We would like to thank Referee #1 for the suggestions for improvements, as we think these will improve the quality of the manuscript. We try to incorporate all the comments of the reviewer, as discussed below, and try to reconcile these comments with the sometimes contradictory comments by the other reviewer. Reviewer comments are in black, responses in blue.

Reviewer 1 feedback:

The authors compare several model frameworks, including different forcing data, different forcing methods and different models, to estimate the sensitivity in the seasonal and interannual variability of dust emission and transport. They find that the correlation in variability is roughly similar between simulations using different meteorological forcings and different models, while simulations forced by sea surface temperature only do not perform as well.

Considering the high uncertainty in dust simulations, investigating the sensitivity to different forcing data sets and models is an important contribution to the field of dust modelling. However, the manuscript presents a lot of information that is not necessarily relevant and that dilutes the actual results of the study. In addition, several parts of the methodology and results are unclear, which make the interpretation further confusing.

We have rewritten the methodology and results sections to be clearer and to focus primarily on interannual variability in surface concentration in order to streamline the paper. Our aim is to eliminate superfluous information that distracts from these core results, but we do use available deposition and AOD observations and mention the implications of these results for deposition and AOD, using the supplemental material.

Note that the recommendations of this reviewer contradict in some places, the recommendations of the other reviewer (who wants more details), so that we do add some details about NAO and ENSO to accommodate the constructive comments of the Reviewer #2.

The manuscript thus likely need to be rewritten before it can be considered for publication in Atmospheric Chemistry and Physics.

We have rewritten the manuscript for clarity, and hope that it is now acceptable for publication.

General comments

A. The model description is often confusing and strongly needs to be reorganized and clarified.

We have rewritten this section and added an introduction to improve clarity.

B. The discussion of the results jumps between a large number of figures and tables split between the manuscript and the supplement, which need to be reduced.

We have reduced the number of figures and tables in the manuscript, as part of our effort to focus on interannual variability in surface concentration and cut down on superfluous information. We had to add back in some figures in the supplement, due to questions from the other reviewers, but streamline whenever possible.
C. A significant part of the results further appears not to be significant, thus the discussion needs to be more focused to emphasize the actual contribution of the manuscript.

We have modified the discussion of our results to emphasize trends in dust interannual variability for particular regions, as well as the differences in model output produced by individual datasets.

Specific comments

1. Introduction

p.2 l.7-15: this part is disconnected from the previous one and needs to be separated into a new paragraph; the different space and time scales and their relevance are confusing and need some hierarchy; what are “4x fluctuations”?

The paragraph has been rewritten to make it clearer. We have rewritten out (4-fold fluctuations) to make the text clearer:

“with 4-fold fluctuations in surface concentration observed across a large region on decadal time scales (Prospero and Lamb, 2003)”

p.2 l.15: this is partially redundant with the following lines and should be reorganized.

This line has been moved to a more relevant position in the next paragraph.

p.2 l.21: this is obvious and could be more specific (interannual variability of dust emissions?).

We clarify that this is true across time scales:

“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; 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.”

p.2 l.28-29: the purpose of the sentence is unclear.

The sentence has been clarified to show that it explains our choice to use the time range 1990-2005 for this study:

“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).”

p.2 l.31: a short paragraph to describe the contents of the paper would be helpful at the end of the
introduction.

A very good suggestion: this has been added:

“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 model simulations, and the implication for observational needs.”

2. Methodology

p.3 l.5-10: the meaning of “model” is confusing here between CESM, CAM, GCM... A description of the nature of these “models” would be helpful.

A very good point. Clarification was added to the beginning of the methods section, but the models are explained in detail farther on in the paper:

“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; 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).”

p.3 l.12: “when the winds are strong” needs to be clarified.

This sentence has been rewritten to clarify (note that we don’t want to repeat the development of dust entrainment mechanisms which are well developed elsewhere in the literature, and we cite):

“Dust is entrained into the atmosphere when strong winds occur in dry, unvegetated regions with easily erodible soils (Marticorena and Bergametti, 1995).”

p.3 l.15: “as described in more detail elsewhere” needs to be rephrased.

“Elsewhere” has been clarified to mean “in previous studies”.

p.3 l.25-27: the purpose of the sentence is unclear.

Good point: we add in reference to the results section, where we discuss this in more detail:

“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).”

p.3 l.31- p.4 l.7: the methodology is unclear. Which variables are nudged? What about the wind? What is the MATCH/CAM approach? Is MATCH used here?

We add more details of this approach, which is widely used in the general circulation community to drive nudged climate model versions. Note that here we are describing the CAM4, but a similar approach is used in the MATCH model, as discussed in that section of the methods.

“In this procedure the 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; developed in Rasch et al., 1997; Mahowald et al., 1997).”

p.4 l.11-13: the purpose of the sentence is unclear; discuss conclusions of these studies?

Our goal is to point out that none of these are perfect: we modify the sentence to make some of the conclusions clearer, although it is beyond the scope of this study to review all these results:

“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; Trenberth 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.”

p.4 l.18-20: more details are needed, e.g. how the model is initialized and what is computed in the atmospheric model (CAM?) exactly (winds cannot be computed alone).
Of course the reviewer is correct: this is an atmospheric model, so all fields are calculated, not just wind, as the first manuscript erroneously indicated. We clarify this:

“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.”

p.4 l.20-22: this is unclear; what is the “correct input format”?

We clarify:

“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.”

p.4 l.31 onwards: the paragraph should be moved and merged with the description of AMIP above.

This is a good point was incorporated into the manuscript.

p.4 l.34: why is there “no inconsistency between the atmospheric model and observations”?

When models are forced to follow observations, there is an inconsistency in the model simulation and real world, which is sampled in the observations. This inconsistency often results in effective heating terms, which cause vertical velocities, for example, and lack of energy conservation. “Observations” was clarified to be “observations from the reanalysis”. AMIP simulations don’t have a conflict because they don’t use the reanalyses. We clarify:

“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).”

p.5 l.9-10: the purpose of the sentence is unclear.

We wish to specify exactly what our CAM5 setup was. We hope the additional sentences at the beginning of section 2.1 will make this clearer.

p.5 l.11-21: the contrast with CAM needs to be better emphasized.

The wording was changed to highlight differences.

p.5 l.30 onwards: this should probably appear earlier, as MATCH is referred to in the previous
section.

We respectfully disagree on this point. Even though MATCH is mentioned earlier, the CAM4 simulations are the core focus of the paper, and we wish to fully discuss them before branching out into other models. It was necessary to refer to MATCH, though, because some of the CAM methodology was developed in MATCH. We hope that the new paragraph at the beginning of the modeling section will clarify this.

p.6 l.9-11: this appears disconnected from the other paragraphs.

