Atmospheric and Oceanic Sciences
Forrestal Campus, Princeton University
NOAA Geophysical Fluid Dynamics Laboratory
Email: bpu@princeton.edu

Professor Philip Stier
Atmospheric Physics
Clarendon Laboratory
University of Oxford
Parks Road
Oxford
OX1 3PU

June 14th, 2018

Dear Editor Stier,

We have submitted a revised paper entitled "How reliable are CMIP5 models in simulating dust optical depth?" by B. Pu and P. Ginoux for consideration for *Atmospheric Chemistry and Physics*. The helpful comments from the two anonymous reviewers are sincerely appreciated. Our replies to each reviewer's comments are attached. We also made some edits in the manuscript.

We gratefully appreciate your time and consideration!

Sincerely,

Bing Pu and Paul Ginoux

Interactive comment on "How reliable are CMIP5 models in simulating dust optical depth?" by Bing Pu and Paul Ginoux
Anonymous Referee #1

We thank the reviewer for very helpful comments. We reply to your comment (in Italic) below.

This work examines the performance of seven CMIP5 climate models with interactive dust emissions schemes against dust optical depth (DOD) from MODIS Deep Blue aerosol products. The performance assessment to reproduce magnitude, spatial pattern and variations of observed DOD is conducted in nine regions, namely North Africa, Middle East, Northern China, North America, India, Southeastern Asia, South Africa, South America and Australia. Furthermore, interannual variations of DOD are also examined together with the impact on it of controlling factors such as 10 m surface wind, precipitation and surface bareness derived from leaf are index (LAI) data. In order to examine the relative contribution of these controlling factors to DOD multiple linear regression is applied on both, observations and models. Calculated regression coefficients in addition to observed and simulated controlling factors are then used to project DOD to the future (both observations and models).
The authors show that although the models can reproduce the global distribution of DOD over land under present conditions, with a better representation over northern than southern hemisphere, the interannual variability of DOD is all in all not well captured by CMIP5 models. Furthermore, models also do not reproduce the observed relations between the DOD and the examined controlling factors. Projected changes of CMIP5 model mean under the RCP8.5 scenario are presented and compared to projections of a regression model.
The research presented is interesting and the paper is well written. As the authors mention in their introduction, performance of CMIP5 models to simulate dust has received little attention and this work is a good first step to change this. I recommend this paper to be published in ACP after some comments have been addressed.

General Comments:
1. The authors highlight the importance of examining the performance of current climate models in simulating dust and they choose to assess this performance by evaluating simulated DOD. In fact, in lines 73-75 the authors claim that evaluating DOD in "CMIP5 models will provide a clear picture of model capability of dust simulation". Although optical depth is a very common variable when it comes to validate models with respect to aerosols (be it dust or any other species), it is an integrated variable and therefore it does not provide any insight into the performance to reproduce the vertical distribution of aerosols. It has been shown that regional and global dust models can present similar performance in simulating AOD but present large diversity in emissions, deposition, surface concentration and vertical distribution (Huneeus et al., 2016). Although that study refers to forecast application, it is consistent with the findings in Huneeus et al. (2011). The authors should acknowledge this limitation in the discussion or conclusions, that although this evaluation is informative and necessary, it does not provide a full picture of current climate models to simulate dust. This similar

performance in optical depth compared to large diversity in other parameters such as emissions, deposition and surface concentration might be linked to the practice to use AOD to tune dust simulations. Is this a practice that is also used in climate models?

We thank the reviewer for pointing out that DOD cannot provide a full picture of dust modeling skill by CMIP5 models. We modified lines 75-77: "A comprehensive evaluation of the climatology and interannual variation of global dust optical depth (DOD) in CMIP5 models will provide insights into models' capability of simulating the integrated aerosol extinction due to dust, which is one of the key variables that determine radative forcing of dust to the climate system." and lines 604-609: "Since DOD is an integrated variable, it does not reflect the vertical distribution of dust aerosols. As pointed by Huneeus et al., (2016), dust models with similar performance in simulating aerosol optical depth may have quite large differences in simulating vertical distribution, emission, deposition, and surface concentration of dust. An overall evaluation of dust modeling capability will require detailed examination of these variables and the life cycle of dust in CMIP5 models in addition to DOD." to better address this issue.

We agree with the reviewer that the similar performance of models in simulating DOD versus their discrepancies in simulating variables such as surface concentration, emission, and deposition may be due to the fact that DOD or AOD is used to tune dust models. Same tuning method may be used in the climate models, too, and thus adds to the need to examine other variables related to dust life cycle in the CMIP5 models.

2. In addition to examining the DOD projections from CMIP5 models, the authors also project DOD using calculated regression coefficients and compare these results to the simulated ones. I have to admit that I have difficulties in seeing the usefulness of this exercise. What is the point of it?

The reason to provide a future DOD projection by the regression model in addition to CMIP5 models' projection was not clearly addressed in the previous version. We added lines 513-522 to better explain the purpose of this analysis: "Here we also present the projected change of DOD from the regression model in Figure 9. The regression model (see section 2.4 for details) is developed based on observed relationships between MODIS DOD and local controlling factors and can largely capture the interannual variations of DOD in the present-day climate (Table S1 in the Supplement). Assuming that the observed connection between DOD and these controlling factors do not change dramatically in the future, we can use this regression model and CMIP5-model projected change of controlling factors to project DOD variations. Compared to DOD projection from CMIP5 models, this approach utilizes additionally observational constrains and is likely to provide a more reliable future projection."

The authors state that similarities are found between both projections "which may be informative" without specifying for what they might me informative. What do differences and similarities of both projections tell us?

We removed "which may be informative", and modified the sentence to: "we find some similarities between the two, which adds to the confidence of projected DOD change in these regions..." (lines 690-691). Although CMIP5 models overestimate the influence of surface wind and precipitation and underestimate the role of bareness, there are some similarities between model and observations over regions such as North Africa in DJF and parts of the Arabian Peninsula in JJA (Fig. 6; lines 450-462), which indicate that models partially capture the connection between the DOD and these controlling factors in some regions. Therefore, the projection of DOD from CMIP5 models (Fig. 7) is not completely unreliable. The similarity between CMIP5 projection and the projection from the regression model thus adds to the confidence of projected change of DOD over North Africa, the Arabian Peninsula, and northern China in some seasons.

3. The authors could improve the description of the methodology applied in the study. Regression coefficients are computed by regressing DOD from MODIS onto the observed controlling factors, the same procedure is repeated with model outputs to obtain "model" regression factors. Now when the interannual variability is examined, in line 320 it is unclear whether the reconstructed DOD using model regression factors or the one based on observations. I would have thought the former but then lines 332-335 refer to the observations making me doubt what reconstruction is then used in the analysis.

The regression coefficients are derived from observations. We modified section 2.4 in the methodology section and lines 412-414 to improve the clarity.

Furthermore, regression analysis on observations is done at 1_x1_ resolution (lines 207-208) while for model outputs the regression analysis is done at 2_x2.5_ resolution. But at what resolution are the reconstructed projections done? at the observation or the model resolution? Potential impacts on the regression coefficients due to different resolution should also briefly be discussed.

For future projection, the regression coefficient is interpolated to a 2° by 2.5° grid to be consistent with model output. So the projected DOD is also on a 2° by 2.5° grid. We modified lines 282-284 to clarify this and discuss potential impacts of the interpolation: "The regression coefficients are interpolated from the 1° by 1° grid to a 2° by 2.5° grid to be consistent with model output. Such an interpolation may smooth out some spatial characteristics from observations."

4. I find it confusing that the paper is build around the seven CMIP5 models with interactive dust emissions to examine their performance to simulate DOD. But when presenting and describing the projections, the reconstructed ones based on the 16 models are considered. I understand and agree with the authors in the reasons to include more models, but then I would have expected that when examining the model performance (both climatology and interannual variability) these reconstruction (from the 16 models) also would be considered in order to be able to draw any conclusion from their projections. How good do these reconstructed projections (16 models) perform when compared with observations in present conditions? Sure, outputs of figure 9 and S8 are similar, but are they for the same reasons? Unfortunately analysis in figure 6 cannot be reproduced for the 16 models. Maybe it would make more sense to base results with respect to reconstructed projections in section 3.3 on figure S8 and move current figure 9 to the supplement (basically swaooing as it is now) and then build on how these results *are also seen (or not) in the 16 models.*

We use seven CMIP5 model with interactive dust emission scheme because we would like to examine the relationship between DOD variations and local controlling factors, while in models with offline dust these connections are lost. We added lines 210-212 to better explain this. The purpose of using variables from 16 CMIP5 models for the future projection is to include as much information (i.e., more model output) about projected change of the controlling factors as possible.

We agree with the reviewer that it is better to show the future projection by the regression model and output from seven CMIP5 models in Figure 9 first and then discuss results from 16-model output later in Figure S7 in the Supplement. We followed the advice to switch the figures and modified text accordingly (lines 522-527, 549-557).

Here we also examine the climatology and interannual variations of reconstructed DOD (using 7-model output). The following figure shows the pattern correlation between MODIS DOD and reconstructed DOD using 7-model output and regression coefficients from observations. Figure R1a shows the pattern correlations between the climatologist of reconstructed DOD (regDOD) and MODIS DOD for 2004-2016 over 9 regions. The pattern correlations are very high, because the constant value in the regression model (i.e., *d* in the equation $regDOD = a \times Precipitation + b \times Wind + c \times Bareness + d$) contains information from MODIS DOD, i.e., has a pattern similar to observed climatology.

We also show the anomalies of the reconstructed DOD where the influence of the constant value is largely removed. Figs. R1b-c show pattern correlations between MODIS DOD and regDOD for the differences of DOD between 2010-2016 and 1861-2005 (Fig. R1b) and between 2010-2016 and 2004-2016 (Fig. R1c). The latter (Fig. R1c) shows slightly better pattern correlations than the former (less green boxes) since the historical condition (1861-2005) is not exactly comparable with the 2004-2016 climatology. Fig. R1d shows the pattern correlation of MODIS and regDOD for the differences of DOD between 2010-2016 and 2004-2009. The pattern correlations are similar to Fig. R1c because relatively short time periods are used (7 years for the 2010-2016 mean and 6 years for the 2004-2009 mean) and values can be largely influenced by interannual variations of the controlling factors in the CMIP5 models.

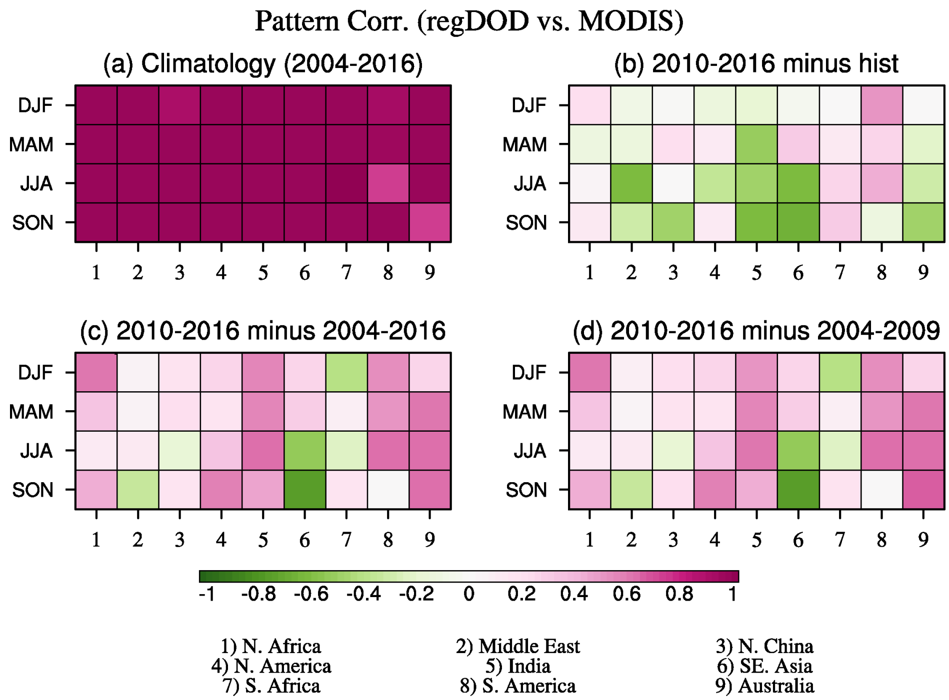


Figure R1. Pattern correlations between MODIS DOD and reconstructed DOD (regDOD) that used output from seven CMIP5 models and observed regression coefficients for (a) 2004-2016 DOD climatology, the differences of DOD (b) between 2010-2016 and historical run, (c) between 2010-2016 and 2004-2016, (d) between 2010-2016 and 2004-2009 over nine regions. MODIS DOD anomaly during 2010-2016 (with reference to the 2004-2016 climatology) is used in calculating pattern correlations in both (b) and (c).

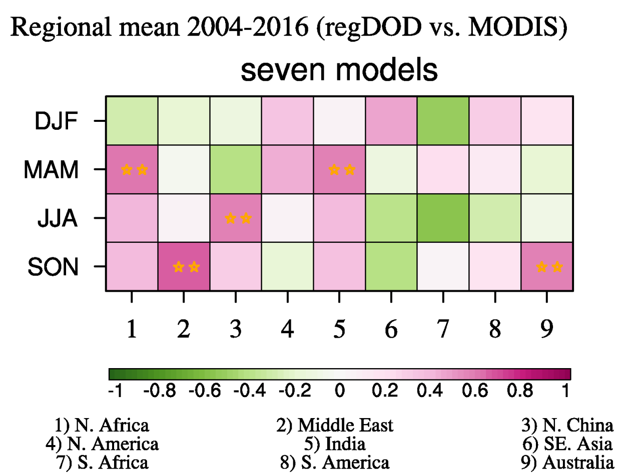


Figure R2. Correlations of regional averaged time series over nine regions between MODIS DOD and reconstructed DOD that used output from seven CMIP5 models and observed regression coefficients. Correlations significant at the 90% confidence level are marked by a star and significance at the 95% confidence level by two stars.

The correlations of regional averaged time series (2004-2016) between MODIS DOD and reconstructed DOD that used 7-model output and regression coefficients from observations are shown in Figure R2. As we mentioned in the paper, CMIP5 models are not expected to capture the interannual variations of the controlling factors, so we would not expect that the reconstructed DOD using CMIP5 output to capture the interannual variations of DOD, either. However, the variations of DOD over Africa in MAM, the Middle East in SON, India in MAM, and Australia in SON are to some extent captured by the regression model (Fig. R2). When we use observed controlling factors to reconstruct DOD (section 2.4.2), interannual variations during the present day is largely captured (Table S1).

The outputs of old Figs. 9 (from 16 models) and S8 (from 7 models) are similar because the projected changes of precipitation, surface wind speed, and bareness from 16-model ensemble mean (Fig. R3) show some features similar to 7-model ensemble mean (Fig. 8). We clarified this in the updated text (lines 549-557) and also added Fig. R3 to the supplement.

