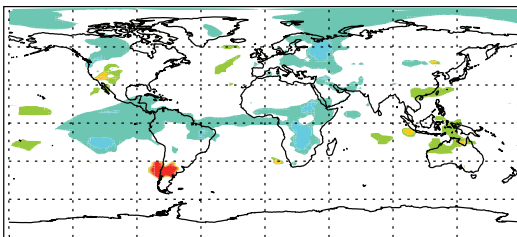
Figure 4: **Slope of the trend in the relative concentrations (linear regression of concentration onto time) of the modelled annual mean surface concentrations (normalized by the mean) in units of fraction change per decade for a) CAM4 (MERRA), b. CAM4 (NCEP), c. CAM4 (ERAI), d. CAM4 (AMIP), e. GCHEM (MERRA), f. MATCH (NCEP), g. CAM5 (AMIP) and the mean slope across the models (h). Only slopes that are larger in magnitude than one sigma from the regression are plotted. Positive slopes imply increasing concentrations.**
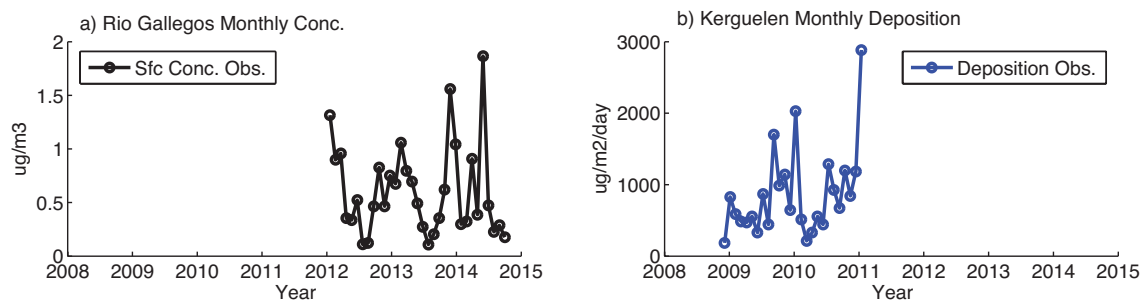
Figure 5: Monthly mean surface concentration observations at the Rio Gallegos site in Argentina (a) in μg/m³ (Zihan, 2016) and monthly mean deposition fluxes in μg/m²/day from Kerguelen (Heimburger et al., 2012) (b).