Good point: we move this paragraph up to the CAM4 part of the text, which makes more sense.

p.7 l.5-7: the definitions are confusing.

This paragraph has been removed from the new draft, because of the emphasis on IAV.

p.8 l.1-4: this is very confusing; does it mean that the method is wrong?

Good point: we clarify:

“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

p.8 l.8: does it mean that the simulations are taken from previous studies?

We conducted all of the CAM simulations for this study, but some of that model results was used for additional studies (with different focuses) published before this one. The MATCH and GEOS-CHEM simulations were conducted for other studies (also with different focuses), and then that data was also used here. We clarify:

“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)…”

p.8 l.11: should it be S3?

Yes, this error has been fixed.

p.8 l.24-25: the sentence is vague.

We rephrase:

“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.”

p.9 l.1: which data? The purpose of the sentence is unclear.
A clarification has been added: “these in situ and sun photometry data”

p.9 l.5 onwards: there is a lot of information presented in many figures and tables split between the main manuscript and the supporting information, which makes the discussion very difficult to follow; the information needs to be presented in a concise way and the discussion better structured.

We have tried to streamline the manuscript and reduce the number of figures in both the paper and the online supplement. We focus on surface concentration and IAV, instead of presenting so many different metrics.

p.10 l.17-32: this should be moved to the introduction.
We have added more information about the role of meteorology in variability in the introduction, in response to both reviewers, and try to explain better the focus of this paper. We think this material needs to be specifically in this section as well, as we discuss the Sahel here, as opposed to all global sources.

p.11: the whole discussion is confusing (is there a trend or not?) and its purpose is unclear, as it is not related neither to the monthly nor to the interannual variability.
This section focuses on IAV, especially trends in IAV, which we try to make sure is more clear. Our goal is to understand both whether the IAV seen in the previous study (Ridley et al., 2015) is robust across models, and whether the mechanisms hypothesized in the previous study looks robust across models. We think this type of hypothesis should be tested across multiple model frameworks, since, as we show in later sections, most IAV is not robust across the models. We try to explain better in the text:

“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.”

p.12 l.7-19: CAM4 exhibits much higher monthly variability than the other models.
We agree that some of the versions of the CAM4 have much higher variability, but not all of them. The text has been expanded to discuss this:

“The CAM4-ERA1 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.”
p.12 1.19-21: models driven by MERRA and NCEP are clearly better than ERAI and AMIP.

We include this point above.

p.12 1.25-27: the statement is weak.

We agree that both the data and our statement is weak. We have included more Southern Hemisphere data to strengthen our argument, but fundamentally we are limited by the lack of available data. We modify to indicate this:

“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.”


This section has been rewritten and modified for the new focus on IAV and surface concentration, so we hope the new section more convincingly illustrates the mentioned points.

p.13 1.28-31: this should be moved to the methods.

This information is already contained in the method section, and we think it is required here for clarify and help interpreting the results.

p.13 1.33-p.14 1.2: are the AMIP simulations meaningful then?

Comparisons with the AMIP simulations show that using reanalyzed meteorology does produced an improvement of the results when reanalyses is used, so this is helpful to those of us who can use AMIP or reanalysis data, and also show what fraction of the dust variability is due just to ocean SSTs. We clarify:

“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.”

p.15 1.8-10: is Figure 13 relevant then?

One of the goals of the study is to explore the coherence of the different models in terms of IAV, so we think Figure 13 (now Figure 9) is extremely relevant to show, and illustrate the variability of IAV across models.

p.16 1.28: should this be Figure 16? Are Figure 14 and 15 needed at all?

The caption should have been for Figure 16, and that has been corrected (the figure in question is now Figure 11). Figures 14 and 15 are less relevant, but still demonstrate important trends in regional deposition and AOD, and have thus been moved to the supplement.
p.16 l.28-30: isn’t it obvious that a whole year of data is required to estimate the annual mean?

Yes, we agree (and add in “as expected” to the text), but in some regions we need more. Our goal is to provide ‘evidence’ for observationalists so they can get funding to conduct longer time series, since this is such a constraint on any understanding of dust distributions

p.17 l.11-12: this again requires the set-up of AMIP simulations to be better explained and asks if they are meaningful.

Our goal is to understand how much AMIP-style simulations can simulate dust, so the fact that they do worse is of course meaningful and helpful. We clarify:

“The model simulations do roughly similarly well 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).”

p.17 l.25 onwards: the results do not appear that clear in section 3.

We have rewritten the results section to present the results more clearly. We also modify the conclusion section to make more clear our results, and the association with specific points in the results section.
We would like to thank Referee #2 for the suggestions for improvements, as we think these will improve the quality of the manuscript. We try to incorporate all the comments of the reviewer, as discussed below, and try to reconcile these comments with the sometimes contradictory comments by the other reviewer. Reviewer comments are in black, responses in blue.

Reviewer 2 feedback:

General comments

The manuscript addresses the highly relevant issue of the representation of the variability of mineral dust in state-of-the-art global model simulations. The study is pursued using a set of simulations including runs with the global model CAM4/5 driven by different boundary conditions and two further other global models. The article is generally well written, and the number and quality of tables and figures is good.

What disappoints is the simplistic and superficial analysis of different correlation parameters without relating them in detail to the varying surface characteristics and meteorological drivers in specific geographic regions, and to global atmospheric circulation patterns (NAO, ENSO, etc.). In the presentation of results, there is no discrimination in the contribution of different desert regions to local dust conditions, and explanations of processes behind the correlations are missing.

The main focus of this paper is to address the question of how sensitive the simulation of dust interannual variability is to the reanalyzed meteorology used, or alternatively, the modeling framework used, and to examine long term trends in global dust variability. Other studies have examined the mechanics by which meteorology drives changes in dust mobilization, transportation, and deposition, and we are not trying to repeat that interesting work here.

We think some of these comments are due to a misunderstanding of the point of the paper, so we add more justification of the goals of the paper in the introduction and change the title of the paper to focus on IAV, and the sensitivity to the reanalyses dataset (not to meteorology). However, we have expanded our analysis to discuss the relation of dust interannual variability to NAO and ENSO.

Note that the comments of this reviewer to have more detail, contradicts the comments of the other reviewer who wants less details, so we have focused the paper more on the coherence of model simulations of IAV in surface concentrations in order to accommodate the other reviewer. Expanding this paper further would make it unwieldy, and we encourage the reviewer or other authors to conduct further studies to look at whether specific mechanisms are driving the coherence of the model results driven by different datasets.

The new introduction section should explain this better:

"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; Mahowald et al.,
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., 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 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.

As an example, the link between precipitation and dust emission in the Sahel is left to the reader’s interpretation.