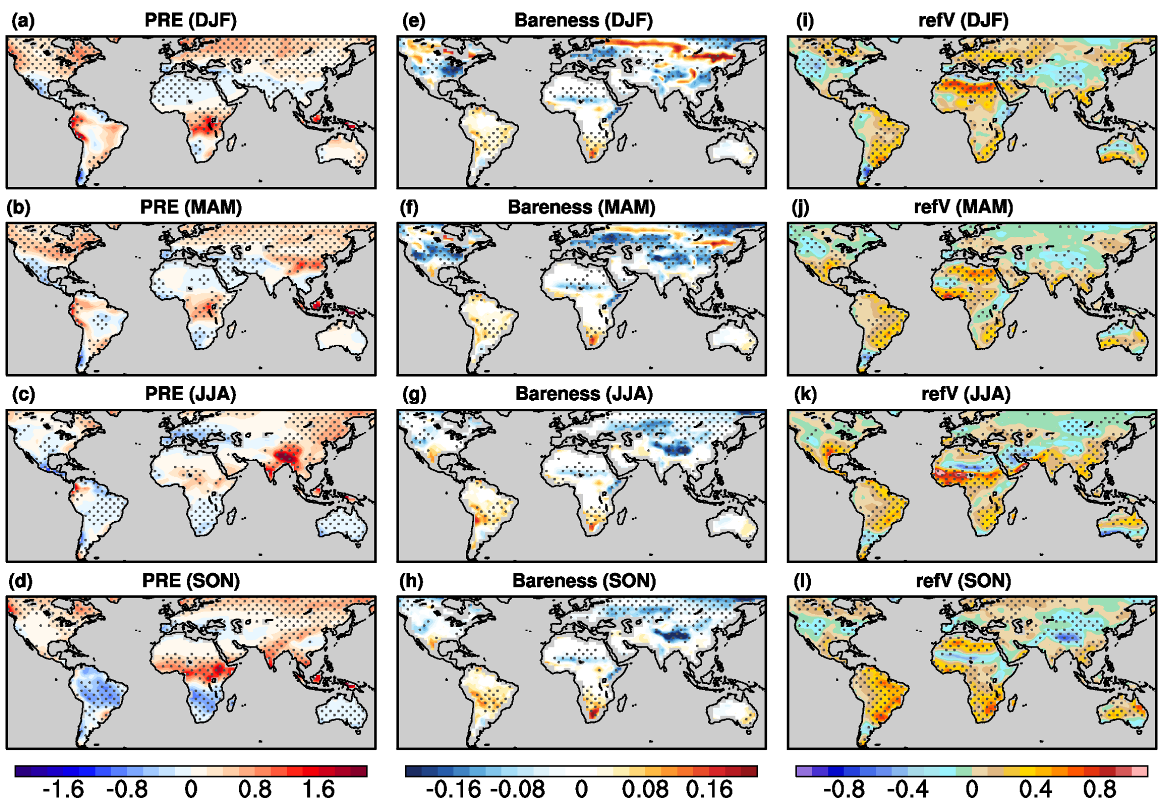


Figure R3. Projected difference of (a)-(d) precipitation (mm day^{-1}), (e)-(h) bareness, and (i)-(l) 10 m wind (m s^{-1}) between the late half of the 21st century (2051-2100; RCP 8.5 scenario) and historical level (1861-2005) from multi-model mean of 16 CMIP5 models. Areas with sign agreement among the models reaches 62.5% (i.e., at least ten out of 16 models have the same sign as the multi-model mean) are dotted.

Specific Comments:
Page 5, lines 73-75: See general comment above, I would suggest reformulating the statement.

We modified those lines to: "A comprehensive evaluation of the climatology and interannual variation of global dust optical depth (DOD) in CMIP5 models will provide insights into models' capability of simulating the integrated aerosol extinction due to dust. DOD is also one of the key variables that determine radative forcing of dust to the climate system.". And also added discussion in lines 604-609 to acknowledge that DOD dose not reflect the vertical distribution of dust aerosols and more variables (such as surface dust concentration, emission, deposition, vertical distributions) are need to provide a whole picture of dust simulation in CMIP5 models.

Page 6, lines 98-100: Given the importance of DOD in this study I suggest you briefly describe the method how DOD was derived from AOD and specify the modifications you applied to adapt the method to collection 6.

Lines 102-114 are added to describe how DOD is derived and adapted to collection 6.

Page 7, line 134: Table 2 is referenced without any reference to Table 1. At present Table 1 corresponds to information on the models used in this study which is addressed in section 2.3. Tables should be arranged according to the order they are referenced in the text.

We actually referred Table 1 in line 78 when introducing the seven models used in this study.

Page 8, lines 138-143: Surface wind speed, bareness and precipitation are defined as controlling factors without providing any evidence or explanation why these parameters. However in lines 321-331 the authors explain why these parameters have been selected. I suggest moving these lines forward to section 2.2.

We follow the advice to move lines 321-331 to section 2.2 (now lines 167-176).

Page 8, line 156: Remove PRECL.

Here we refer to the precipitation data from PRECL and so will keep "PRECL precipitation".

Page 9, lines 164-182: A reference to (current) Table 1 should be made in this section. In addition, information on the 16 models used in the future projections needs to be provided.

We added "Table 1" in line 207. We also modified lines 286-288 to clarify that information of 16 CMIP5 models can be found from the Supplementary Table S1 of Pu and Ginoux (2017).

Page 10, line 188: Provide a reference for the mass extinction efficiency used.

The mass extinction efficiency used here is from Ginoux et al. (2012a) as mentioned in line 231. We also added discussion on this variable in lines 237-241.

Page 10, lines 191-201: The authors illustrate the difference between the derived DOD and simulated one from one of the seven CMIP5 models with interactive dust emissions. It seems arbitrary why this model is used and not any other of the seven models? Is the intention of these lines to validate the derived DOD and therefore the chosen method? If that's the case then a more thorough validation should be done such as comparing the derived model mean DOD from all 16 models to the model mean from the seven CMIP5 models. Otherwise I don't see the point of having these analysis.

In these lines we compare the derived DOD versus model calculated DOD in GFDL-CM3 to valid the method we used to derive DOD (i.e., Eq. e). We did not use this analysis to select models. We chose seven models with interactive dust emission schemes to examine DOD climatology and interannual variations because DOD in these models are influenced by environmental factors and the can be compared with observations, while in models with offline dust, these connections do not exist in the models.

We used GFDL-CM3 as an example to validate the DOD derivation because it's the only model among the seven that we can access model calculated DOD. We modified lines 241-253 to better present the analysis.

Page 11, lines 216-220: What period is considered in this analysis, same as observations, ie 2004-2016?

Yes. We modified line 271 to clarify this.

Page 11, line 226: Please provide some information on these 16 models, which models are they? Are the seven model with interactive dust emission part of these 16 models? Do they have prescribed emissions? A similar table as Table 1 should be included with relevant information of these 16 models.

Seven models are part of these 16 models. We modified line 286-288 to clarify that models information and dust emission schemes can be found from Supplementary Table S1 of Pu and Ginoux (2017).

Page 13, lines 258-260: How do the authors explain the shift to the north in the DOD by HadGEM2?

In lines 258-260 (original version) we referred the multi-model mean shown in Fig. 2b: "The peak around 19° N in North Africa and Middle East is well captured by the multi-model mean, although the magnitude is slightly underestimated." The overestimation of DOD around 28° N in the HadGeM2 model may be caused by its overestimation of DOD over the Middle East and India in summer (Figs. 3b, e).

Page 13, line 269: remove "than other seasons".

Done.

Page 13, line 270: add "by the model mean" after "captured".

Done.

Page 13, lines 269-271: Since individual models are illustrated, authors should not only focus on the multi model mean but also on the individual models and their differences with respect to the multi model mean and the observations. For instance, MIROC and

GFDL do not present the observed variability, in particular over N. America and India and they also present a different variability than the other models over northern China, with the peak closer to the observed one.

We revised lines 334-337, 345-347, 358-359 to add discussion on a few models' performance over North America, northern China, and Australia.

Page 14, lines 276-277: The MODIS DOD peak in Australia is hardly seen.

We have scaled MODIS DOD over Australia ten times in Fig. 3 and modified figure caption accordingly to better display the seasonal cycle of DOD.

Page 15-16, lines 319-321: Which reconstructed DOD is used here? is it the one considering observed regression coefficients and simulated controlling factors? Or is it the one using simulated regression coefficients derived from model DOD and model controlling factors? Also, are only the seven CMIP5 models with interactive dust emission considered? The authors should be more specific which reconstruction they refer. Also, couldn't the correlation based on reconstructed DOD be integrated in the figure as an additional column?

The reconstructed DOD used observed regression coefficients and observed controlling factors. We modified lines 398, 412-414 to clarify this. We actually considered adding a column to Fig. 5 to show the correlations between MODIS DOD and reconstructed DOD (see Fig. R4 below) in the early version of the paper. However, since the reconstructed DOD here used observed controlling factors, which make it slightly "unfair" to compare the results with those from CMIP5 DOD, we decide to present the results separately in Table 2.

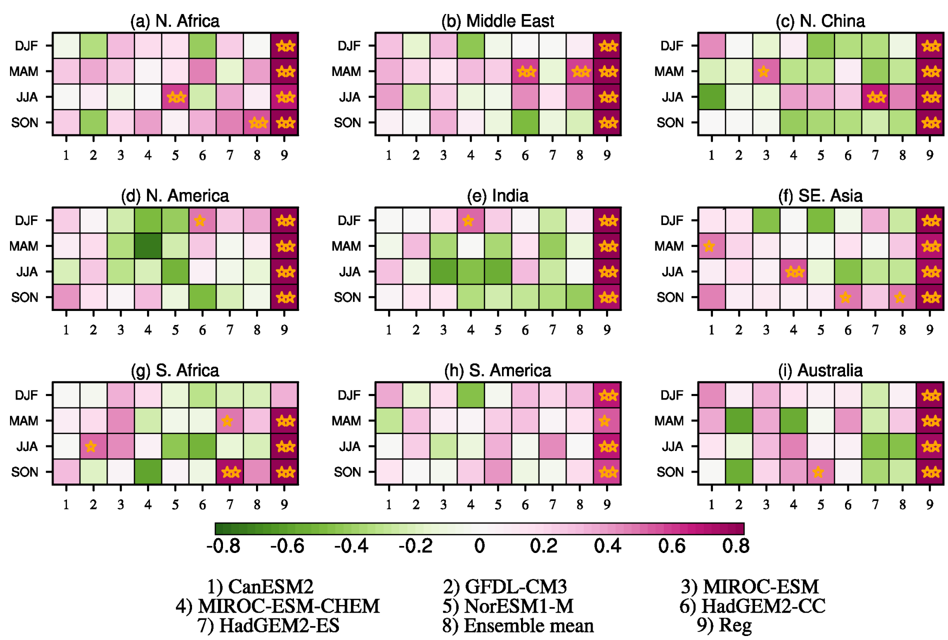


Figure R4. Correlations (color) between regional averaged time series from CMIP5 DOD and MODIS DOD from 2004 to 2016 for four seasons. Numbers in the X-axis denotes each model (1-7), multi-model mean (8), reconstructed DOD (9). Correlations significant at the 90% confidence level are marked by a star and significance at the 95% confidence level by two stars.

Page 16, lines 321-332: move these lines to section 2. See general comment.

Done.

Page 18, lines 378-380: I have difficulties seeing the similarities in North Africa and the Middle east between the MODIS and CMIP5 regression coefficients pointed out by the authors. I actually see more the differences between both regressions in both regions. I would suggest the authors review the analysis in these lines.

Lines 458-462 are modified clarity this: "In JJA, the influences of precipitation and bareness over the eastern Arabian Peninsula in the multi-model mean (Fig. 6g) also show some similarity to observation (Fig. 6c), although an underestimation of the influence from bareness and an overestimation of precipitation are still there. "

Page 22, lines 470-473: On which results are the authors basing this statement. I suggest specifying.

We added "in the present-day (Fig. 6)" after "Multi-model mean also overestimates the connection between DOD and precipitation and surface wind and underestimates the influence of bareness" to specify this argument. In section 3.2 we compared the multiple linear regression coefficients from CMIP5 models with those from the observations (Fig. 6) and found multi-model mean overestimates the connection between DOD and precipitation and surface wind while underestimates the influence of bareness.

Pages 22-23, lines 465-490: These lines would fit better in the discussion section.

We prefer to discuss the uncertainties of CMIP5 and regression model projections right after showing the results of the two methods (Figs. 7-10) in section 3.3. In section 4, more general issues such as including other variables from CMIP5 models to examine model performance, studies on future dust projection, and the implication of the regression model, are discussed.

Page 24, line 522-524: The statement seems something that would fit better in the conclusion section. Consider moving it.

This is not the key conclusion of the paper, so we prefer to keep it in the discussion.

Page 25, line 546: Suggest replacing "quite well" with something more academic.

We modified the line to: "In JJA, the simulated zonal mean DOD from multi-model mean largely resembles MODIS DOD".

Page 27, line 583: In which way are similarities between both projections "informative"? What information do they provide.

See our detailed reply to Comment #2. We removed "which may be informative", and modified the sentence to: "we find some similarities between the two, which adds to the confidence of projected DOD change in these regions, for instance..."

Interactive comment on "How reliable are CMIP5 models in simulating dust optical depth?" by Bing Pu and Paul Ginoux
Anonymous Referee #2

We thank the reviewer for very helpful comments. We reply to your comment (in Italic) below.

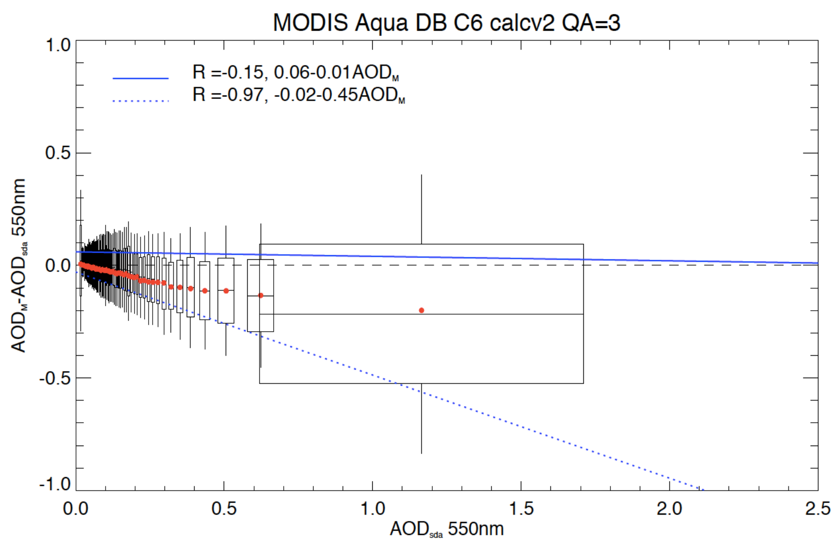
The article presents an in-depth analysis of the CMIP5 models ability to reproduce the dust optical depth (DOD), considering both seasonal and inter-annual variability, as well as the driving factors behind those DOD levels. The observational data used are DOD over land derived from MODIS Terra-aqua data; bareness derived from AVHRR; 10m wind speed from ERA-Interim reanalysis; and precipitation from PRECL. The analysis of the driving factors is performed by regressing the observed DOD from MODIS over land to the observed/reanalyzed driving factors. The analysis is then extended to future climate scenarios (RCP8.5) using both the CMIP5 models' dust outputs and the regression based on present day observed relationships between DOD and the driving factors.
The main results/conclusions are: 1) Models behave better over the NH large dust sources. 2) Models do not reproduce interannual variability. 3) The constraints from bareness in models are underestimated and the influences of wind speed and precipitation are overestimated. 4) A corrected projection of DOD based on the regression model is proposed. There are some similarities between the projections and the corrected projections.
The paper is very interesting, includes novelties and deserves publication. However, I have several doubts and comments that need clarification and further discussion.

General comments
1) DOD from MODIS: It is not clear what the DOD derived from MODIS refers to. Is it total dust optical depth or coarse dust optical depth? I understand that it refers to the total dust optical depth (fine and coarse) but I was confused when the product was compared to the coarse (O'Neill) product from AERONET. Can you please explain better the derivation of DOD from AOD in the paper? Given the importance of the dataset for the paper I feel it is not enough to refer the reader to other publications. Also, can you provide an estimation of the uncertainty of this product?

We added lines 102-114 to better explain how DOD is derived. It is coarse dust optical depth. The formula is derived from the work of Anderson et al. (2005). Uncertainty of this product is added to the supplementary information as shown below. We also modified lines 119-122 to include these information.