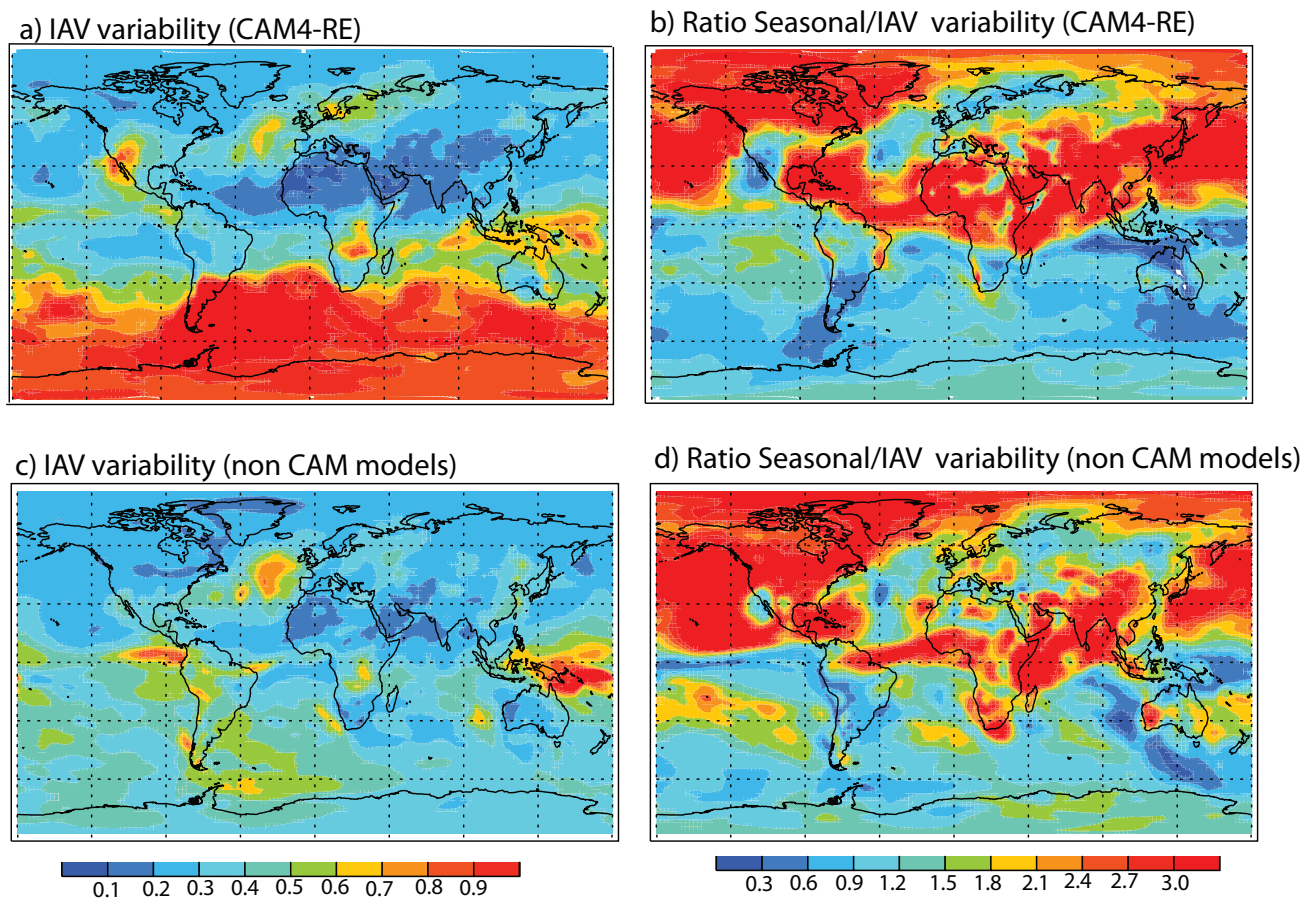
Figure 6: Spatial plot of the modelled IAV variability in the model simulations at each grid box (CAM4-RE is the mean of CAM4 (MERRA), CAM4 (NCEP) and CAM4 (ERAI)), where variability is unitless and is the standard deviation divided by the mean of the annual mean between 1990 and 2005 for a) surface concentration. The ratio of the seasonal variability over the IAV variability (calculated using the 12 climatological monthly means) is shown in the right hand panel for b) surface concentration. Similar diagnostics for the non-CAM models (GCHEM(MERRA) and MATCH (NCEP)) are shown in the bottom panel for a) IAV in variability and b) ratio of seasonal over IAV variability for concentration.