There is substantial discussion of the role of precipitation and dust emission in both the introduction and the Sahel source trend section, and we add more of this discussion.

The contribution of Asian dust sources to the northern hemispheric dust cycle seems underrepresented in the discussion.

The model results shown here are compared to the available observations, and seem to support the dust strength used here for the Asian sources (we include Figures S1-S6, which were previously excluded to streamline the paper to show this better). Perhaps the reviewer means that we were unable to find observations isolating the impacts of these regions, so that we did not discuss them as in depth as the North African sources, which is true, and an unfortunate fact that these sources are difficult to isolate due to other aerosols and a lack of data. There are so many other aerosols that these areas are difficult to obtain high quality information to discuss. Note also that our study focuses on the time period 1990-2005, and there is more data covering other time periods.

A more detailed description and explanation of the figures would be desirable.

We have tried to both simplify the number of figures as well as improve the figure captions, especially as noted below in detail by the reviewer.

Specific comments:

1. Page 2: The Introduction lacks a brief overview of what is already known about source strength, seasonal and inter-annual variability of dust emissions, as well as their main drivers for important dust source regions.
As indicated above, we have some details about important mechanisms to the introductory paragraph above, although perhaps not in the detail the reviewer requests, because our goal is not to repeat these studies NOR to detail all causes of IAV, but rather to see how robust model results are for IAV.

What is expected to be seen in the model results? An outline of the paper at the end of the Introduction would be helpful.

These are good suggestions, and we have added the requested information at the end of introduction:

“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 model simulations, and the implication for observational needs.”

2. Page 3, line 18: Could you provide the scale factors here?

The scale factors are different for each model and are really much more detail than needed here. The important factor is where the model was enhanced substantially for the rest of the discussion, which comes in the following sentence. More details are provided in 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).”

3. Page 6, lines: 9-11: I might have overseen this but I could not find the results in the manuscript. So, please omit the paragraph or present the other species in the text.

We move this paragraph to be within the CAM4 discussion, in response to the other reviewer, and include a reference to the section where it is used:

“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 (SO4)), and we use these to contrast the correlations between dust aerosols and other aerosols for a sensitivity study, referenced in Section 3.3.”

4. Page 6, lines: 23-24: You say that only sites where more than 50% of the modelled AOD was from dust are included in this comparison. To me it seems that due to this filtering the comparison is not independent. Would it be better to compare observed coarse-mode AOD with modelled dust AOD or to use the angstrom coefficient to filter dust events? There is no further description of the filtering below. Which wavelength has the AOD used? For an overview it would be good to have the location of observation sites on a map.
We add the location of the detailed observing sites to the maps for the appropriate variable (Figures S1, S3 and S5). We cannot use the coarse/fine mode in the AERONET observations because this could also be seasalts, especially in coastal regions, so we chose to use model results for this calculation. It is unfortunate that there is so little data across the globe to consider IAV, and one of the focus points of this paper is to try to motivate for more data collection, especially long term sites that can be clearly identified to be dust sites. Notice that, as we say in the text in several places, our goal is not to use every observation (especially low quality observations like visibility or AOD, which can be multiple aerosols), but rather to show the differences in model simulations, that all do about equally against observations, but still show substantial differences. We hope the new introduction makes this point clearer.

We add in an explanation:

“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.”

5. Page 7, lines 26-28: I do not understand why these two records have been chosen, which have a much shorter and inconsistent time coverage, while considerably less but still several AERONET observations are available in the southern hemisphere.

The problem we encountered with the AERONET sites in the southern hemisphere is that, there is substantial seasalts in the Southern Hemisphere, so that coarse mode is often dominate by that aerosol. We did find one station, where there is a dominant contribution from dust in the coarse mode, at least, and thus we have added an Australian site (Tinga Tingana), which has a large dust percentage, to our analysis, from the coarse mode. Since this is likely to have substantial other aerosols, we include it only in the Southern Hemisphere analysis in the online supplement. Much more data is needed to better resolve the dust cycle and variability in the Southern Hemisphere.

6. Page 9, lines 5-23: The analysis is too general. At least it should be clearly described which desert contributes to the dust load at the individual sites, and then the variability and relevant changing atmospheric and/or surface conditions have to be evaluated respectively. I do not see why for the AODs you are limited to regions close to source regions (line 20).

With this paper, we are not focusing so much on regional dust sources and their contribution at individual sites as we are on the modeled interannual variability of the dust once it is in the atmosphere, and whether it is consistent across models. As we try to explain in the introduction, the reviewer seems to misunderstand the goal of this paper. Other papers have examined the observed causes of variability in dust sources at different sites, and we are not trying to recreate that work, but to examine differences between dust simulations caused by the use of different reanalyses and see what fraction of the IAV is robust.

7. Page 10, lines 12-13: “[. . .] insight into how much of the variability in dust is driven by SSTs.” Can this be quantified any further?

The results are discussed in more detail later. We use the metrics described later in the Section (2.3) to do this more quantitatively.

8. Page 10, line 17: Here and elsewhere in the text, even though it is obvious, please mention the
dust source affecting the geographic region under consideration.

This is a good idea. We do not include it in this section, but rather lower, where we discuss different source regions and downwind regions.

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

9. Page 10, lines 29-32: Could also vegetation-related changes in surface roughness explain the trend in the Sahelian dust emissions as discussed by Cowie et al. (GRL, 2013), and would these changes be considered by any of the models used? What exactly are the mechanisms behind these correlations?

Thank you for bringing this paper to our attention: this is a very important paper that we had not seen. We modify the text throughout the paper to include this hypothesis, which links the other hypotheses in a novel manner. We are unable to check this hypothesis and whether it is operating our model simulations, because we did not archive surface roughness, but bring up this point. As we discuss elsewhere, this paper has too much material already, so we cannot add more information about the specific mechanisms, which is also not the point of this paper, and is better explored in other papers (like the Cowie et al., 2013 paper). In addition, Ridley et al. (2014) tested the potential role for vegetation changes on the trend in dust and found it was not a factor (at least in that simulation).

10. Page 11, line 13: Correct: “[...] supporting the idea that the source strength decreased within the time period 1990 – 2005.”

The clarification has been added.

11. Page 11, lines 29-31: Here, you may include a study by, e.g., Marsham et al. (GRL, 2011) who showed that coarse-resolution models usually fail to reproduce dust emission events related to moist convection.

Yes, this is a good point. We add the point that all of our model simulations may miss small scale features like dust devils or moist convection. We add this in the introduction:

“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).”

12. Page 12, lines 1-27: The database for the analysis in the southern hemisphere is too small to
be statistically significant. Problematic is also the shorter and, in particular, inconsistent time coverage of the measurements. Moreover, the two sites are affected by mineral dust from different sources, which is not discussed at all. You could easily broaden your database by using AERONET sun photometer observations.