Figures R1-2 compares aerosol optical depth (AOD) between MODIS and AErosol RObotic NETwork (AERONET) sites data (top), and between MODIS DOD and AERONET coarse mode aerosol optical depth (COD; bottom). AERONET COD is processed by the Spectral Deconvolution Algorithm (O'Neill et al., 2003). We used an evaluation method following Levy et al. (2003; their Fig. 11) for AOD and COD errors. The AERONET Level 2 (quality assured) 10 minutes AOD and COD (500 nm) are extracted for Aqua equatorial crossing time (1:30 PM) and Terra equatorial crossing time (10:30 AM) plus or minus 30 minutes, and are considered if there is at least 2 measurements per day and there should be at least 100 days with data. We select AERONET sites within a spatial radius of 15 km of MODIS measurement. 883 AERONET sites are used. Total number of valid data is about 35,747. In box-whisker plots (e.g., Fig. R1), all collocated MODIS and AERONET data are grouped into bins of 500 measurements. The last bin will contain a larger number of values corresponding to the remaining of the division.

As shown in Fig. R1, MODIS slightly underestimated Aqua AOD and DOD for most of the AOD and DOD ranges. Compared to AERONET station data, Aqua AOD is underestimated, and DOD largely inherits this error. For Aqua DOD around 0.50, the median error is around 0.08, with estimated errors ranging from -0.29 to 0.16. Terra DOD is better than Aqua DOD in terms of the median of errors (Fig. R2 bottom vs. Fig R1 bottom). The median error for Terra DOD around 0.50 is very close to zero, with estimated errors ranging from -0.23 to 0.25.



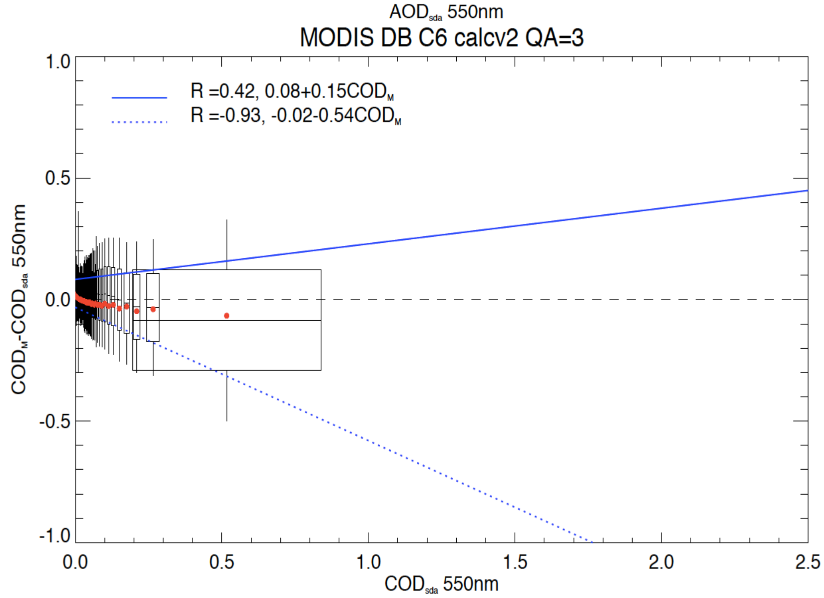


Figure R1. Comparison between grouped Aqua AOD error (i.e., the differences between MODIS AOD and AERONET AOD versus AERONET AOD, top), and grouped coarse mode aerosol optical depth (COD) error (i.e., the differences between MODIS DOD and AERONET COD versus AERONET COD, bottom). For each box-whisker, its width is 1σ of the AOD (COD) bin, while its height, whiskers, middle line and red dots are the 1σ, 2σ, mean, and median of AOD (COD) error, respectively. The envelope of estimated errors are blue and the one-one line (zero error) is dashed black.

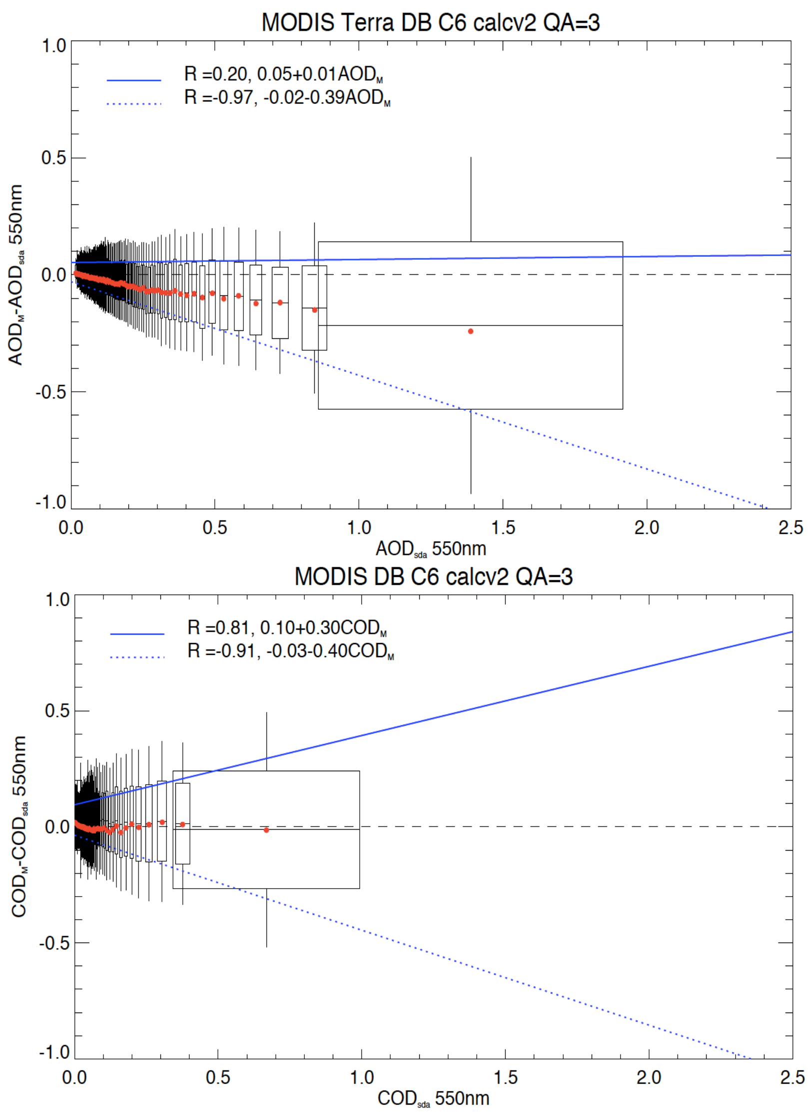


Figure R2. Same as Fig. R1 but for Terra DOD.

The confidence of satellite data over the different regions is assessed by comparison with AERONET (few stations, low spatial coverage), CALIOP, and considering the number of days with available DOD per season. Results show that while in Africa, South America, Middle East and some Asian regions confidence seems to be high, for some regions in

Asia/North America it largely depends on the season. In my view, the strength or confidence on the DOD data by region should be considered when discussing: the modelled DOD evaluation at the regional level, the regression method projections and discrepancies with CMIP5 models.

Major uncertainties we found in terms of days of coverage and comparison with AERONET and CALIOP are: 1) low coverage over northern China and Southeastern Asia in JJA; 2) DOD is slightly higher than COD from AERONET over Arabian Peninsula in DJF and SON; 3) DOD is lower than CALIOP COD over northern India in MAM. We added lines 159-164, 338-343 to discuss the uncertainties associated with DOD.

2) DOD from CMIP5 models: The authors compare the DOD derived from the selected CMIP5 models using Eq. (2). Using a value of 0.6 everywhere and for every model is an important simplification as it depends on model-dependent assumptions on size distribution and other issues such as the size range considered. While 0.6 may be a reasonable value for GFDL-CM3, how can we be sure it is ok for other models? Is there any other model for which you could compare this assumption in addition to GFDL-CM3.

We agree that using 0.6 for all models is a simplification and adds uncertainties to our analysis. We modified text to address this issue, e.g., added lines 237-241: "Applying the same mass extinction efficiency everywhere and to all the CMIP5 model output used here is a simplification, as different models may have quite different mass extinction efficiency. For instance, *e* can range from 0.25 to 1.28 $m^2\ g^{-1}$ in AEROCOM models, with a multi-model medium of 0.72 $m^2\ g^{-1}$ (Huneeus et al., 2011)." and lines 243-244: "A full validation of this method will require modeled DOD from all the other CMIP5 models, which are currently not available."

3) Clear sky vs all sky values: While the authors have made an effort to gather the largest possible amount of DOD data by using both Aqua and Terra, the results of the comparison between MODIS DOD and model DOD may be quite affected by the use of all sky values from the models instead of clear sky values. Can you at least quantify this effect by for example using clear sky DOD from GFDL-CM3? How large is this effect? This may be potentially important in areas with seasonal clouds and precipitation. Could this be one of the reasons for the strong disagreement in some regions?

As the reviewer pointed out, MODIS AOD removed pixels contaminated by cloud, and therefore AOD (and DOD) is retrieved toward a clear-sky condition. On the other hand, the derived (or modeled) DOD in CMIP5 models does not have any cloud-screening process and therefore is under an all-sky condition. The inconsistence between the two may add some uncertainties in regions with more cloud coverage/amount, such as the central U.S., northern China, southeastern Asia, and northern South America, but less so over North Africa, South Africa, the middle East, Australia, India (except JJA), and Australia (Figure R3).

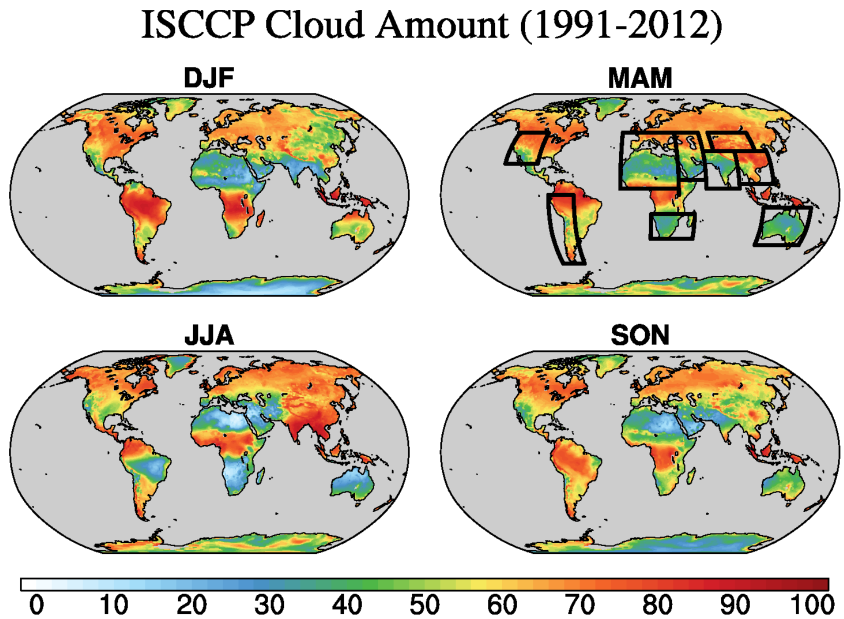


Figure R3. Total cloud amount (%) from the International Satellite Cloud Climatology Project (ISCCP) averaged over 1991-2012. Black boxes denote the nine averaging regions.

In GFDL-CM3 model, DOD at each grid point is calculated under all-sky condition and model does not have output of clear-sky DOD. We compared DOD from CALIOP level 3 data under all-sky condition and cloud-free (i.e., clear sky) condition (Figure R4). The differences of global mean DOD over land under all-sky and clear-sky conditions range from -0.003 in MAM to 0.001 in DJF. The differences are larger (> ±0.05) over cloudy regions in MAM and JJA, particularly over Guinea coast in West Africa, northern China, southeastern Asia, India (Fig. R4, bottom). The differences are largely due to the fact that much less samples are collected to produce cloud-free DOD over these cloudy regions (not shown). The disagreement between MODIS DOD and CMIP5 DOD in the above regions (i.e., Guinea coast in West Africa, northern China, southeastern Asia, India) is not particularly higher than other regions (e.g., Fig. 4).

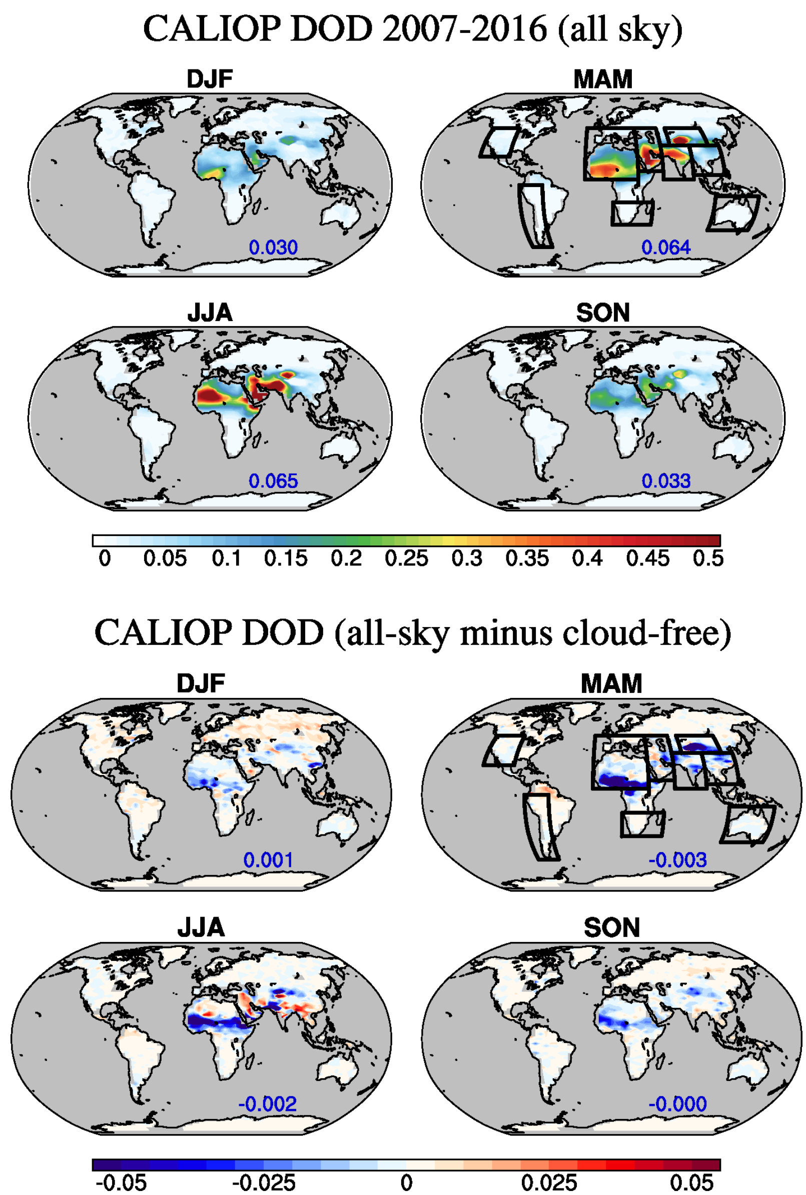


Figure R4. Climatology (2007-2016) of CALIOP DOD under all-sky condition (top) and the differences between all-sky and cloud-free conditions (bottom). Blue numbers denote global mean DOD over land.

4) Interannual variability: One of the findings of this study is that the interannual DOD variation is not very well captured by the CMIP5 models. It is stated that "models probably misrepresented these [controlling factor] relationships, in addition to their incapacity of capturing the interannual variations of individual controlling factors". Because of their nature, CMIP5 models cannot (and are not meant to) represent year-to-year variations of the driving factors in such a short time period. Therefore, the first part of the statement is just speculative, i.e., one cannot know whether the relationships are misrepresented from that analysis alone. I strongly believe that this part should be better discussed both in the results section and the conclusions. I also believe that the *comparison between CMIP5 model output and observations in Figures S4 to S6 is not needed. Isn't it obvious that CMIP5 models cannot represent year-to-year variations of each season in a 12-year period?*

The reviewer questioned our argument "...models probably misrepresented these relationships..." following the discussion on Figure 5 and Table S1. We did not intend to make any conclusion at that line, but to bring up a question. Later in Figure 6 we examined the connections between CMIP5 DOD and controlling factors. We revised line 425 to avoid misunderstanding: "... models may misrepresent these relationships, in addition to their incapacity of capturing the interannual variations of individual controlling factors in general". We also followed reviewer's suggestion to remove Figures S4-6 and modified text accordingly (line 427, lines 669-674).