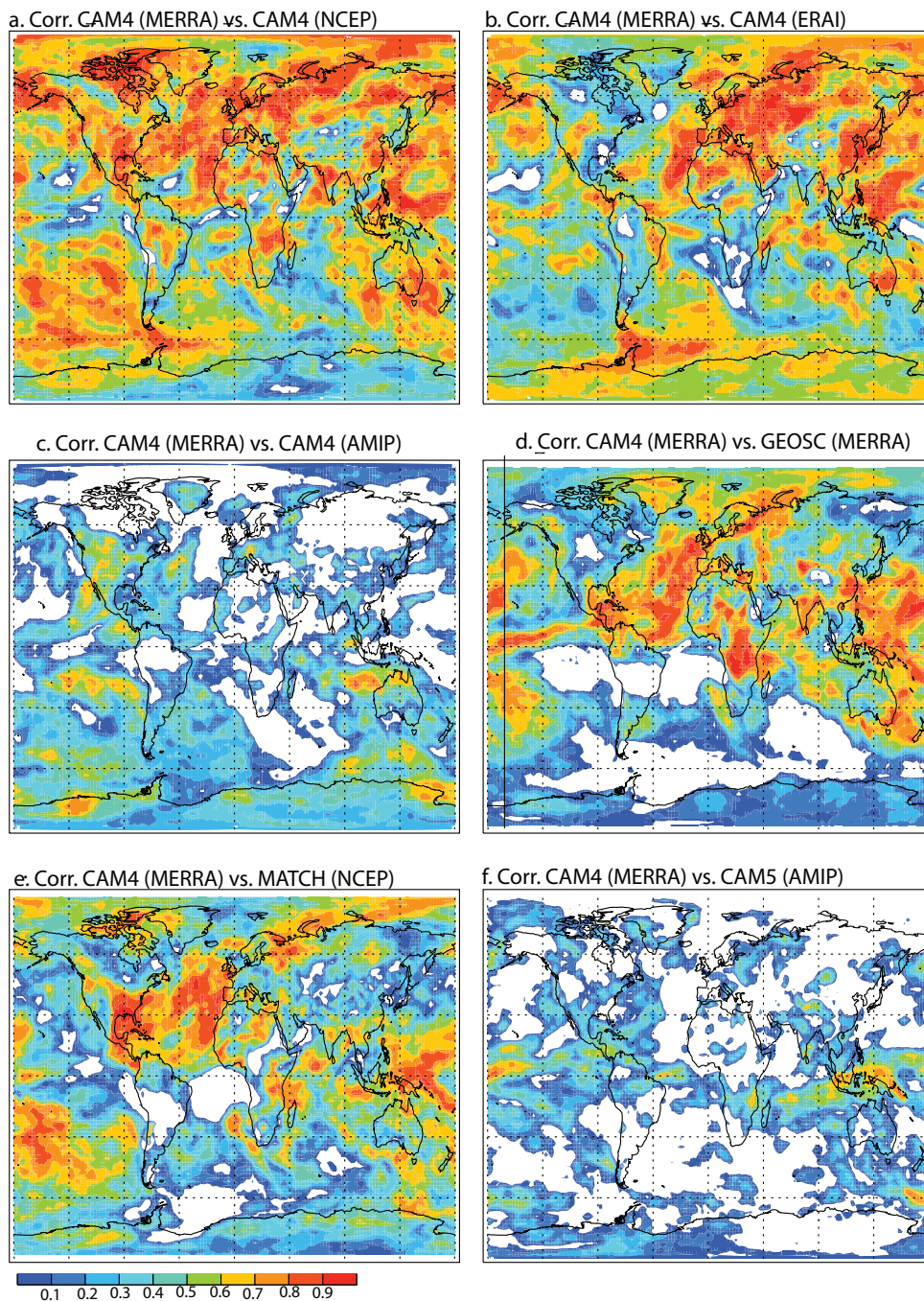
Figure 7: Spatial plot of the temporal rank correlation of the annual mean modelled surface concentration in the CAM4 (MERRA) case compared to each of the other model simulations: a) CAM4 (NCEP), b) CAM4 (ERAI), c) CAM4 (AMIP), d) GCHEM (MERRA), e) MATCH (NCEP) and f) CAM5 (AMIP). At each grid-box the 16-year time series are correlated between the two model versions, and the color indicates the value of the correlation.