We have added one additional AERONET observations in the Southern Hemisphere. As discussed above, the rest of the AERONET sites fail even if we focus on coarse mode, because they have potentially too much sea salt influence. We need more observations for the southern hemisphere that can be robustly thought of as dust.

13. Page 15, lines 15-17: The ‘global average’ is not a larger source. Could you please specify the smaller sources meant?

We have removed the reference to the global average and added examples of smaller sources.

Technical corrections:

1. Page 2, line 8: Do you mean “4-fold fluctuations”?

we mean 4-fold fluctuations. We clarify the sentence:

“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),”

2. Page 5, line 26: Wording: “[. . .] it [the model] can only be operated using reanalysis datasets, [. . .]”.

The sentence has been reworded.

3. Page 7, Eq. 1: A proper equation would look more scientific.

The equation has been formatted.

4. Page 11: IAV needs to be resolved.

We change the definition the IAV to be more clear, we hope.

5. Page 18: Correct: “whether other drivers”.

The typo has been corrected and clarified.
Sensitivity of the Interannual Variability of Mineral Aerosol Simulations to Meteorological Forcing Dataset

Molly B. Smith1,2, Natalie M. Mahowald1, Samuel Albani1,3, Aaron Perry1, Remi Losno4, Zihan Qu4,5, Beatrice Marticorena6, David A. Ridley6, Colette L. Heald6

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.
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 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, 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 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

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 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 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 precipitation, winds, surface roughness, or land use (e.g. Tegen and Miller, 1998; Mahowald et al., 2003; Miller and Tegen, 1998b; Mahowald et al., 2002; 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; Mahowald et al., 2002; Prospero and Lamb, 2004; Engelstaedter 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., 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 interannual 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 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).
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 different modeling 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 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 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; 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).

2.1.1 CAM4 dust model

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 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.

Marshall et al., 2011)
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 m²/m²) 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 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 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 factors to existing soil erodibility parameters for macro regions, in order to best match available data (Albani et al., 2014). In the CAM model, the reanalysis forcing can be used to nudge the model close to specific weather patterns so that large scale meteorology is input into the model, allowing the model flexibility in terms of 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 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; 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 observations.
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 (SO₄)), 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.
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 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 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. 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 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 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., 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 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 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 (Heimbürger et al., 2012). Because of the limited datasets in the 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 seasalts 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 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...
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):

$$\text{Variability} = \frac{\sigma}{\mu}$$

where $\sigma$ is the standard deviation of the modeled and observed values, and $\mu$ is the mean.

We also analyze the observations and model output for trends, by calculating the least squares fit slope, as well as the standard deviation of the modeled and observed values, which test the ability of the models to simulate variability, but some information about seasonal correlation is also provided, in order to provide context and comparison to previous studies.

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 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.

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 simulations using the exact same meteorology (CAM4-MERRA vs. CHEM-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 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 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; Huneues 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; Huneues 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 (Huneues 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 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 (Figure 1 vs Figure S7 or Figure 2 vs. Figure S8, Table S1 vs. S2; or Figure 3a vs. 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 2e 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 3g and 3h). 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 3g and 3h), and that the model can simulate seasonal variability better than interannual variability in both surface concentrations and AOD (Figure 1 vs. Figure 2; Figure 3a vs. 3b, 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 11). 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). 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).

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 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.

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 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 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 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,
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, 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 0.2 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. Gaiser, 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-ERA1 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 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 Tingana, the IAV variability is 0.42 in the observations, and tends to be smaller in the models.

At both Rio Gallegos and 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 5a). 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 5b), with smaller ratios of seasonal to IAV variability over the southwest US, North Atlantic to Northern Europe, and across large part 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 7a vs. Figures S10a and S10b), and similar importance of the seasonal cycle (Figure 7b 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 (Alban 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 [35]. 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 2a) had similar correlation coefficients as using different winds and model framework (e.g. CAM4 (MERRA) vs. MATCH (NCEP); Figure 7a) or different models with the same winds (e.g. CAM4 (MERRA) vs. GCHEM (MERRA); Figure 2a). 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 are some statistically significant trends in the model simulations over the 1990-2011 time period, as seen in Barbados (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 8c). 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 8a). 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 8b), 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 SO4, 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 S13). 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 6). 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 6), 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 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.

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 0.75), 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 al., 2014, Figure S1), and suggest source-receptor type relationships consistent with previous studies (e.g. Tanaka and Chiba, 2006; Mahowald, 2007). These 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 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 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 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 3) and previous studies (e.g. Huneau 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 8), 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).

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 1). The models’ ability to simulate the 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 2). Surface concentration and deposition have similar distributions of variability, while AOD tends to have less variability (Figure 2). There is more variability, especially interannual variability, in parts of the Southern Hemisphere (Figure 4). 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 the very limited observations at some stations in the Southern Hemisphere suggest that could be 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.

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 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 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 principle and previous studies (e.g. Engelstaedter 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).

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 annual mean concentrations (Figure 1). 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 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 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


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.


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:10.5194/acp-10-8287-2010, 2010.


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., 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.,


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>
<thead>
<tr>
<th>Case name</th>
<th>Model base</th>
<th>Meteorology</th>
<th>Retuned sources</th>
<th>Years available</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM4 (MERRA)</td>
<td>CAM4</td>
<td>MERRA</td>
<td>Y</td>
<td>1980-2008</td>
<td>(Albani et al., 2014)</td>
</tr>
<tr>
<td>CAM4 (NCEP)</td>
<td>CAM4</td>
<td>NCEP</td>
<td>Y</td>
<td>1989-2006</td>
<td>(Albani et al., 2014)</td>
</tr>
<tr>
<td>CAM4 (ERA)</td>
<td>CAM4</td>
<td>ERA-Interim</td>
<td>Y</td>
<td>1989-2008</td>
<td>(Albani et al., 2014)</td>
</tr>
<tr>
<td>CAM4 (AMIP)</td>
<td>CAM4</td>
<td>Online/AMIP</td>
<td>Y</td>
<td>1980-2006</td>
<td>(Albani et al., 2014)</td>
</tr>
<tr>
<td>GCHEM (MERRA)</td>
<td>GEOS-</td>
<td>MERRA</td>
<td>N</td>
<td>1982-2008</td>
<td>(Ridley et al., 2014)</td>
</tr>
<tr>
<td>MATCH (NCEP)</td>
<td>MATCH</td>
<td>NCEP</td>
<td>N</td>
<td>1982-2008</td>
<td>(Luo et al., 2003)</td>
</tr>
<tr>
<td>CAM5 (AMIP)</td>
<td>CAM5</td>
<td>Online/AMIP</td>
<td>Y</td>
<td>1990-2008</td>
<td>(Albani et al., 2014)</td>
</tr>
</tbody>
</table>
Table 2: Description of observational sites included here. Locations are plotted in the maps (Figure S1, S3 and S5).