5) The role of surface bareness: one of the important conclusions of the study is that "constraints from surface bareness are largely underestimated while the influences of surface wind and precipitations are overestimated". I have a few doubts/comments on this:

a. How can you know that the constraint from surface bareness is largely underestimated? Given your methodology, couldn't it be that the constraint of surface bareness is correct in absolute terms but the effect of precipitation (through soil humidity) is overestimated? This should be clarified.

It is possible that the magnitude of one controlling factor in the model is closer to the observation while the others are systematically underestimated/overestimated. So we standardized each controlling factor before regression. Therefore, the differences due to their absolute values are removed. The regression coefficients thus reflect how the interannual variations of each factor may contribute to the variations of DOD.

b. While I think that the methodology is sound, it is not clear to me how year-to-year variations of around 2-3 % in LAI (Figure S7) can have such an impact in the interannual variability of dust in Northern Africa. Because this conclusion has important implications, could you further discuss this point? What would be the physical mechanisms that could explain this?

First of all, we'd like to clarify that Fig. S7 shows bareness instead of LAI. Year-to-year variations of LAI are above 10% over the Sahel and parts of North Africa (Figure R5, right column). Bareness, or LAI, is a key non-erodible factor that can prevent wind erosion. The reason bareness shows a stronger influence on the interannual variations of DOD than the other two factors (precipitation and surface wind speed) is because its variations are more consistent with DOD changes. Here we show an example. We select an area over the Sahel (10°-16°N, 0°-25°E) where bareness is the dominant controlling factor in MAM based on multiple liner regression (Figure R6a). As shown in Figure R6b, the interannual variation of bareness (orange) is more consistent with DOD (black) variations than surface wind speed (green) or precipitation (purple) in the region. The correlation between DOD and standardized bareness is 0.61 ($p=0.03$), also higher than the correlations between DOD and precipitation (-0.46) or between DOD and surface wind (0.55).

We also examined multiple-linear regression using LAI from GLASS during 2004-2014 (Xiao et al. 2014). GLASS LAI is derived from MODIS products for years after 2001. The results using GLASS LAI are very similar to what we obtained from AVHRR LAI (Figure R7).

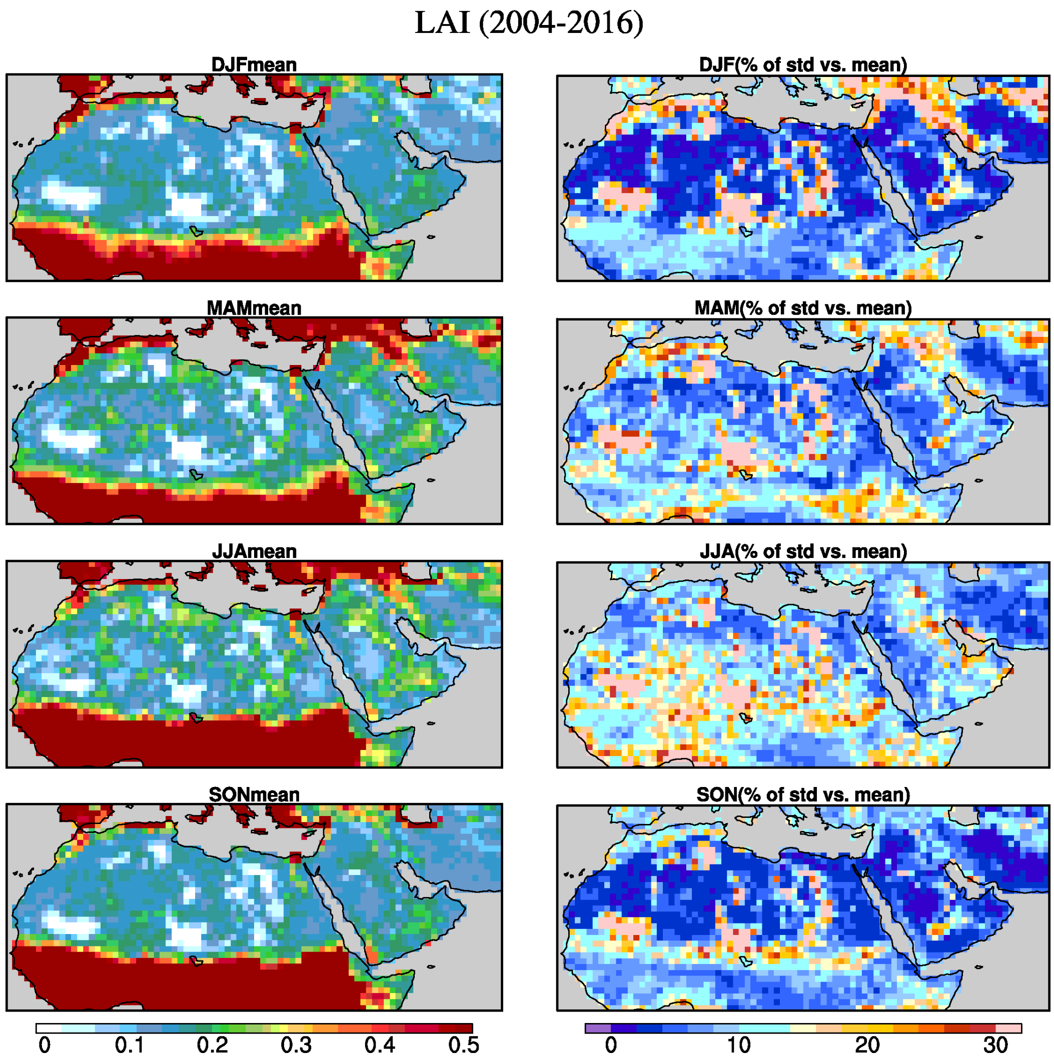


Figure R5. Seasonal mean of LAI averaged over 2004-2016 (m^2/m^2; left) and ratio (%) of standard deviation of LAI to seasonal mean LAI (right) over North Africa and the Arabian Peninsula from AVHRR during 2004-2016.

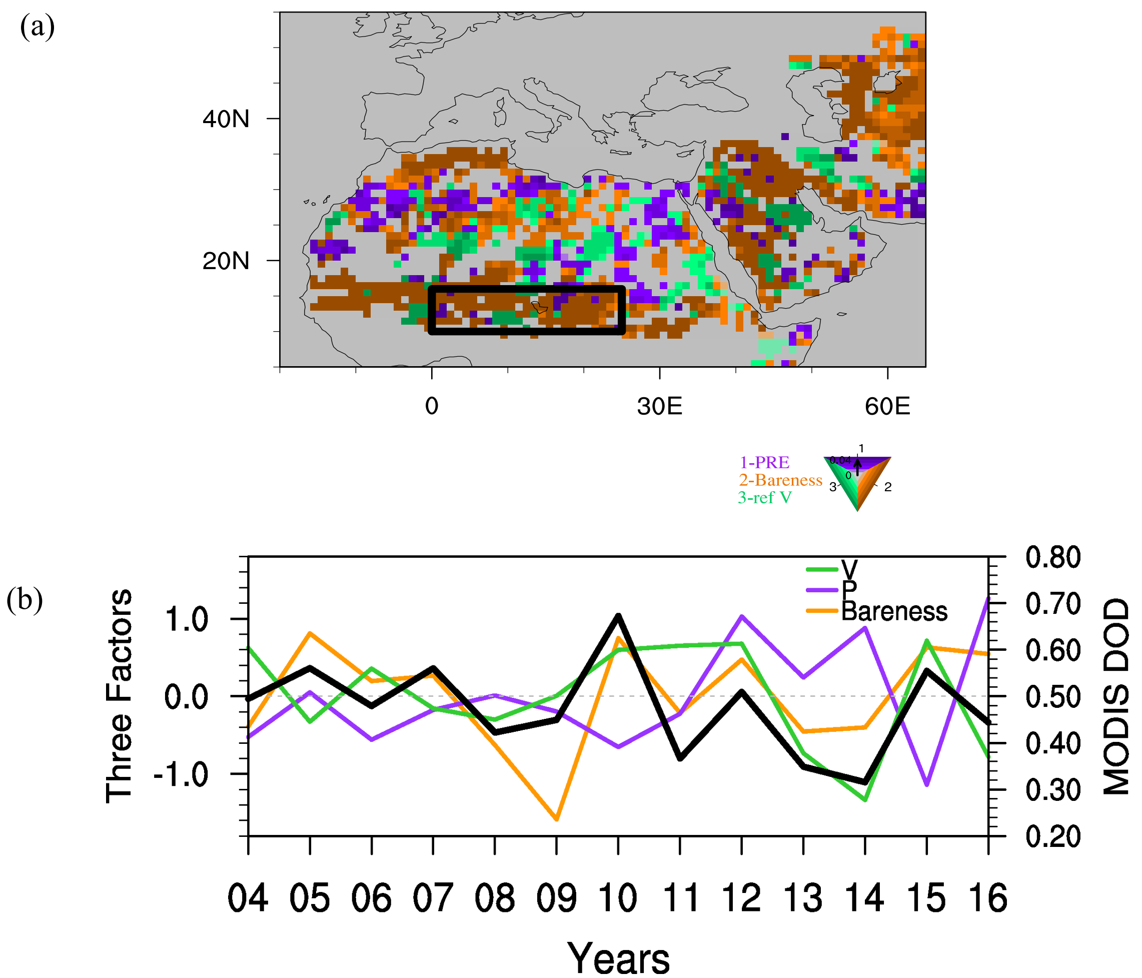


Figure R6. (a) Same as Fig. 6b but on a 1° by 1° grid for North Africa and the Middle East. Black box indicate an averaging area between 10°-16°N and 0°-25°E. (b) Time series of standardized controlling factors of bareness (orange), surface wind (green), precipitation (purple) and MODIS DOD (black) averaged over the area shown in (a).

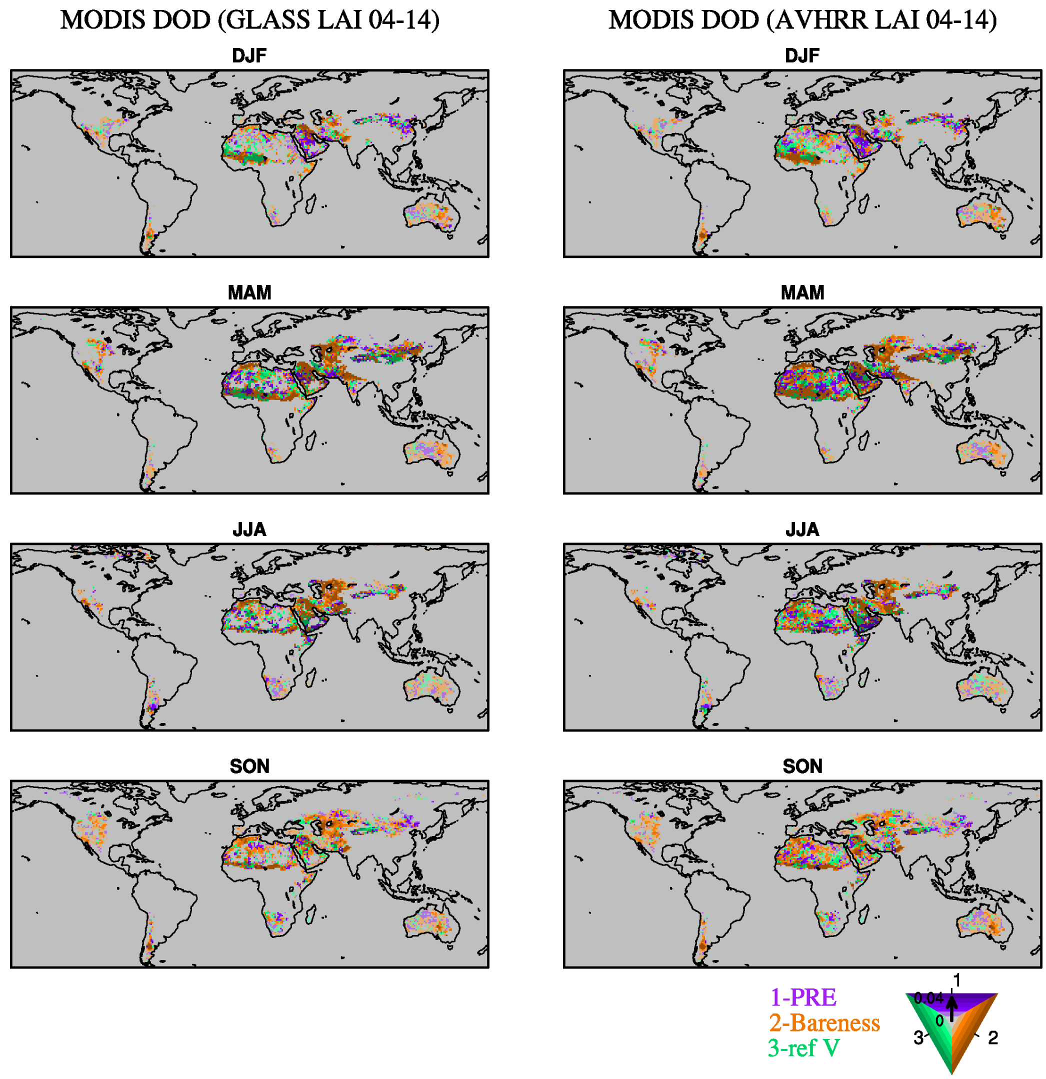


Figure R7. Regression coefficients calculated by regressing DOD in each season onto standardized precipitation (purple), bareness (orange), and surface wind speed (green) from 2004 to 2014. Coefficients obtained using MODIS DOD and observed controlling factors. Plots in the left used LAI from the GLASS, while on the right used LAI from the AVHRR. All the other variables are the same. The color of the shading denotes the largest coefficient in absolute value among the three, while the saturation of the color shows the magnitude of the coefficient (from 0 to 0.04). All regression coefficients regardless of their statistical significance are shown. Missing values are shaded in grey. To highlight coefficients near dust source regions, a mask of LAI ≤ 0.5 is applied.

Can you provide the same figure (S7) but for the model derived bareness (both present day and future projections)? *How well do the models compare with the observed range of variability of the LAI in arid regions (the Sahara for example)?*

Modeled climatology of bareness is higher over North Africa (Figure R8) than that in the AVHRR, and the standard deviation is lower over northern North Africa but much higher over the Sahel.