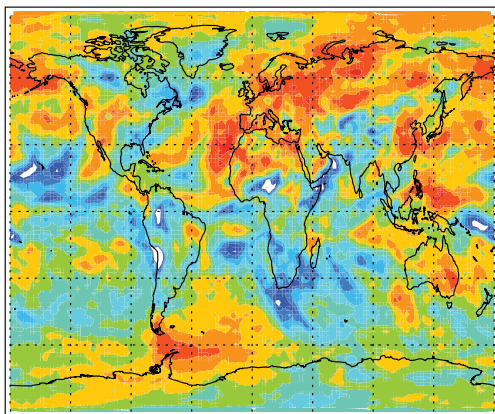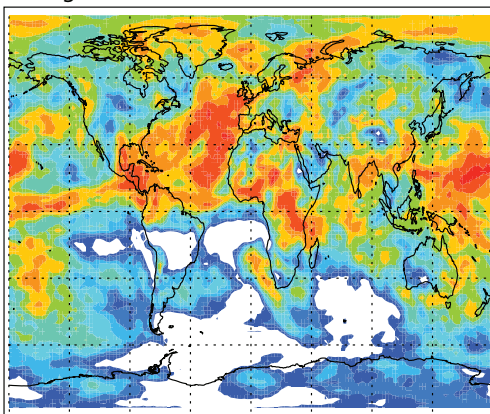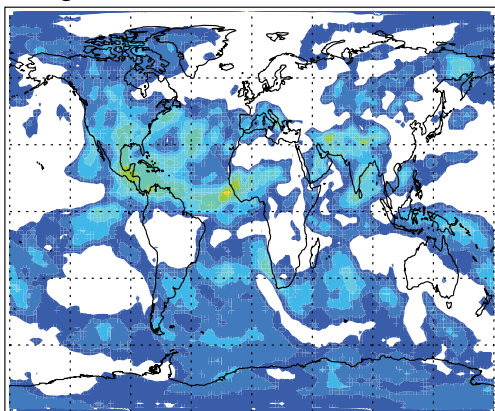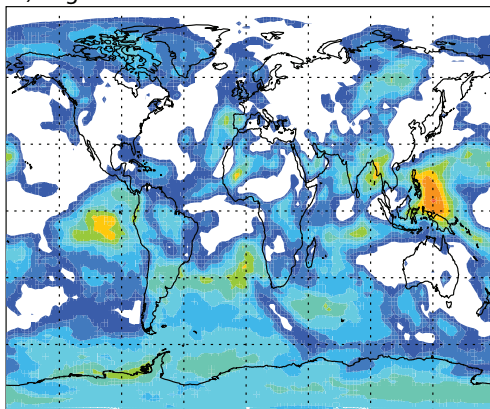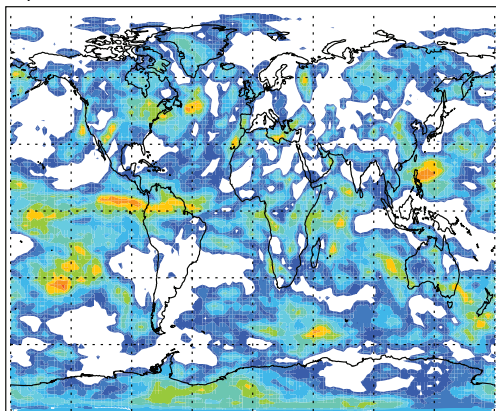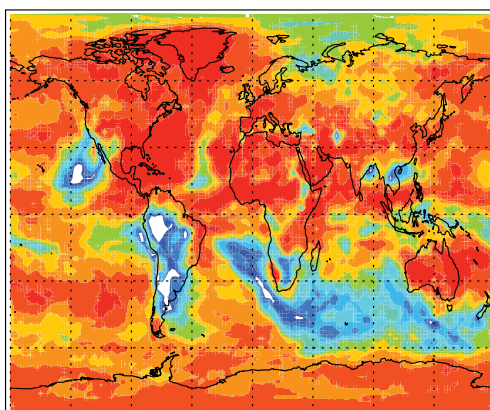
Figure 8: Spatial plot of the average temporal rank correlation of the annual mean modelled values. The correlation is calculated between the CAM4-Reanalysis models (CAM4 (MERRA);

CAM4 (NCEP) and CAM4 (ERAI) and averaged; left hand column (a) and the models driven by the same meteorology (average of CAM4 (MERRA) vs. GCHEM (MERRA) and CAM4 (NCEP) vs. MATCH (NDEP) (b) for surface concentration.   At each grid-box the 16-year time series are correlated between the models, and the color indicates the value of the correlation. The temporal

5   correlation of the climate index time series and the modelled annual mean concentration is shown in (c) NAO and (d) ENSO, where the values are the average of the correlations in the CAM4-Renalysis models. The temporal correlation of the annual mean concentration between the CAM4 (AMIP) and CAM5 (AMIP) simulations is shown (e). The temporal correlation for the seasonal cycle is shown (f), which represents the average temporal correlation for the 3 CAM4-

10   reanalysis models, using the climatological mean for each of the 12 months.