<table>
<thead>
<tr>
<th>Site</th>
<th>Latitude</th>
<th>Longitude °E</th>
<th>Years</th>
<th>Citation or PI for Aeronet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface Concentration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Banizombou</td>
<td>14</td>
<td>3</td>
<td>2006-2013</td>
<td>(Marticorena et al., 2010)</td>
</tr>
<tr>
<td>Bermuda</td>
<td>32</td>
<td>-65</td>
<td>1989-1997</td>
<td>(Arimoto et al., 1995)</td>
</tr>
<tr>
<td>Cinzana</td>
<td>13</td>
<td>-6</td>
<td>2006-2013</td>
<td>(Marticorena et al., 2010)</td>
</tr>
<tr>
<td>Izana</td>
<td>26</td>
<td>-16</td>
<td>1989-1998</td>
<td>(Arimoto et al., 1995)</td>
</tr>
<tr>
<td>Machead</td>
<td>53</td>
<td>-10</td>
<td>1989-1994</td>
<td>(Arimoto et al., 1995)</td>
</tr>
<tr>
<td>Mbour</td>
<td>14</td>
<td>-17</td>
<td>2006-2013</td>
<td>(Marticorena et al., 2010)</td>
</tr>
<tr>
<td>Miami</td>
<td>26</td>
<td>-80</td>
<td>1974-1999</td>
<td>(Prospero, 1999)</td>
</tr>
<tr>
<td>Midway</td>
<td>28</td>
<td>-177</td>
<td>1982-2000</td>
<td>(Prospero and Savoie, 1989)</td>
</tr>
<tr>
<td>AOD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bahrain</td>
<td>26</td>
<td>50</td>
<td>1998-2006</td>
<td>B. Holben</td>
</tr>
<tr>
<td>Dalanzadgad</td>
<td>43</td>
<td>104</td>
<td>1997-2012</td>
<td>B. Holben</td>
</tr>
<tr>
<td>Ilorin</td>
<td>8</td>
<td>4</td>
<td>1998-2009</td>
<td>R. Pinker</td>
</tr>
<tr>
<td>Sede Boker</td>
<td>36</td>
<td>34</td>
<td>1998-2010</td>
<td>A. Karnel</td>
</tr>
<tr>
<td>Southern Hemisphere</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rio Gallegos</td>
<td>-52</td>
<td>-69</td>
<td>2011-2014</td>
<td>(Zihan, 2016)</td>
</tr>
<tr>
<td>Kerounen deposition</td>
<td>-49</td>
<td>70</td>
<td>2008-2010</td>
<td>(Heimbucher et al., 2012)</td>
</tr>
<tr>
<td>Tinga Tingana AOD</td>
<td>-29</td>
<td>140</td>
<td>2002-2012</td>
<td>R. Mitchell</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Case</th>
<th>Surface Concentration Correlation (log space)</th>
<th>Deposition Correlation (log space)</th>
<th>AOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM4 (MERRA)</td>
<td>0.73 (0.89)</td>
<td>0.94 (0.84)</td>
<td>0.73</td>
</tr>
<tr>
<td>CAM4 (NCEP)</td>
<td>0.67 (0.86)</td>
<td>0.79 (0.84)</td>
<td>0.67</td>
</tr>
<tr>
<td>CAM4 (ERA-I)</td>
<td>0.48 (0.81)</td>
<td>0.58 (0.76)</td>
<td>0.87</td>
</tr>
<tr>
<td>CAM4 (AMIP)</td>
<td>0.79 (0.78)</td>
<td>0.73 (0.84)</td>
<td>0.41</td>
</tr>
<tr>
<td>GCHEM (MERRA)</td>
<td>0.73 (0.90)</td>
<td>0.63 (0.84)</td>
<td>0.43</td>
</tr>
<tr>
<td>MATCH (NCEP)</td>
<td>0.83 (0.84)</td>
<td>0.43 (0.81)</td>
<td>0.32</td>
</tr>
<tr>
<td>CAM5 (AMIP)</td>
<td>0.79 (0.89)</td>
<td>0.59 (0.85)</td>
<td>0.70</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM4 (MERRA)</td>
<td>-0.014 (0.0054)</td>
<td>-0.0065</td>
<td>-0.0005</td>
<td>0.0035</td>
</tr>
<tr>
<td>CAM4 (NCEP)</td>
<td>-0.0017 (0.016)</td>
<td>-0.0169</td>
<td>-0.0074</td>
<td>0.0425</td>
</tr>
<tr>
<td>CAM4 (ERAI)</td>
<td>-0.0058 (0.0051)</td>
<td>0.0006</td>
<td>0.002</td>
<td>0.0008</td>
</tr>
<tr>
<td>CAM4 (AMIP)</td>
<td>-0.0079 (0.0035)</td>
<td>-0.0061</td>
<td>-0.0037</td>
<td>0.0186</td>
</tr>
<tr>
<td>GCHEM (MERRA)</td>
<td>-0.025 (0.0047)</td>
<td>-0.021</td>
<td>-0.0072</td>
<td>0.0035</td>
</tr>
<tr>
<td>MATCH (NCEP)</td>
<td>-0.0087 (0.0059)</td>
<td>-0.0047</td>
<td>0.027</td>
<td>0.0425</td>
</tr>
<tr>
<td>CAM5 (AMIP)</td>
<td>-0.01 (0.01)</td>
<td>0.0027</td>
<td>0.029</td>
<td>0.0089</td>
</tr>
</tbody>
</table>

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.
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).

<table>
<thead>
<tr>
<th>Model/Observations</th>
<th>Rio Gallego Surface concentrations</th>
<th>Kerguelen deposition</th>
<th>Tinga Tingana AOD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IAV variability</td>
<td>Ratio variability from seasonal cycle over IAV</td>
<td>IAV variability</td>
</tr>
<tr>
<td>CAM4 (MERRA)</td>
<td>1.78</td>
<td>1.06</td>
<td>0.22</td>
</tr>
<tr>
<td>CAM4 (NCEP)</td>
<td>2.13</td>
<td>0.79</td>
<td>0.27</td>
</tr>
<tr>
<td>CAM4 (ERAI)</td>
<td>2.66</td>
<td>0.84</td>
<td>1.42</td>
</tr>
<tr>
<td>CAM4 (AMIP)</td>
<td>3.86</td>
<td>0.55</td>
<td>2.06</td>
</tr>
<tr>
<td>GCHEM (MERRA)</td>
<td>0.67</td>
<td>1.26</td>
<td>0.32</td>
</tr>
<tr>
<td>MATCH (NCEP)</td>
<td>0.68</td>
<td>1.37</td>
<td>0.17</td>
</tr>
<tr>
<td>CAM5 (AMIP)</td>
<td>0.42</td>
<td>1.42</td>
<td>0.46</td>
</tr>
<tr>
<td>Obs</td>
<td>0.10</td>
<td>1.39</td>
<td>0.08</td>
</tr>
</tbody>
</table>
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.