Bareness (CMIP5) 2004-2016

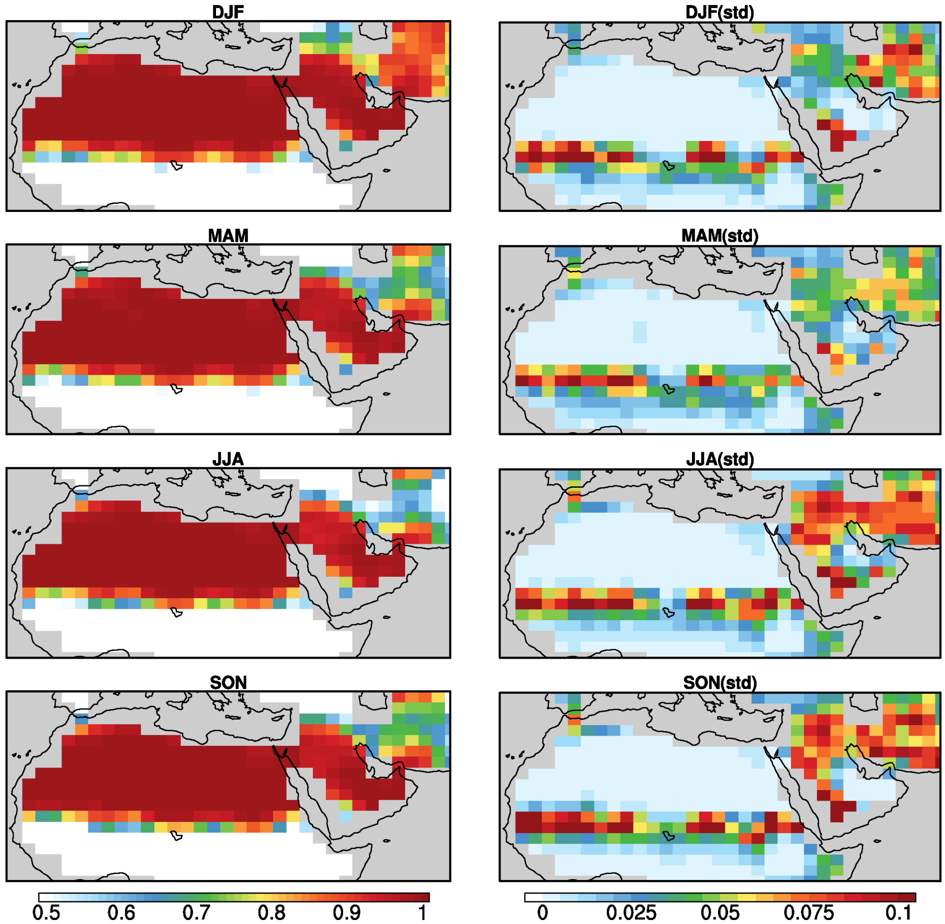


Figure R8. Seasonal mean (left) and standard deviation (right) of bareness over North Africa and the Arabian Peninsula from CMIP5 7-model ensemble mean during 2004-2016.

6) Regression method projections vs. CMIP5 projections: The regression method used to derive DOD in future scenarios is based upon 16 CMIP5 model variables (surface wind speed, bareness and precipitation) and compared to dynamical projections of only 7 CMIP5 models (those with online dust schemes). Partly, differences in future trends might come by differences in driving variables. You state [line 439] that projected DOD changes using the full sample or only 7 models are very similar. If so, why not using the same 7 model outputs as drivers? This would enhance consistency. Finally, why the similarities between the two approaches in some regions may be informative?

The purpose of using variables from 16 CMIP5 models for the future projection is to include as much information (i.e., more model output) about projected change of the controlling factors as possible. As the reviewer pointed that, different number of CMIP5 models used for the regression model may add to the differences between CMIP5 model projected DOD and regression model projected DOD. So we follow the suggestion to show regression model projected DOD change using 7-model output in Figure 9 to keep the consistency and show results from 16-model in Figure S7.

We removed "which may be informative", and modified the sentence to: "we find some similarities between the two, which adds to the confidence of projected DOD change in these regions, for instance...". We also modified lines 513-522 to better explain approach of future projection using the regression model: "Here we also present the projected change of DOD from the regression model in Figure 9. The regression model is developed based on observed relationships between DOD and local controlling factors and can largely capture the interannual variations of DOD in the present-day climate (Table S1 in the Supplement). Assuming that the observed connection between DOD and these controlling factors do not change dramatically in the future, we can use this regression model and CMIP5-model projected change of controlling factors to project DOD variations. Compared to DOD projection from CMIP5 models, this approach utilizes additionally observational constrains and is likely to provide a more reliable future projection." Although CMIP5 models overestimate the influence of surface wind and precipitation and underestimate the role of bareness, there are some similarities between model and observations (Fig. 6; lines 450-462), which indicate that models partially capture the connection between the DOD and these controlling factors in some regions. Therefore, the projection of DOD from CMIP5 models is not completely unreliable. The similarity between CMIP5 projection and the projection from the regression model thus adds to the confidence of projected change of DOD over North Africa, the Arabian Peninsula, and northern China in some seasons.

Minor comments:

- I suggest to list multiple references to the same topic chronologically, unless there are reason to order them differently, e.g. in the introduction.

Done.

- I think the column heading "Dust emission scheme" is somewhat misleading as the given references describe the implementation of a dust emission scheme, not the scheme itself. Perhaps rewording to "Dust emission implementation" or similar would help.

We follow the suggestion to change column head to "Dust emission implementation".

- I suggest changing Eq. (1) to Bareness = exp(-LAI). Also, is there a reference for this equation?

We change Eq. (1) following the comment and added a reference.

- L. 146-147: I would normally not consider a resolution of 80km "very suitable to study the influence of wind speed on dust emission and transport on small scales". I understand the intent of this statement, but I suggest rephrasing this to avoid misunderstanding.

Thanks for the suggestion. We modified lines 185 to "We choose this analysis because of its relatively high spatial resolution".

- L. 160: I suggest to delete "relatively high" as well as "quite"

Done.

- L. 192: GFLD-CM3 should be GFDL-CM3

Done.

- Line 205: clarify which DOD is regressed onto observed values, i.e. satellite derived DOD

We changed the sentence to "by regressing MODIS DOD onto..."

- Fig. 3i: It is very hard to see the MODIS DOD pattern for Australia. Can this be improved?

We re-plotted the figure to better display MODIS DOD for Australia.

- Fig. 4 is ok, but quite dense

We updated the figure to make it look better.

- L. 311: variability instead of variations

Done.

- L. 328: wind erosion instead of "soil erosion from wind"

Done (now line 173).

- Line 431: centaury should be century

Done.

- L. 457 ff: Sometimes it is not clear if "models" refers to the CMIP5 models or projection 'models'.

We changed "models" to "regression projections" to avoid confusion.

- Figure 6. It is difficult to sort out the different elements, e.g. the strength of the regression depending on the shading intensity is not visible. I would suggest: to make a zoom per region, or to display dependencies from the 3 variables in independent maps, and to use the same resolution for MODIS and CMIP5 maps to make easier a direct comparison.

We updated Figures 6 by interpolating results from MOIDS and observed controlling factors to model grids (2° by 2.5°, Figs. 6a-d) and changed color scale from 0~0.02 to 0~0.04 to better show the shading intensity. We also zoomed in and plotted a few figures for different regions (Figures R9-R13) here. The patterns in new Fig. 6 are very similar to the old one. The connection between DOD and bareness is underestimated on the interannual time scale in CMIP5 models. On the other hand, DOD's connection with precipitation and surface wind speed are overestimated.

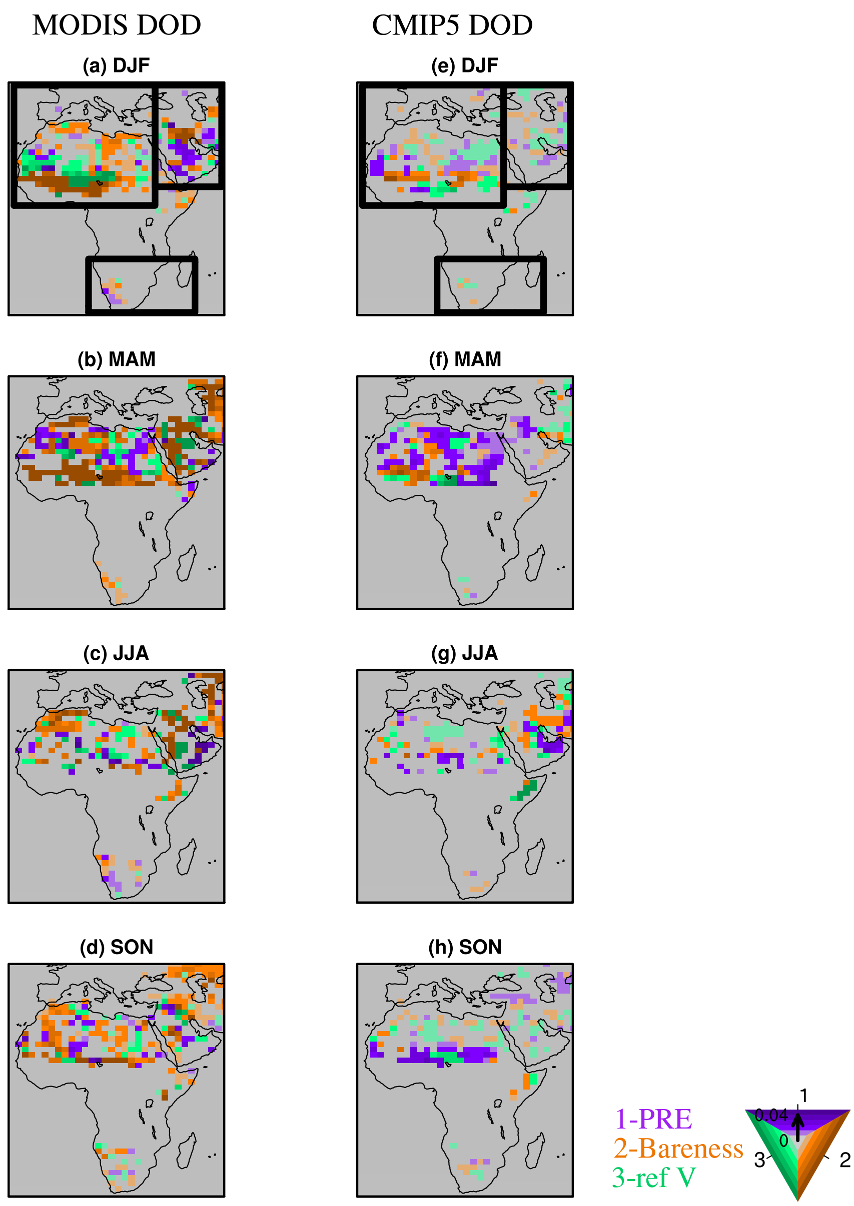


Figure R9. Same as Fig. 6 but for Africa and the Middle East. Black boxes denote the averaging regions defined in Table 2: North Africa, South Africa, and the Middle East.

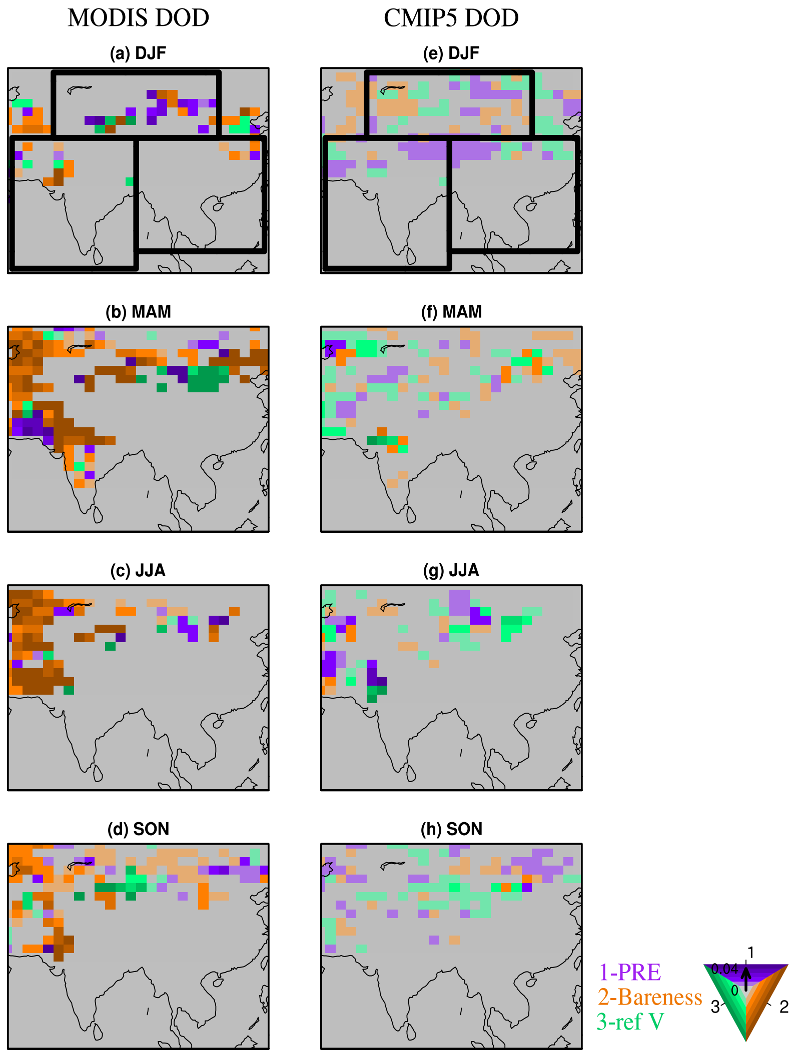


Figure R10. Same as Fig. 6 but for Asia. Black boxes denote the averaging regions defined in Table 2: northern China, India, and southeastern Asia.

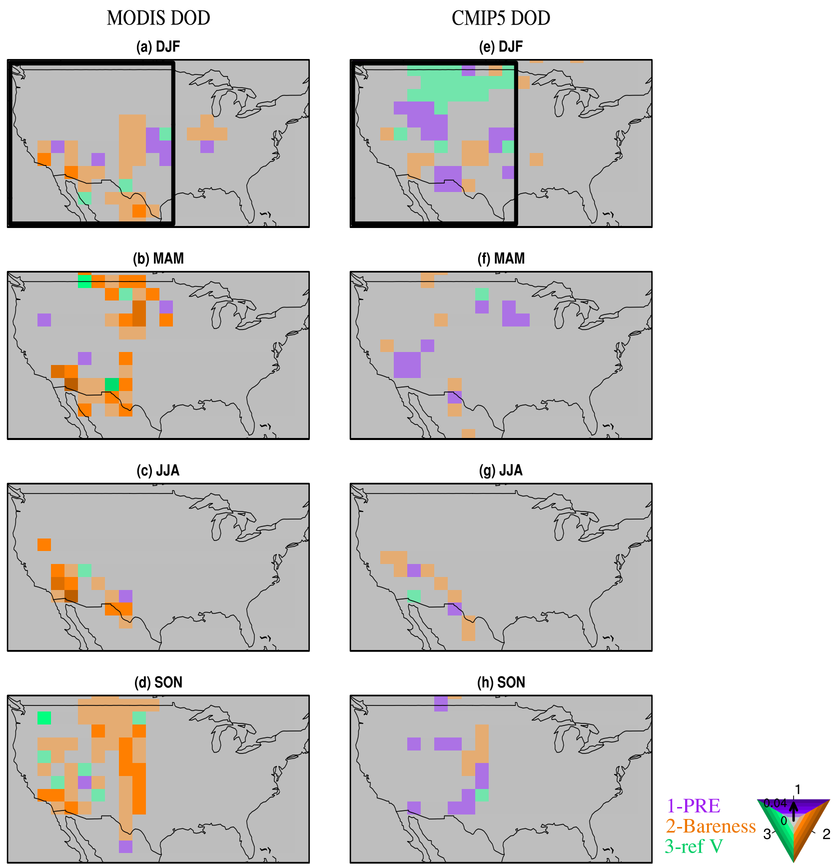


Figure R11. Same as Fig. 6 but for North America. Black boxes denote the averaging regions defined in Table 2: North America.