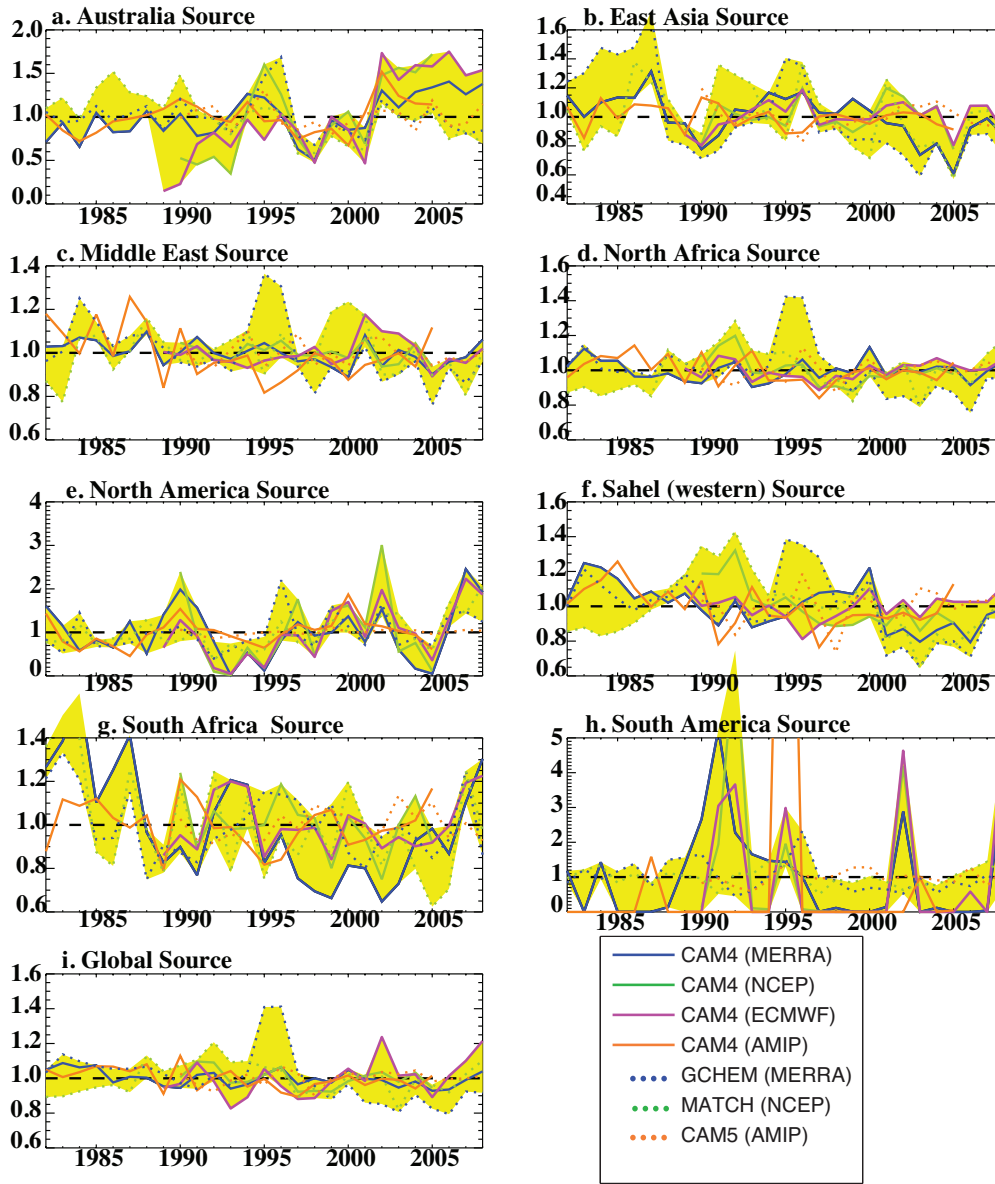
Figure 9: Time series of the annual mean source strength in different regions as simulated in the different model versions (different colors, as in legend). The regions are defined as: Australia (130 to 150°E, -35 to -25°N), East Asia (80 to 112°E, 35 to 50°N), Middle East (40 to 70°E, 10 to 45°N), North Africa (-20 to 40°E, 10 to 35°N), North America (235 to 265°E, 25 to 40°N), Sahel (western) (-20 to 13°E, 13 to 22°N), South Africa (15 to 40°E, -35 to -20°N), South America (285 to 310°E, -50 to -30°N). All time series are normalized by the climatological mean (Table S8) in order to focus on interannual variability. The yellow