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

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

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 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 30°N; 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); 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).

<table>
<thead>
<tr>
<th>Region</th>
<th>CAM4-RE across model Correlation</th>
<th>NAO correlation</th>
<th>El Nino Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Atlantic</td>
<td>0.66</td>
<td>0.10</td>
<td>0.45</td>
</tr>
<tr>
<td>North Pacific</td>
<td>0.51</td>
<td>0.19</td>
<td>0.62</td>
</tr>
<tr>
<td>North Central</td>
<td>0.75</td>
<td>0.04</td>
<td>-0.10</td>
</tr>
</tbody>
</table>
Table 8: Deposition into 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 following columns show the climatological mean deposition flux (Tg/year) for each model simulation. 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 30°N, 120 to 270°E); North Indian (10 to 30°N, 30 to 120°E); Equatorial Atlantic (-10 to 10°N; 120 to 285°E); Equatorial Pacific (-10 to 10°N; 120 to 285°E); 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).
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). 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; Jlo: 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 variability 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 concentrations.

**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 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-Reanalysis 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-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 10°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 1: 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.
While p

We consider separately here the seasonal cycle in the surface concentrations and the AOD produced by each simulation at fifteen specific sites (Figures 1a-1i and 2a-2f), compared with the interannual variability at those sites (Figure 3a-3i and 4a-4f), which when combined give the monthly mean variability (Figure S1 and S2). Notice that each time series in these figures is divided by the mean value across the time period, so that the seasonal variability (Figures 1 and 2) can be contrasted with the interannual variability (Figures 3 and 4) without concerns about differences in scale. The amount of variability (represented here by the standard deviation of the monthly mean over the annual mean; Eq. 1 in Section 2.3) in surface concentration simulated by the range of the models encompasses that of the observations at the different sites (Table S1; Figure 5a), which is between 0.75 (Mbour) and 2.0 (Bermuda). Note that as seen previously (e.g. Tegen and Miller, 1998; Mahowald et al., 2003), the largest variabilities in both surface concentration and AOD are seen downwind from the source regions, in ocean regions that intermittently experience events (e.g. Bermuda, where there was a specific event in the Atlantic described by Ginoux et al., 2001). The models simulate the stronger variability at Bermuda (for example), and the smaller variability over the source regions (i.e. Mbour). Much, but not all of this variability in the observations comes from the seasonal cycle, both in the models and in the observations in these stations (Table S2; Figure 5c). The fraction of the monthly mean variability that is due to the seasonal cycle varies from 0.5 (Izana) to 0.94 (Midway) (Figure 5c; Table S2).

The variability in the AOD (Table S1; Figure 2 and 4; Figure 5b), tends to be smaller than in the surface concentrations, which is likely to be due to both the fact that for the AODs, we are limited to regions close to source regions, and the fact that AODs are more integrated quantities as seen in previous studies (Mahowald et al., 2003). Notice that again, close to the source regions, the seasonal variability in AODs is the dominant portion of the variability in both the models and observations (Figure 5d; Table S2).

Most of the ability of the models to simulate the variability of the observations comes from the seasonal cycle (Figure 5e,f vs. Figure 5g, h). There is very little skill in the models’ ability to simulate interannual variability at any of the stations (Figure 5i,j).

In all variables (surface concentration, deposition or AOD), the southern hemisphere has much larger variability, and much of this extra variability is due to interannual variability (Figure 8).
All of the models included here, except the GCHEM (MERRA) and MATCH (NCEP) simulations, include seasonal and interannual variability in vegetation as a control on sources (CAM4 and CAM5) (Albani et al., 2014; Ridley et al., 2014). It appears that in many locations the model simulations including time varying LAI have larger variability (Figure S3; contrast f vs. others, especially in the Southern Oceans). However, over the North Atlantic, for example, one simulation without time varying LAI (MATCH (NCEP)) has greater variability than the CAM4 and CAM5 simulations, which have time varying LAI (Figure S3f vs. others). In the North Pacific, the MATCH (NCEP) simulation has more variability than the CAM4 (ERAI) simulation, which is consistent with the MATCH (NCEP) simulation being able to simulate the seasonal cycle in transport to the Pacific (Luo et al., 2003) without having a seasonal cycle in LAI, in contrast to other models (e.g. (Tegen et al., 2002) which is based on ERAI winds). This highlights how the details of model construction and meteorology can change model behavior, making it difficult to broadly generalize about what model requirements (e.g. time varying LAI) are needed to match observed seasonal or interannual variability.

, which is consistent with the statistically significant but low correlations between such ocean-forced phenomenon as NAO and El Nino and dust (e.g. Moulin and Chiapello, 2004; Mahowald et al., 2003).

If we examine the regional correlation coefficients of dust surface concentration to ENSO (Table 6), we see that the strongest correlations occur in the ocean basins where this oscillation occurs. The North Pacific had the highest correlation coefficient (0.62), followed by the North Atlantic (0.45) and Equatorial Pacific (0.42), with a strong anticorrelation over the Antarctic Ocean (-0.63). The NAO also featured a strong Antarctic anticorrelation (-0.42), and had the highest correlation coefficient over the South Indian Ocean (0.49). Again, this is similar to previous studies (e.g. Mahowald, 2003). The correlations tend to be lower near the southwestern US, and in the Southern hemisphere (Figure 10).

The correlation coefficients averaged over the CAM4-reanalysis models (Figure 11a, c and f) suggest that the surface concentration and deposition are similarly correlated, but that AOD has smoother and higher correlations. This is consistent with AOD being a more integrated quantity (Figure 8) (Huneeus et al., 2011; Albani et al., 2014). The lowest correlations occur over remote ocean regions, especially in the Southern Hemisphere, for all the variables. Note also that the correlation coefficients between the simulations using the same modeling framework but different meteorology (Figure 11a, c and f) had similar correlation coefficients, with a similar spatial structure, as the correlations between simulations using different modeling frameworks but the same meteorology (Figure 11b, d, and e). This suggests that meteorology and the modeling framework are equally important for simulating variability.