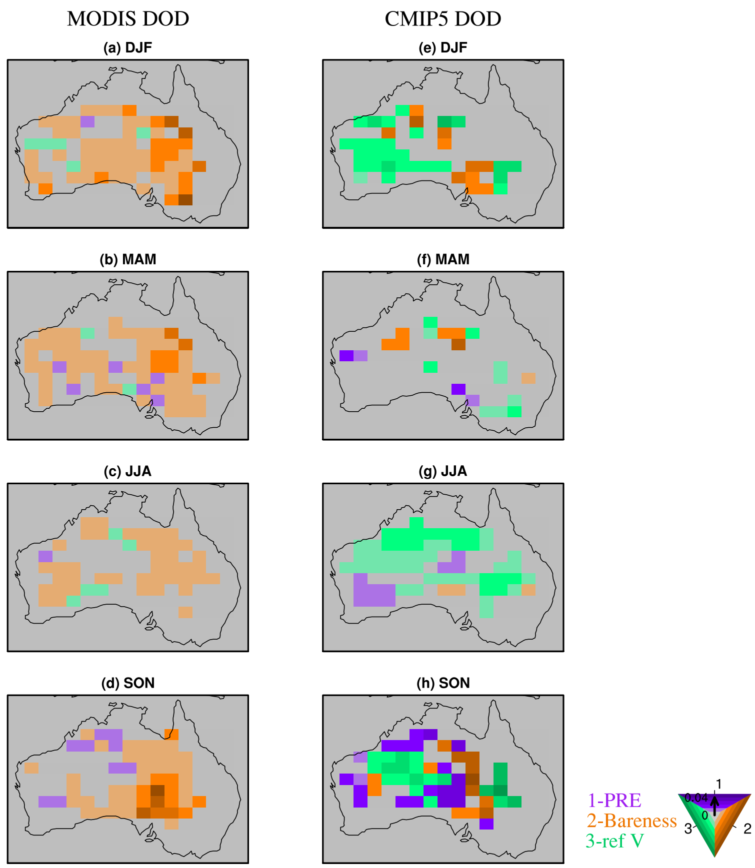


Figure R12. Same as Fig. 6 but for Australia.

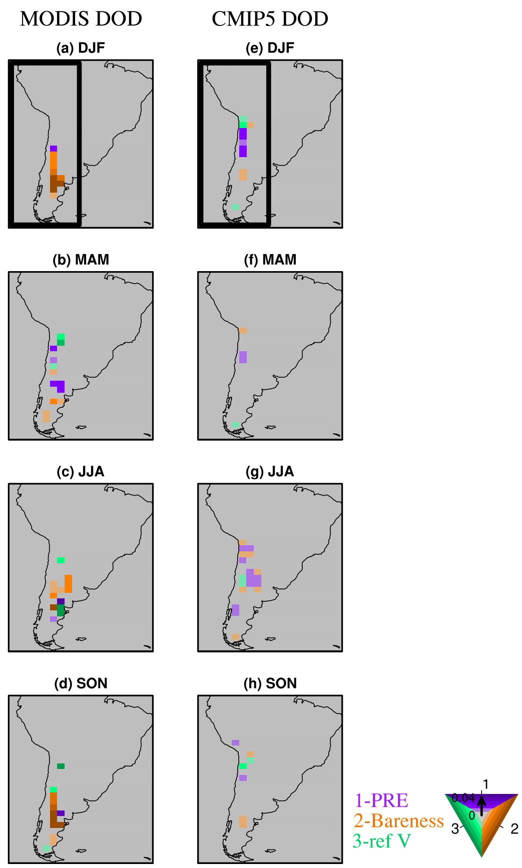


Figure R13. Same as Fig. 6 but for South America. Black boxes denote the averaging regions defined in Table 2: South America.

Reference:

Xiao ZQ, Liang SL, Wang JD, Chen P, Yin XJ, Zhang LQ, et al. Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance. IEEE T. Geosci. Remote 52, 209-223 (2014)

How reliable are CMIP5 models in simulating dust optical depth?

Bing Pu[1,2] and Paul Ginoux[2]

[1]Atmospheric and Oceanic Sciences Program, Princeton University,

Princeton, New Jersey 08544

[2]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey 08540

Correspondence to: Bing Pu (bpu@princeton.edu)

Abstract. Dust aerosol plays an important role in the climate system by affecting the radiative and energy balances. Biases in dust modeling may result in biases in simulating global energy budget and regional climate. It is thus very important to understand how well dust is simulated in the Coupled Model Intercomparison Project Phase 5 (CMIP5) models. Here seven CMIP5 models using interactive dust emission schemes are examined against satellite derived dust optical depth (DOD) during 2004-2016.

It is found that multi-model mean can largely capture the global spatial pattern and zonal mean of DOD over land in present-day climatology in MAM and JJA. Global mean land DOD is underestimated by -25.2% in MAM to -6.4% in DJF. While seasonal cycle, magnitude, and spatial pattern are generally captured by multi-model mean over major dust source regions such as North Africa and the Middle East, these variables are not so well represented by most of the models in South Africa and Australia. Interannual variations of DOD are neither captured by most of the models nor by multi-model mean. Models also do not capture the observed connections between DOD and local controlling factors such as surface wind speed, bareness, and precipitation. The constraints from surface bareness are largely underestimated while the influences of surface wind and precipitation are overestimated.

Projections of DOD change in the late half of the 21st century under the Representative Concentration Pathways 8.5 scenario by multi-model mean is compared with those projected by a regression model. Despite the uncertainties associated with both projections, results show some similarities between the two, e.g., DOD pattern over North Africa in DJF and JJA, an increase of DOD in the central Arabian Peninsula in all seasons, and a decrease over northern China from MAM to SON.

1. Introduction

Dust is the second most abundant aerosols by mass in the atmosphere after sea salt. It absorbs and scatters both shortwave and longwave radiation and thus modifies local radiative budget and consequently vertical temperature profile, influencing global and regional climate. For instance, studies found dust influences the strength of the West African monsoon (e.g., Miller and Tegen, 1998; Miller et al., 2004; Mahowald et al., 2010; Strong et al., 2015) and Indian monsoonal rainfall (e.g., Vinoj et al., 2014; Jin et al., 2014, 2015, 2016; Solmon et al., 2015; Kim et al., 2016; Sharma and Miller, 2017). Dust aerosols are also found to amplify droughts during the U.S. Dust Bowl and Medieval Climate Anomaly (Cook et al., 2008, 2009, 2013), and affect Atlantic tropical cyclones (e.g., Dunion and Velden, 2004; Wong and Dessler, 2005; Evan et al., 2006; Sun et al., 2008; Strong et al., 2018). Dust particles can also serve as ice cloud nuclei and influence the properties of the cloud (e.g., Levin et al., 1996; Rosenfield et al., 1997; Wurzler et al., 2000; Nakajima et al., 2001; Bangert et al., 2012) and affect regional radiative balance and hydrological cycle. When deposited in the oceans, iron-enriched dust also provides nutrients for phytoplankton, affecting ocean productivity and therefore carbon and nitrogen cycles and ocean albedo (e.g., Fung et al., 2000; Jickells et al., 2005; Shao et al., 2011~~; Jickells et al., 2005~~).

Globally, the estimated radiative forcing from dust aerosol is 0.10 (-0.30 to +0.10) W m^{-2}, a magnitude about one fourth of the radiative forcing of sulfate aerosol or black carbon from fossil fuel and biofuel (Myhre et al., 2013; their Table 8.4). Biases in dust simulation may potentially affect global energy budgets and regional climate simulation.

Thus, it is very important to examine the capability of current state-of-the-art climate models in simulating dust.

Only a few studies examined the Coupled Model Intercomparison Project Phase 5 (CMIP5) model output of dust and most of them are regional evaluations. For instance, Evan et al. (2014) examined model output for Africa, but mainly focused on an area over the northeastern Atlantic (10°–20°N and 20°–30°W) where a long-term proxy of dust optical depth data over Cape Verde islands is available (Evan and Mukhopadhyay, 2010). They found models underestimated dust emission and mass path and failed to capture the interannual variations from 1960 to 2004, as models did not capture the negative connection between dust mass path and precipitation over the Sahel.

Another work examined CMIP5 aerosol optical depth (AOD) is by Sanap et al. (2014) for India. They compared dust distribution in the models with Earth Probe total ozone monitoring system (EPTOMS)/ Ozone monitoring Instrument (OMI) aerosol index (AI) from 2000 to 2005. They found most of CMIP5 models, except two HadGEM2 models, underestimated dust load over Indo-Gangetic Plains, and suggested the biases are due to a misrepresentation of 850 hPa winds in the models. Later, Misra et al. (2016) also examined CMIP5 modeled AOD for India but did not specifically focus on dust.

Shindell et al. (2013) examined the output of 10 models from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) for one year (2000), among which eight models also participated in the CMIP5. They noticed that simulated dust AOD vary by more than a factor of two across models. However, this study also did not focus on dust, but emphasized the radiative forcings from anthropogenic aerosols.

None of the above studies examined global dust simulation in CMIP5 models. What's more, most studies focused on annual mean, not seasonal averages. It is very possible that models perform better in some seasons than others. AeroCom multiple-dust model intercomparison was performed on both global and regional scales (Huneeus et al., 2011) but only focused on one year, thus models' capability of simulating interannual or long-term variability of dust is not clear. A comprehensive evaluation of the climatology and interannual variation of global dust optical depth (DOD) in CMIP5 models will provide ~~a clear picture of~~insights into models' capability of ~~dust simulation~~simulating the integrated aerosol extinction due to dust, which is one of the key variables that determine the radative forcing of dust to the climate system.

Here we examine the results of seven CMIP5 models (Table 1) by comparing model output with DOD derived from Moderate Resolution Imaging Spectroradiometer (MODIS) Deep Blue aerosol products. Projections on changes of DOD in the late half of the 21st century by CMIP5 models and also by a regression model (Pu and Ginoux, 2017) are examined and analyzed.

The following section introduces data and methods used in this study. Results are presented in section 3, including examinations on the climatology and interannual variations of ~~modeled~~ CMIP5 DOD and future projections. Discussion and major conclusions are presented in sections 4 and 5, respectively.

2. Data and Methodology

2.1 DOD from MODIS

DOD is a widely used variable that describes optical depth due to the extinction by mineral particles. It is one of the key factors (single scattering albedo and asymmetry factor being the two others) controlling dust interaction with radiation. Monthly DOD are derived from MODIS aerosol products retrieved using the Deep Blue (MDB2) algorithm, which employs radiance from the blue channels to detect aerosols globally over land even over bright surfaces, such as desert (Hsu et al., 2004, 2006). Ginoux et al. (2012b) used collection 5.1 level 2 aerosol products from MODIS aboard the Aqua satellite to derive DOD. Here, both MODIS aerosol products (collection 6, level 2; Hsu et al., 2013) from the Aqua and Terra platforms are used. Aerosol products such as AOD (550 nm), single scattering albedo, and the Ångström exponent are first interpolated to a regular 0.1° by 0.1° grid using the algorithm described by Ginoux et al. (2010). The DOD is then derived from AOD following the methods of Ginoux et al. (2012b) with adaptions for the newly released MODIS collection 6 aerosol products (Pu and Ginoux, 2016). To separate dust from other aerosols, we use the Ångström exponent (α) and single scattering albedo (ω). Ångström exponent has been shown to be highly sensitive to particle size (Eck et al., 1999). A continuous function relating the Ångström exponent to fine-mode aerosol optical depth established by Anderson et al. (2005; their Eq. 5) based on ground-based data is used to separate dust from fine particles. We also screen the data by setting single scattering albedo at 470 nm to be less than one for dust due to its absorption of solar radiation. This separates dust from scattering aerosols such as sea salt, which is purely scattering. The formula can be summarized as the following:

$$DOD = AOD \times (0.98-0.5089\alpha + 0.0512\alpha^2) \quad \text{if } (\omega < 1) \quad . \qquad (1)$$

Note that DOD represents the coarse mode fraction of dust only. It is estimated that the fine mode dust at emission is less than 10% (Kok et al., 2017).

Aqua and Terra DOD have previously been used to study global dust sources (Ginoux et al., 2012b), and their geomorphological signature (Baddock et al., 2016), dust variations in the Middle East (Pu and Ginoux, 2016) and the U.S. (Pu and Ginoux, 2017), and have been validated with Aerosol Robotic NETwork (AERONET) stations over the U.S. (Pu and Ginoux, 2017). Here we compare Aqua and Terra DOD against AERONET stations globally (Section 1 and Figures. S1-2 in the Supplement). Both Aqua and Terra DOD is slightly underestimated, with respective errors of 0.08+0.52DOD and 0.10+0.48DOD.

Daily DOD ~~is derived for both~~from Aqua and Terra ~~satellites and then~~are averaged to monthly data and interpolated to a 1° by 1° grid. Terra passes the Equator from north to south around 10:30 local time while Aqua passes the Equator from south to north around 13:30 local time. To reduce missing data and also to combine the information from both morning and afternoon hours, a combined monthly DOD (here after MODIS DOD) is derived by averaging Aqua and Terra DOD when both products exist or using either Aqua or Terra DOD when only one product is available. As shown in Figure S3~~1~~ in the Supplement, the mean available days in each season and also spatial coverage are enhanced in combined DOD than using Aqua or Terra (not shown) DOD alone. This combined DOD is available from January 2003 to December 2016.

~~Aqua and Terra DOD product has previously been used to study global dust sources (Ginoux et al., 2012b), dust variations in the Middle East (Pu and Ginoux, 2016) and the U.S. (Pu and Ginoux, 2017), and has been validated with Aerosol Robotic NETwork (AERONET) stations over the U.S. (Pu and Ginoux, 2017). Here we~~

We also compared MODIS DOD climatology with both AERONET observation and DOD retrieved from Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP; Winker et al., 2004; 2007) aboard the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite. AERONET stations provide cloud-screened and quality assured (level 2) coarse mode aerosol optical depth (COD) at 500 nm, which is processed by the Spectral Deconvolution Algorithm (O'Neill et al., 2003). Only nine sites have long-term COD records during 2003-2016, and the climatological mean of MODIS DOD generally compares well with these sites (Figure S4~~2~~ in the Supplement).

CALIOP measures backscattered radiances attenuated by the presence of aerosols and clouds and retrieves corresponding microphysical and optical properties of aerosols. Monthly dust AOD (or DOD) on a 2° latitude by 5° longitude grid are available since June 2006. The climatology of CALIOP DOD during 2007-2016 is similar to that of MODIS DOD during the same period (Figure S5~~3~~ in the Supplement). The global mean (over land) MODIS DOD is slightly higher than that from CALIOP, probably due to the lower horizontal resolution of the latter. The pattern correlations (e.g., Pu et al., 2016) between the two products range from 0.83 in boreal spring and summer to 0.63 in boreal winter (Figure S5~~3~~ in the Supplement).

Due to higher spatial resolution (compared with CALIOP) and coverage (compared with AERONET sites), MODIS DOD is chosen as the primary product to validate CMIP5 model output. Nine regions (Table 2) are selected to study the DOD magnitude, spatial pattern, and variations. These regions cover major dust source regions previously identified (Ginoux et al. 2012).

Given the analysis above (Figs. S3-5), there are some uncertainties associated with DOD in a few regions in some seasons: (1) relatively low coverage (<30 days per season) over northern China and southeastern Asia in JJA; (2) DOD is slightly higher than COD from AERONET over the Arabian Peninsula in DJF and SON; (3) DOD is lower than CALIOP over northern India in MAM. We will consider these uncertainties in the following analysis wherever is relevant.

2.2 Reanalysis and observation datasets

Previous study found that the variations of dust event frequency over the U.S. in the recent decade could be largely represented by the variations of three local controlling factors: seasonal mean surface wind speed, bareness, and precipitation (Pu and Ginoux, 2017). These factors have previously been found to constrain dust emission or variability on multiple time scales (e.g., Gillette and Passi, 1988; Fecan et al., 1999; Zender and Kwon, 2005). While surface wind is positively related to the emission and transport of dust, vegetation is an important non-erodible element that prevents wind erosion. Precipitation is generally negatively related to dust emission and transport processes. While the scavenging effect of precipitation on small dust particles only lasts a few hours or days, influences of precipitation on soil moisture lasts longer.

To examine the interannual variations of DOD and its connection with local controlling factors such as surface wind speed, bareness, and precipitation, monthly data of 10 m wind speed from the ERA-Interim (Dee et al., 2011), leaf area index (LAI) data from Advanced Very High Resolution Radiometer (AVHRR; Claverie et al., 2014, 2016), and precipitation from the Precipitation Reconstruction over Land (PRECL; Chen et al., 2002) are used.