highlighted area is the area encompassed by the 5 reanalysis-based simulations (CAM4 (MERRA), CAM4 (NCEP), CAM4 (ERAI), GCHEM (MERRA), MATCH (NCEP)).
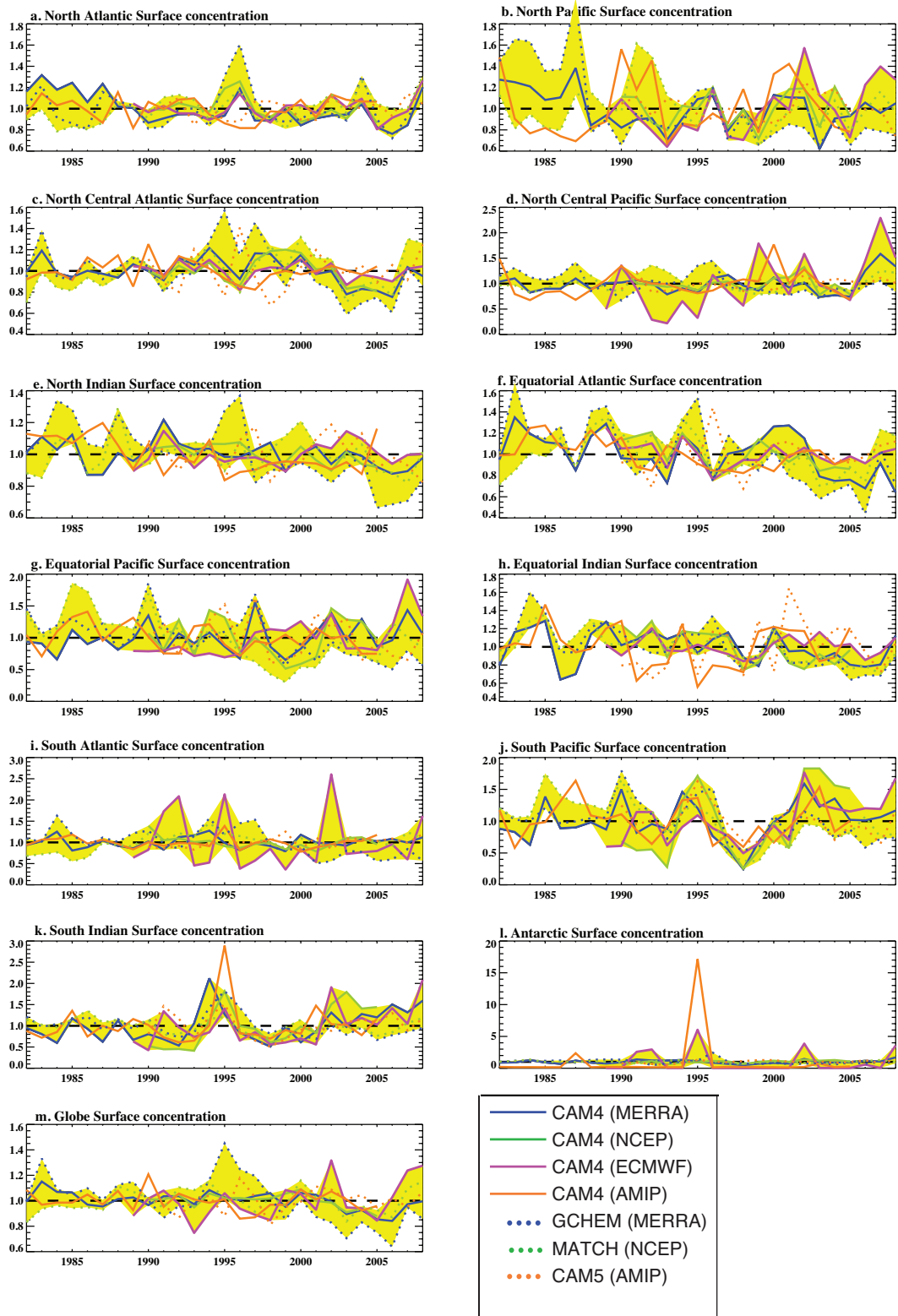
Figure 10: Time series of the annual mean surface concentration in different regions as simulated in the different model versions (different colors, as in legend), similar to Figure 9. The regions are defined as: Australia (130 to 150°E, -35 to -25°N), East Asia (80 to 112°E, 35 to 50°N), Middle East (40 to 70°E, 10 to 45°N), North Africa (-20 to 40°E, 10 to 35°N), North America (235 to 265°E, 25 to 40°N), Sahel (western) (-20 to 13°E, 13 to 22°N), South Africa (15 to 40°E, -35 to -20°N), South America (285 to 310°E, -50 to -30°N). All time series are normalized by the climatological mean in order to focus on interannual variability. The yellow highlighted area is the area encompassed by the 5 reanalysis-based simulations (CAM4 (MERRA), CAM4 (NCEP), CAM4 (ERAI), GCHEM (MERRA), MATCH (NCEP)).
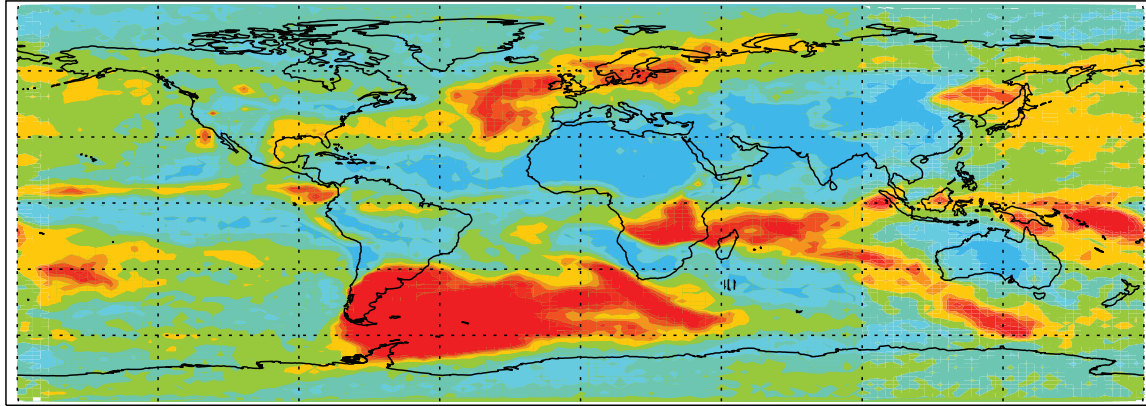
Figure 11: Monte Carlo-based estimation of the number of monthly mean observations required to obtain a 95% chance of obtaining a mean and standard deviation consistent with the true model mean between 1990-2005 at each gridbox, based on model simulations using the CAM4 (MERRA) model. More details on methods in Section 2.3.