The magnitude and the spatial structure of the correlation in the seasonal cycle is similar for all the variables as that seen using all monthly means (Figure 12a, c, e vs. 11a, c, e), suggesting that most of the correlation between the model versions is due to the correlation in the seasonal cycle (Figure 12). This is made clearer when considering just
the correlation of the annual means, showing that the models simulations are much less similar in their interannual variability (Figure 12b, d and f). Note however, that the spatial structure in the correlations for IAV is very different than considering either the monthly mean or seasonal cycle (Figure 12b, d, f vs. 12 a, c, e).

A comparison of other aerosols in two of the CAM4-reanalysis based simulations available here (Figure S5), is consistent with the idea that dust is highly variable, and that there is some correlation between the models driven by different meteorology far from the sources as well as close, but that interannual variability can be quite different for transport of aerosols (Figure S5). 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.
Surface concentrations
Table 5: Variability in Southern Hemisphere
Values for the monthly variability (monthly average standard deviation divided by mean) and the fraction of the variability from the seasonal cycle (climatological monthly average standard deviation, divided by monthly mean standard deviation) for the surface concentration in the model cases and data from Rio Gallego and deposition data from Kerguelen (locations listed in Table 1).
<table>
<thead>
<tr>
<th>Model/Observations</th>
<th>Rio Gallego Surface concentrations</th>
<th>Kerguelen deposition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monthly variability</td>
<td>Fraction variability from seasonal cycle</td>
</tr>
<tr>
<td>CAM4 (MERRA)</td>
<td>5.67</td>
<td>0.33</td>
</tr>
<tr>
<td>CAM4 (NCEP)</td>
<td>5.89</td>
<td>0.29</td>
</tr>
<tr>
<td>CAM4 (ERAI)</td>
<td>7.52</td>
<td>0.30</td>
</tr>
<tr>
<td>CAM4 (AMIP)</td>
<td>8.98</td>
<td>0.24</td>
</tr>
<tr>
<td>GCHEM (MERRA)</td>
<td>1.94</td>
<td>0.43</td>
</tr>
<tr>
<td>MATCH (NCEP)</td>
<td>2.18</td>
<td>0.42</td>
</tr>
<tr>
<td>CAM5 (AMIP)</td>
<td>2.01</td>
<td>0.71</td>
</tr>
<tr>
<td>Obs</td>
<td>0.70</td>
<td>0.20</td>
</tr>
</tbody>
</table>
Table 6: 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 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); 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).

<table>
<thead>
<tr>
<th>Region</th>
<th>CAM4-RE across model Correlation</th>
<th>NAO correlation</th>
<th>El Nino Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Atlantic</td>
<td>0.66</td>
<td>0.10</td>
<td>0.45</td>
</tr>
<tr>
<td>North Pacific</td>
<td>0.51</td>
<td>0.19</td>
<td>0.62</td>
</tr>
<tr>
<td>Region</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>North Central Atlantic</td>
<td>0.75</td>
<td>0.04</td>
<td>-0.10</td>
</tr>
<tr>
<td>North Central Pacific</td>
<td>0.46</td>
<td>-0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>North Indian</td>
<td>0.30</td>
<td>0.13</td>
<td>0.38</td>
</tr>
<tr>
<td>Equatorial Atlantic</td>
<td>0.59</td>
<td>-0.02</td>
<td>-0.31</td>
</tr>
<tr>
<td>Equatorial Pacific</td>
<td>0.19</td>
<td>-0.12</td>
<td>0.42</td>
</tr>
<tr>
<td>Equatorial Indian</td>
<td>0.31</td>
<td>-0.18</td>
<td>-0.15</td>
</tr>
<tr>
<td>South Atlantic</td>
<td>0.11</td>
<td>-0.22</td>
<td>-0.42</td>
</tr>
<tr>
<td>South Pacific</td>
<td>0.65</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>South Indian</td>
<td>0.46</td>
<td>0.29</td>
<td>0.16</td>
</tr>
<tr>
<td>Antarctic</td>
<td>0.28</td>
<td>-0.42</td>
<td>-0.63</td>
</tr>
<tr>
<td>Global</td>
<td>0.42</td>
<td>0.01</td>
<td>-0.03</td>
</tr>
<tr>
<td>Region</td>
<td>CAM4-RE across model Correlation</td>
<td>NAO correlation</td>
<td>El Nino Correlation</td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------------------------</td>
<td>-----------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>North Atlantic</td>
<td>0.66</td>
<td>0.10</td>
<td>0.45</td>
</tr>
<tr>
<td>North Pacific</td>
<td>0.51</td>
<td>0.19</td>
<td>0.62</td>
</tr>
<tr>
<td>North Central Atlantic</td>
<td>0.75</td>
<td>0.04</td>
<td>-0.10</td>
</tr>
<tr>
<td>North Central Pacific</td>
<td>0.46</td>
<td>-0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>North Indian</td>
<td>0.30</td>
<td>0.13</td>
<td>0.38</td>
</tr>
<tr>
<td>Equatorial Atlantic</td>
<td>0.59</td>
<td>-0.02</td>
<td>-0.31</td>
</tr>
<tr>
<td>Equatorial Pacific</td>
<td>0.19</td>
<td>-0.12</td>
<td>0.42</td>
</tr>
<tr>
<td>Equatorial Indian</td>
<td>0.31</td>
<td>-0.18</td>
<td>-0.15</td>
</tr>
<tr>
<td>South Atlantic</td>
<td>0.11</td>
<td>-0.22</td>
<td>-0.42</td>
</tr>
<tr>
<td>South Pacific</td>
<td>0.65</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>South Indian</td>
<td>0.46</td>
<td>0.29</td>
<td>0.16</td>
</tr>
<tr>
<td>Antarctic</td>
<td>0.28</td>
<td>-0.42</td>
<td>-0.63</td>
</tr>
<tr>
<td>Global</td>
<td>0.42</td>
<td>0.01</td>
<td>-0.03</td>
</tr>
<tr>
<td>Page 35: [474] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [475] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [476] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [477] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [478] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [479] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt, Not Bold, Not Italic, Font color: Auto</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [480] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [481] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [482] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [483] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [484] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>None, Space Before: 0 pt, Line spacing: single, Don't keep with next, Don't keep lines together</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [485] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt, Not Bold, Not Italic, Font color: Auto</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page 35: [486] Formatted</td>
<td>Smith, Molly B</td>
<td>11/24/16 2:05:00 PM</td>
<td></td>
</tr>
<tr>
<td>Font: 12 pt</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 86: 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 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); 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).