ERA-Interim is a global reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF). Its horizontal resolution is T255 (about 0.75° or 80 km~~),~~). We choose this analysis because of its relatively high spatial resolution~~very suitable to study the influence of wind speed on dust emission and transport on small scales~~. The monthly data are available from 1979 to present day.

Monthly LAI derived from the version 4 of Climate Data Record (CDR) of AVHRR is used to calculate surface bareness. The data are produced by the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) and the University of Maryland. Monthly gridded data on a horizontal resolution of 0.05° by 0.05° degree are available from 1981 to present. This product is selected due to its high spatial resolution and long temporal coverage. Surface bareness is calculated from seasonal mean LAI (Pu and Ginoux, 2017) as the following,

$$Bareness = exp\ (-\cancel{1\times}LAI) \qquad . \qquad (2\cancel{1})$$

Bareness is originally defined as *exp (-LAI-SAI)*, where *SAI* is stem area index (Evans et al. 2016). Since satellite does not retrieve brownish SAI, we only use LAI to calculate bareness.

PRECL precipitation from the National Oceanic and Atmospheric Administration (NOAA) is a global analysis available monthly from 1948 to present at a 1° by 1° resolution. The dataset is derived from gauge observations from the Global Historical Climatology Network (GHCN), version 2, and the Climate Anomaly Monitoring System (CAMS) datasets. Its long coverage and ~~relatively high~~ spatial resolution is ~~quite~~ suitable to study the connections between DOD and precipitation.

2.3 CMIP5 model output

Among CMIP5 models we selected seven models (Table 1) that used interactive dust emission schemes, in which dust emission varied in response to changes of climate. The output of 10 m wind speed, precipitation, and LAI are also available from these models. In models that dust is simulated offline, i.e., dust emission did not interactively respond to meteorological and climate changes, the connections between DOD and modeled controlling factors are lost. Other models (to our best knowledge) either used offline dust as an input~~, in which dust emission did not interactively respond to meteorological and climate changes,~~ or did not write out the variables needed for this analysis.

Both historical run from 1861 to 2005 and future run under the Representative Concentration Pathways 8.5 (RCP 8.5) scenario (Riahi et al., 2011) from 2006 to 2100 are used. Here the RCP 8.5 scenario is chosen because it represents the upper limit of the projected greenhouse gas change in the twenty-first century and thus likely is the worst-case scenario for future DOD variation under climate change. Also, studies found that observed CO_2 emission pathway during 2005-2014 matches RCP 8.5 scenario better than other scenarios (e.g., Fuss et al., 2014), which makes the RCP8.5 output suitable to examine present-day DOD variations after 2005.

Monthly model output of dust load, surface 10 m wind speed, precipitation, and LAI are used. Historical output from 2003 to 2005 and RCP 8.5 output from 2006 to 2016 are combined to form time series and climatology during 2003-2016 to compare with MODIS DOD during the same time period.

2.3.1 DOD derived from modeled dust load

Most CMIP5 models did not save DOD, so we used monthly dust load and converted them to DOD using the relationship derived by Ginoux et al. (2012a) as the following

$$\tau = M \times e\ , \qquad (3\cancel{2})$$

where τ is DOD at 500 nm, M is the load of dust in unit of (g m^{-2}), and e = 0.6 m^2 g^{-1} is the mass extinction efficiency. Dust load from different models is first interpolated to a 2° by 2.5° grid and then converted to DOD. The same method was used by Pu and Ginoux (2017) for the U.S. Applying the same mass extinction efficiency everywhere and to all the CMIP5 model output used here is a simplification, as different models may have quite different mass extinction efficiency. For instance, e can range from 0.25 to 1.28 m^2 g^{-1} in AEROCOM models, with a multi-model medium of 0.72 m^2 g^{-1} (Huneeus et al., 2011). Here, w~~W~~e compare~~d~~ the derived DOD with modeled DOD from one historical simulation of GFDL~~LD~~-CM3 model (Donner et al., 2011) as an example. A full validation of this method will require modeled DOD from all the other CMIP5 models, which are currently not available. The pattern correlation of the climatology (1861-2005) between the derived DOD and modeled DOD in GFDL-CM3 are very high, all above 0.99 for four seasons (not shown). The percentage differences between derived DOD and modeled DOD averaged over global land range from -3.6% in DJF and SON to 1.3% in MAM and JJA. ~~Over Africa, DOD is slightly overestimated by 0~6.7% (regional mean), while over the Middle East, there is a small underestimation by -1.6% in SON and up to 8.2% overestimation in JJA. Among the nine regions we focused in this analysis, three regions (North America, South Africa, and South America) show an underestimation of~~

~~more than 20% in some seasons and two regions (Northern China and Australia) show an overestimation of more than 10% in some seasons.~~

2.4 ~~Multiple linear~~ A linear r~~r~~egression model

2.4.1 Multiple linear regression

In order to examine the relative contribution of each local controlling factor to DOD variations, multiple linear regression is applied by regressing MODIS DOD onto standardized seasonal mean ERA-Interim surface wind speed, AVHRR bareness, and PRECL precipitation at each grid point. All the data are re-gridded to a 1° by 1° grid before the calculation. Over regions where values are missing for any of the explanatory variables (i.e., precipitation, bareness, and surface wind speed) or DOD, the regression coefficients are set to missing values. The collinearity among these explanatory variables is examined by calculating variance inflation factor (VIF) (e.g., O'Brien, 2007; Abudu et al., 2011), and in most regions the VIF is below 2 (not shown), indicating a low collinearity (5–10 is usually considered high). Bootstrap resampling is used to test the significance of the regression coefficients, following the method used by Pu and Ginoux (2017).

Multiple linear regression is also applied to CMIP5 model derived DOD and output of surface wind speed, bareness, and precipitation to obtain regression coefficients from the models from 2004 to 2016. All variables are interpolated to a 2° by 2.5° grid before regression. The results are compared with regression coefficients derived from observational datasets.

2.4.2 DOD reconstruction and future projection

Using regression coefficients obtained from observations and observed variations of precipitation, bareness, and surface wind speed from 2004 to 2016, we can reconstruct DOD in the present day and compare it with MODIS DOD (see discussion in section 3.2).

Similar to the method used by Pu and Ginoux (2017), the regression coefficients derived from MODIS DOD and observed controlling factors ~~from 2004 to 2016~~ and CMIP5 model output of surface wind speed, bareness, and precipitation are used to project variations of future DOD. The regression coefficients are interpolated from the 1° by 1° grid to a 2° by 2.5° grid to be consistent with model output. Such an interpolation may smooth out some spatial characteristics from observations. Here we tried two groups of CMIP5 output for these controlling factors. One group used seven models with interactive dust emission scheme (Table 1), and the other used 16 CMIP5 models ~~as did by~~ (see Supplementary Table S1 of Pu and Ginoux, ~~(~~2017~~; their Supplementary Table S1~~) that include the seven models with interactive dust emission scheme. The reason to test the latter is to include as much model output of the controlling factors as possible. The differences between the historical run (1861–2005 average) and that of the RCP 8.5 run for the late half of the twenty-first century (2051–2100) are standardized by the standard deviation of the historical run for each explanatory variable. The projected change reveals how DOD will vary with reference to the historical conditions (mean and standard deviation).

3. Results

3.1 Climatology (2004-2016)

Figure 1 shows the climatology of MODIS DOD (top panel) in four seasons during 2004-2016 and that from the CMIP5 multi-model mean (bottom). Globally, the dustiest regions are largely located over the northern hemisphere (NH) over North Africa, the Middle East, and East Asia (Figs. 1a-d). In these regions, DOD is higher in boreal spring and summer than fall and winter. Modeled global DOD over land is generally lower than that from MODIS DOD, ranging from -0.028 (-25.2%) in MAM to -0.005 (-6.4%) in DJF. The global spatial pattern is better captured in MAM and JJA, with pattern correlations of 0.74 and 0.85, respectively (Figs. 1f-g). In DJF, DOD is overestimated over central Africa and Australia, but underestimated over the Middle East and Asia (Fig. 1e), while in SON there is a similar overestimation in Australia and an underestimation in the Middle East (Fig. 1h).

Figure 2 shows the zonal mean of CMIP5 DOD from individual models (thin colorful lines) and multi-model ensemble mean (thick black), in comparison with MODIS DOD (thick red). In DJF, DOD is underestimated in the NH from 15° N to 50°N but overestimated over the tropics and southern hemisphere (SH) (Fig. 2a). While the overestimation in the SH is largely contributed by three models, the underestimation in the NH appears in all the seven models. The overestimation of DOD in HadGEM2-ES has also been identified in a previous study (Bellouin et al., 2011) and will be discussed later. In MAM, a similar overestimation of DOD in the tropics and SH also occurs in some models, and the multi-model mean slightly overestimates DOD around 20°-30°S (Fig. 2b). In NH, there is a weak underestimation too, but the overall gradient is largely captured. In JJA, the multi-model mean resembles MODIS DOD very well (Fig. 2c), consistent with the highest pattern correlation in this season shown in Fig. 1. The peak around 19° N in North Africa and Middle East is well captured by the multi-model mean, although the magnitude is slightly underestimated. In SON, different from MODIS DOD that peaks around 19°N, the multi-model mean has two peaks around 15°N and 28°S, respectively, a pattern somewhat similar to that in DJF (Fig. 2d). Consequently, DOD in CMIP5 multi-model mean is overestimated at 15°-40°S and 0°-15°N but underestimated at 15°S -0° and 15°-40°N.

Seasonal cycles of CMIP5 DOD are compared with MODIS DOD in nine regions in Figure 3. The annual means of DOD in each region from multi-model mean (black) and MODIS (red) are also listed in each plot. The spread of DOD among individual models is greater during boreal spring and summer for regions in the NH and during austral spring and summer for regions in the SH ~~than other seasons~~. Seasonal cycles over North Africa, the Middle East, North America, and India are generally captured by multi-model mean, with modeled DOD peaking during the same seasons as MODIS DOD (Figs. 3a-b, d-e). While some models overestimate the seasonal peaks over the Middle East, North America, and India (e.g., CanESM2, HadGEM2-ES, and HadGEM2-CC), a few models have very weak seasonal cycles and underestimate DOD over North America and India (e.g., GFDL-CM3, NorESM1-M, MIROC-ESM, and MIROC-ESM-CHEM). Note that MODIS DOD is slightly lower than CALIOP DOD over India in MAM (Fig. S5), therefore for these models the underestimation may be larger than shown in Fig. 3e.

Since the temporal coverage of MODIS DOD over northern China and southeastern Asia is relatively low in JJA compared with other regions (Fig. S3), we also examined the seasonal cycle of CALIOP DOD (not shown) and results are similar but with weaker magnitude. Over northern China, MODIS DOD peaks in spring (Fig. 3c), consistent with previous studies (e.g., Zhao et al., 2006; Laurent et al., 2006; Ginoux et al., 2012b), while multi-model mean peaks ~~much~~ later in May-June. Individual models have quite different seasonal cycles, with GFDL-CM3 model having a peak (in April) closer to the timing of MODIS maximum. Similar misrepresentation occurs over the southeastern Asia (Fig. 3f).

In South Africa and South America the observed maxima in early austral spring (i.e., September) are also ~~missed~~not captured by the multi-model mean (Figs. 3g-h). Note that CanESM2 largely captures the seasonal cycle of DOD over South America, although the magnitude is overestimated (Fig. 3h). In Australia, DOD is largely overestimated and the peak from November to January in MODIS DOD is ~~also misrepresented in the~~shifted about one month earlier in the multi-model mean (Fig. 3i). Similar to the finding here, Bellouin et al. (2011) also found that HadGEM2-ES model overestimated DOD over Australia and Thar desert region in northwestern India and suggested that these overestimations were likely due to model's overestimation of bare soil fraction and underestimation of soil moisture. Despite overestimation, the seasonal cycle in HadGEM2-CC model is more similar to MODIS DOD than other models (Fig. 3i).

We further examine the magnitudes and spatial patterns of CMIP5 DOD in these regions. Figure 4 shows the ratio of pattern standard deviations (standard deviations of values within the domain) and pattern correlation between CMIP5 DOD and MODIS DOD climatology (2004-2016) in each region for four seasons. While the former reveals the magnitude differences, the latter demonstrates the spatial resemblance.

Over North Africa, the Middle East, and India, the ratio of CMIP5 DOD from individual models and multi-model mean versus MODIS DOD are all within ± one order of magnitude (Fig. 4). Most models underestimate DOD in northern China, although the magnitudes are largely within the range of -one order of magnitude to one. Over North America, South Africa, and Australia, some models underestimate the DOD by more than two orders of magnitudes, while over Australia three models overestimate DOD by more than one order of magnitude. In general, magnitudes of multi-model mean are closer to satellite DOD than most individual models and are largely within ± one order of magnitude of MODIS DOD.

The spatial patterns are better captured over North Africa and the Middle East than other regions (Fig. 4), with pattern correlations above 0.6 in most models (with highest pattern correlation of 0.92 and 0.83, respectively). Pattern correlations from multi-model mean are also high, reaching 0.87 (0.78) over North Africa and 0.75 (0.73) over the Middle East in JJA (MAM). Nonetheless, some models show negative pattern correlations over North Africa, northern China, North America, southeastern Asia, South Africa, South America, and Australia. Overall, spatial patterns are less well represented in regions over the SH than over the NH in CMIP5 models.

In short, in terms of both magnitudes and spatial pattern, DOD climatology is best represented over North Africa and the Middle East among the nine regions. The multi-model mean shows that DOD over North Africa is slightly better simulated than over the Middle East, somewhat similar to the finding of AeroCom multi-model analysis (Huneeus et al. 2011).

3.2 Interannual variations

An important aspect of dust activity is its long-term variability~~variations~~, including interannual and decadal variations. Dust emission in North Africa is known to have strong decadal variations (e.g., Prospero and Nees, 1986; Prospero and Lamb, 2003; Mahowald et al., 2010; Evan et al., 2014, 2016), while over Australia, strong interannual variations have been related to El Niño–Southern Oscillation (e.g., Marx et al., 2009; Evans et al., 2016). Due to the short time coverage of high quality satellite products, we focus on interannual variations of DOD from 2004 to 2016.

Figure 5 shows the correlations of regional mean time series of DOD between MODIS and CMIP5 models and multi-model mean for each season in nine regions. We also show correlations between the reconstructed DOD (see section 2.4.2 for details) and MODIS DOD for reference (Table S1 in the Supplement). ~~Previous study found that the variations of dust event frequency over the U.S. in the recent decade could be largely represented by the variations of three local controlling factors: seasonal mean surface wind speed, bareness, and precipitation (Pu and Ginoux, 2017). These factors have previously been found to constrain dust emission or variability on multiple time scales (e.g., Gillette and Passi, 1988; Fecan et al., 1999; Zender and Kwon, 2005). While surface wind is positively related to the emission and transport of dust, vegetation is an important non-erodible element that prevents soil erosion from wind. Precipitation is generally negatively related to dust emission and transport processes. While the scavenging effect of precipitation on small dust particles only lasts a few hours or days, influences of precipitation on soil moisture lasts longer. Here we extend our regression model (Pu and Ginoux, 2017) to a global scale. Regression coefficients are obtained by regressing MODIS DOD onto observed surface wind, bareness, and precipitation during~~

~~2004-2016 (see methodology section for details).~~ The reconstructed DOD is ~~then~~ calculated using ~~these~~ observed regression coefficients and time-varying controlling factors from observations (i.e., surface wind speed, bareness, and precipitation).