<table>
<thead>
<tr>
<th></th>
<th>CAM4-RE across model Correlation</th>
<th>NAO correlation</th>
<th>El Nino Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Atlantic</td>
<td>0.66</td>
<td>0.10</td>
<td>0.45</td>
</tr>
<tr>
<td>North Pacific</td>
<td>0.51</td>
<td>0.19</td>
<td>0.62</td>
</tr>
<tr>
<td>North Central Atlantic</td>
<td>0.75</td>
<td>0.04</td>
<td>-0.10</td>
</tr>
<tr>
<td>North Central Pacific</td>
<td>0.46</td>
<td>-0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>North Indian</td>
<td>0.30</td>
<td>0.13</td>
<td>0.38</td>
</tr>
<tr>
<td>Equatorial Atlantic</td>
<td>0.59</td>
<td>-0.02</td>
<td>-0.31</td>
</tr>
<tr>
<td>Equatorial Pacific</td>
<td>0.19</td>
<td>-0.12</td>
<td>0.42</td>
</tr>
<tr>
<td>Equatorial Indian</td>
<td>0.31</td>
<td>-0.18</td>
<td>-0.15</td>
</tr>
<tr>
<td>South Atlantic</td>
<td>0.11</td>
<td>-0.22</td>
<td>-0.42</td>
</tr>
<tr>
<td>South Pacific</td>
<td>0.65</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>South Indian</td>
<td>0.46</td>
<td>0.29</td>
<td>0.16</td>
</tr>
<tr>
<td>Antarctic</td>
<td>0.28</td>
<td>-0.42</td>
<td>-0.63</td>
</tr>
<tr>
<td>Global</td>
<td>0.42</td>
<td>0.01</td>
<td>-0.03</td>
</tr>
</tbody>
</table>
Table 8: Deposition into 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 following columns show the climatological mean deposition flux (Tg/year) for each model simulation. 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); 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).

<table>
<thead>
<tr>
<th>Region</th>
<th>Correlation</th>
<th>Mean CAM4-Re</th>
<th>Mean all</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Atlantic</td>
<td>0.68</td>
<td>54 +/- 7%</td>
<td>61 +/- 17%</td>
</tr>
<tr>
<td>North Pacific</td>
<td>0.50</td>
<td>30 +/- 30%</td>
<td>30 +/- 60%</td>
</tr>
<tr>
<td>North Central Atlantic</td>
<td>0.66</td>
<td>91 +/- 8%</td>
<td>120 +/- 30%</td>
</tr>
<tr>
<td>North Central Pacific</td>
<td>0.33</td>
<td>20 +/- 90%</td>
<td>20 +/- 90%</td>
</tr>
<tr>
<td>North Indian</td>
<td>0.10</td>
<td>100 +/- 30%</td>
<td>110 +/- 40%</td>
</tr>
<tr>
<td>Equatorial Atlantic</td>
<td>0.39</td>
<td>77 +/- 20%</td>
<td>80 +/- 35%</td>
</tr>
<tr>
<td>Equatorial Pacific</td>
<td>0.33</td>
<td>4 +/- 100%</td>
<td>6 +/- 130%</td>
</tr>
<tr>
<td>Equatorial Indian</td>
<td>0.31</td>
<td>23 +/- 25%</td>
<td>21 +/- 23%</td>
</tr>
<tr>
<td>South Atlantic</td>
<td>0.16</td>
<td>3 +/- 60%</td>
<td>3 +/- 90%</td>
</tr>
<tr>
<td>South Pacific</td>
<td>0.62</td>
<td>1 +/- 40%</td>
<td>3 +/- 60%</td>
</tr>
<tr>
<td>South Indian</td>
<td>0.38</td>
<td>2 +/- 40%</td>
<td>4 +/- 75%</td>
</tr>
<tr>
<td>Antarctic</td>
<td>0.35</td>
<td>15 +/- 100%</td>
<td>30 +/- 100%</td>
</tr>
<tr>
<td>Global</td>
<td>0.45</td>
<td>450 +/- 10%</td>
<td>510 +/- 24%</td>
</tr>
</tbody>
</table>
Figure 1: Climatological monthly average time series of dust concentration divided by the mean of each time series (unitless) at stations a) Banizoumbou, b) Barbados, c) Bermuda, d) Cinzana, e) Izana, f) Macehead, g) Mbour, h) Miami, i) Midway. Observations are in black and are from the locations and citations in Table 2. 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 monthly means are divided by the long term annual mean to allow comparison with interannual variability, since they are similarly normalized (Figure 3).
Figure 2: Climatological monthly average aerosol optical depth (AOD) for model simulations (based on dust only), compared with AERONET observations for: a) Bahrain b) Dahkla c) Dalanzadgad, d) Dhabi, e) Ilorin and f) Sede Boker for each of the different model versions (Colors and information are the same as in Figure 1, but for the AOD). Observational data from AERONET stations (citations and locations listed in Table 2). The monthly means are divided by the long term annual mean to allow comparison with interannual variability, since they are similarly normalized (Figure 4).
a. Variability in Conc

b. Variability in AOD

c. Frac Seasonal Variability in Conc

d. Frac Seasonal Variability in AOD

e. Corr in Conc Monthly

f. Corr in AOD Monthly

g. Corr in Conc Seasonal

h. Corr in AOD Seasonal

i. Corr in Conc IAV

j. Corr in AOD IAV

Page 38: [607] Deleted  Cornell University  11/15/16 2:51:00 PM  monthly
Page 38: [607] Deleted  Cornell University  11/15/16 2:51:00 PM  monthly
Page 38: [607] Deleted  Cornell University  11/15/16 2:51:00 PM  monthly
Page 38: [607] Deleted  Cornell University  11/15/16 2:51:00 PM  monthly
Page 38: [607] Deleted  Cornell University  11/15/16 2:51:00 PM  monthly
Page 38: [608] Deleted  Smith, Molly B  11/24/16 1:03:00 PM  the
a. Surf. conc. variability

b. Fraction of surf. conc. variability from seasonal cycle

c. Deposition variability

d. Fraction of deposition variability from seasonal cycle

e. AOD variability

f. Fraction of AOD variability from seasonal cycle
Figure 104: Time series of the annual mean deposition flux in different ocean regions as simulated in the different model versions (different colors, as in legend). All time series are normalized by the climatological mean in order to focus on interannual variability (shown...
in Table S9). The yellow highlighted area is the area encompassed by the 5 reanalysis-based simulations (CAM4 (MERRA), CAM4 (NCEP), CAM4 (ERAI), GCHEM (MERRA), MATCH (NCEP)). The ocean basin areas are defined as over ocean in the regions in Table S8 from (Gregg et al., 2003).

Figure 115: Time series of the annual mean Aerosol Optical Depth (AOD) in different ocean regions as simulated in the different model versions (different colors, as in legend). All time series are normalized by the climatological mean (Table S10) 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)). The ocean basin areas are defined as over ocean in the regions in Table S8 from (Gregg et al., 2003).