The interannual variations of DOD are in general not well captured by CMIP5 models. This is consistent with previous study by Evan et al. (2014) who found dust variability downwind of North Africa over the northeastern Atlantic was misrepresented in CMIP5 models. In most regions, only one or two models show significant positive correlation with MODIS DOD in some seasons, and negative correlations exist in all regions (Fig. 5). North Africa, the Middle East, southeastern Asia, South America, and Australia show less negative correlations than other dusty regions. On the other hand, reconstructed DOD shows significant positive correlations with MODIS DOD over most regions in all seasons (Table S1 in the Supplement). This suggests that the interannual variations of DOD can be largely attributed to the variations of these controlling factors, and models ~~probably~~ may misrepresent~~ed~~ these relationships, in addition to their incapacity of capturing the interannual variations of individual controlling factors in general (~~Figures S4-6 in the Supplement~~not shown), which is not uncommon for coupled models.

We further examine the connection between those controlling factors and DOD in CMIP5 models. Figure 6 shows the dominant controlling factors among the three (surface wind speed, bareness, and precipitation) on DOD variations in four seasons from MODIS (left column) and from CMIP5 multi-model mean (right column), respectively. To highlight factors controlling DOD variations near the dust source regions, a mask of AVHRR LAI≤ 0.5 is applied to both coefficients.

Bareness plays the most important role in many dusty regions in observations, e.g., over Australia, central U.S., and South America (Figs. 6a-d). Note that while bareness plays an important role over the Sahel during DJF and MAM, it also shows strong signal over some areas in the northern North Africa (Figs. 6a-b). The reliability of this information is limited by the accuracy of LAI retrieval in these areas. The value of bareness in this region is actually quite high (as LAI is very low), but still has weak interannual variability (Figures S6~~7~~ in the Supplement). Over some areas of North and South Africa, the Middle East, and East Asia, surface wind and precipitation are also quite important.

The role of bareness is largely underestimated in CMIP5 models, while surface wind and precipitation become the dominant factors (Figs. 6e-h). The misrepresentation of the connection between DOD and these controlling factors may cause the misrepresentation of the dust load and its variability. Taking Australia for an example, the overestimation of DOD magnitudes may be related to an overestimation of the influence of surface wind on DOD and a lack of constraints from surface bareness.

Despite the large differences between the observed and modeled connections between DOD and the controlling factors, some regions show similarities. For instance, over North Africa in DJF, both show an important influence from surface winds (Figs. 6a, e), although the locations of surface wind-dominant areas are not exactly the same. Evan et al. (2016) also found a dominant role of surface wind on African dust variability, but they focused on monthly means, not seasonal averages. In MAM, precipitation starts to play a role in some parts of North Africa, while surface wind still dominates in some areas (Fig. 6b). Same increasing influence of precipitation is shown in the multi-model mean, but such an influence seems overestimated (Fig. 6f). In JJA, the influences of ~~surface wind in North Africa and~~ precipitation and bareness over the eastern Arabian Peninsula~~in the Middle East~~ in the multi-model mean (Fig. 6g) also show some similarity to observation (Fig. 6c), although an underestimation of the influence from bareness and an overestimation of ~~surface wind~~precipitation are still there.

Also, note that in CMIP5 models, due to lack of constraints from low surface temperature (e.g., over frozen land) and snow cover on dust emission or misrepresentations of dust transport, DOD and also the regression coefficients still exist over NH high latitudes in boreal winter and spring in the multi-model mean (Figs. 6e-f).

3.3 Future projections

How will DOD change in response to increasing greenhouse gases? The results from CMIP5 multi-model mean are shown in Figure 7. We compare the DOD during the late half of the 21st century under the RCP 8.5 scenario with that in the historical level (1861-2005 average).

Over land, CMIP5 model projects a decrease of global mean DOD in all seasons except JJA (Figs. 7a-d). The inter-model standard deviation is much greater than the multi-model mean, suggesting large discrepancies among individual models. The projected decrease is largely over northern North America, southern North Africa, eastern central Africa, and East Asia, while the increase is largely over northern North Africa, the Middle East, southern North America, South Africa, South America, and southern Australia (Fig. 7). Regional means of DOD change (in percentage) with reference to CMIP5 historical run are summarized in Table 3.

What might be the causes of DOD change? Figure 8 shows the projected change of precipitation, bareness, and surface wind speed from CMIP5 multi-model mean. These factors play important role in DOD variations in the present day, although models tend to underestimate the role of bareness and overestimate the influences of precipitation and surface wind (Fig. 6). Increases in precipitation can increase soil moisture and remove airborne dust, thus usually favors a decrease of DOD. As shown in Figs. 8a-d, the increases of precipitation in northern Eurasia, northern North America, the Congo basin in Africa, and Australia (DJF and MAM) may contribute to the decrease of DOD in these regions, while the decreases of precipitation over northern North Africa and the Middle East (DJF and MAM), South Africa, and South America may contribute to the increase of DOD (DJF-SON). Also note that in JJA both precipitation and DOD increase over northern North Africa and the Middle East (Fig. 8c), suggesting other factors dominate the variation of DOD in the multi-model mean.

A decrease (increase) of bareness indicates a growth (decay) of vegetation and is usually associated with a decrease (increase) of DOD. In general, except regions such as southern North America, South America, South Africa, part of northern Eurasia, and central Sahel, the pattern of bareness change does not resemble DOD change (Figs. 8e-h). This is probably due to the fact that the overall influence of bareness on DOD variation is underestimated in CMIP5 models (Fig. 6).

Increases in surface wind can enhance dust emission and transport, and vise versa. The changes of surface wind in DJF and MAM are similar and likely to contribute to the increase of DOD over northern North Africa, the Middle East, eastern South America, southern South Africa, and southern Australia (Figs. 8i-j). The decrease of DOD over northwestern North America, the Sahel, and northern Australia may also relate to the decrease of surface wind there, in addition to an increase of precipitation and a reduction of bareness. In JJA and SON (Figs. 8k-l), the increases of surface wind in South America, South Africa, central Australia and the decreases of wind in northwestern North America, northern Eurasia, and the central Sahel are also consistent with patterns of DOD change.

In short, variations of CMIP5 DOD in the late half of the 21st century ~~centaury~~ are more consistent with changes of precipitation and surface wind speed than with surface bareness, consistent with the analysis above regarding to the present-day condition.

Here we also present t~~T~~he projected change of DOD from the regression model ~~is shown~~in Figure 9. The regression model (see section 2.4 for details) is developed based on observed relationships between MODIS DOD and local controlling factors and can largely capture the interannual variations of DOD in the present-day climate (Table S1 in the Supplement). Assuming that the observed connection between DOD and these controlling factors do not change dramatically in the future, we can use this regression model and CMIP5-model projected change of controlling factors to project DOD variations. Compared to DOD projection from CMIP5 models, this approach utilizes additionally observational constrains and is likely to provide a more reliable future projection. ~~The results are calculated using~~We use ~~the regression coefficients obtained from observations during 2004-2016 and~~ projected changes of precipitation, bareness, and surface wind speed from seven~~16~~ CMIP5 models with interactive dust emission scheme (see methodology). A similar method is applied to the model output from ~~seven~~ 16 CMIP5 models ~~with interactive dust emission scheme~~, and results are similar (Figure

S7~~8~~ in the Supplement). A mask of present-day LAI ≤ 0.5 is also applied to highlight the changes of DOD near dust source regions. By doing this, we assume the location of major dust sources will not change much at the late half of the 21st century. The unmasked figure is presented in the supplementary file (Figure S8~~9~~ in the Supplement). The reason we did not use the projected future LAI as a mask is that there're large uncertainties associated with LAI projection, especially over northern hemisphere subtropical regions (e.g., Figs. 8e-h).

In DJF, regression model projected change of DOD over Mexico, North Africa, the Middle East and part of northern China (Fig. 9a) are similar to those projected by CMIP5 models over those dust source regions (Fig. 7a), but with a greater magnitude. In MAM, a decrease of DOD is projected over large area of North Africa (Fig. 9b), which is different from the pattern projected from the CMIP5 multi-model mean (Fig. 7b). The decrease of DOD over northern central U.S. is also different from the overall increase projected by CMIP5 DOD~~, as also noted by Pu and Ginoux (2017)~~. However, the increase of DOD over the Middle East and the decrease of DOD over northern China are similar to that of CMIP5 DOD. During JJA and SON, DOD decreases over the Sahel and northern China but increases over a belt to the north of central Sahel and parts of the Middle East (Figs. 9c-d). The weak increase of DOD over the southern corner of South Africa in JJA and a slight decrease in SON also has high agreement among the ~~models~~regression projections (dotted areas in Figs. 9c-d). Changes of DOD over Australia are very small in all seasons and show little consistency among the ~~models~~ regression projections.

The regression model projection using 16-model output shows very similar patterns (Figure S7 in the Supplement), largely because the projected changes of precipitation, surface wind speed, and bareness from 16-model ensemble mean are similar to those from 7-model ensemble mean in dusty regions (Figure S9 in the Supplement). But there are also some discrepancies in terms of magnitude and pattern that are revealed in the projected DOD patterns, e.g., the projected reduction of DOD is greater and more widespread over the northern Asia in MAM if using 16-model output and the increase of DOD along the southern edge of the Sahara is weaker in JJA and SON (Fig. S7 in the Supplement vs. Fig. 9).

The contribution of each controlling factor to the total DOD change is shown in Figure 10. While changes of bareness over North Africa~~, northern Middle East~~ and northern China play an important role in DOD change, changes of precipitation, e.g. over northwestern China in MAM, and surface wind, e.g., over northern North Africa and the Middle East in DJF and MAM, also play vital roles.

Both projections from the CMIP5 models and that from the regression model have ~~large~~ some uncertainties. The reliability of future projection by CMIP5 models is limited by models' capability of capturing present-day climatology and observed connection between DOD and local controlling factors. As discussed earlier, the overall performance of models is better in those very dusty regions in the NH, such as North Africa and the Middle East, than other regions. Multi-model mean also overestimates the connection between DOD and precipitation and surface wind and underestimates the influence of bareness in the present-day (Fig. 6), which can cast doubts on the projected variation of DOD in response to climate change.

The uncertainties associated with regression model are two folds. First, there're uncertainties associated with the regression model itself. Since the regression coefficients are derived from observed relationships between DOD and controlling factors in a relatively short time period, factors controlling the low frequency variation of DOD (e.g., decadal variations) may not be included. Other meteorological factors that could play an important role in regional dust variability, e.g., nocturnal low-level jets (e.g., Todd et al., 2008; Fiedler et al., 2013; Fiedler et al., 2016) and haboobs over Africa (e.g., Ashpole and Washington, 2013), are not directly considered in the model. The influences of anthropogenic land use/land cover change are also not included in the regression model. Anthropogenic land use/land cover change has been found to have played an important role in long-term dust variability in some regions (e.g., Neff et al., 2005; 2008; Moulin and Chiapello, 2006; McConnell et al., 2007), although previous modeling study found its influences on future dust emission was minor compared to climate change (Tegen et al., 2004). So the projection made by the regression model only reveals the change of DOD in association with climate change. Second, uncertainties associated with model projected change of controlling factors, such as bareness in U.S. in JJA as pointed by Pu and Ginoux (2017), also limit the accuracy of the results.

Despite these uncertainties, both methods make similar projections particularly in some dusty regions. For instance, the DOD pattern over North Africa in DJF and JJA, an increase of DOD in the central Arabian Peninsula in all seasons, and a decrease of DOD over northern China from MAM to SON (Figs. 7, 9).

4. Discussion

We examined DOD in seven CMIP5 models with interactive dust emission schemes. Other important variables that influence the radiative property ~~and concentration~~ of dust, such as Angström exponent and single scattering albedo~~, dust emission, and surface concentration~~, are also worth further examination, if these variables are archived. A better quantification of the radiative forcing of dust may also require an examination on the size distribution of dust particles, as studies (e.g., Kok et al., 2017) found in current AeroCom models fraction of coarse dust particles were underestimated and so was the warming effect of dust. Whether this is the case in the CMIP5 models is not clear.

Also note that since DOD is an integrated variable, it does not reflect the vertical distribution of dust aerosols. As pointed by Huneeus et al., (2016), dust models with similar performance in simulating aerosol optical depth may have quite large differences in simulating vertical distribution, emission, deposition, and surface concentration of dust. An overall evaluation of dust modeling capability will require detailed examination of these variables and the life cycle of dust in CMIP5 models in addition to DOD.

Early studies on future dust projection used offline dust models driven by climate model output under different scenarios. For instance, Mahowald and Luo (2003) used an offline dust model and output from National Center of Atmospheric Research's coupled Climate System Model (CSM) 1.0 (Boville and Gent, 1998) under A1 scenario (Houghton et al., 2001) and projected a decrease of dust emission by the end of the 21st century by -20% to -63%, depending on different scenarios. In general, when they included vegetation change, the projected dust reduction became greater, but including land use change slightly weakened such reduction. Similarly, Tegen et al. (2004) used output from ECHAM4 and HadCM3 and a dust model (Tegen et al., 2002) to examine the change of dust emission by 2040-2050 and 2070-2080 and found results were model and scenario dependent, from -26% to 10%. However, including anthropogenic cultivation practices tended to increase dust emission in both models. They also pointed out that such an influence from anthropogenic land-use was not big enough to overcome the effect of climate change.

The interactive dust emission schemes and new generations of climate models used in CMIP5 are likely to provide more reliable projections, but this may also depend on how changes of dust and its radiative forcing are fed back to the climate system in the models. While these projections are largely model-dependent, based on our analysis on the DOD climatology in CMIP5 models, the multi-model mean has a better chance to provide a more reliable projection than individual models.

Here a regression model combined with MODIS DOD is used to identify key local factors that control the variation of DOD on the interannual time scale. The results are then compared with model output to examine models' capability of capturing observed connections between DOD and controlling factors. This method may be applied to other dust model intercomparison projects as well, such as AeroCom (Huneeus et al. 2011), to help examine model performance.

5. Conclusion

Dust aerosol plays an important role in the climate system by directly scattering and absorbing solar and longwave radiation and indirectly affecting the formation and radiative properties of cloud. It is thus very important to understand how well dust is simulated in the state-of-the-art climate models. While many features and variables are systematically examined in the CMIP5 multi-model output, we found that to our best knowledge an evaluation of global dust modeling in CMIP5 models is still in blank. In this study we examined a key variable associated with dust radiative effect, dust optical depth (DOD), using seven CMIP5 models with interactive dust emission schemes and DOD retrieved from MODIS Deep Blue aerosol products.

We found that the global spatial pattern and magnitude are largely captured by CMIP5 models in the 2004-2016 climatology, with an underestimation of global DOD (over land) by -25.2% in MAM to -6.4% in DJF. The spatial pattern is better captured in boreal dusty seasons during MAM and JJA. In JJA, the simulated zonal mean DOD from multi-model mean largely ~~captures~~ resembles MODIS DOD~~ quite well~~.

The magnitudes of multi-model mean are closer to MODIS climatology than most individual models and are largely within ± one order of magnitude of MODIS DOD in the nine regions examined here (North Africa, the Middle East, n~~N~~orthern China, North America, India, southeastern Asia, South Africa, South America, and Australia; see Fig. 1 and Table 2 for domains). While some models underestimate DOD in North America and South America by more than two orders of magnitude, a few also overestimate DOD in Australia by more than one order of magnitude. Both the magnitude and spatial patterns of DOD are better captured over North Africa and the Middle East than other regions.

The multi-model mean also largely captures the seasonal cycle of DOD in some very dusty regions, such as North Africa and the Middle East. Seasonal variations in North America and India are also generally captured by the multi-model mean, with the modeled DOD peaking at approximately the same season as in MODIS DOD, but not so in n~~N~~orthern China and southeastern Asia. Seasonal cycles in those dusty regions in the southern hemisphere is generally not well captured, with modeled DOD over South Africa and South America peaking later than that in MODIS DOD but earlier in Australia.

The interannual variations of DOD are not captured by most of the CMIP5 models during 2004-2016. ~~This is likely due to models'~~Models also ~~underestimation~~ underestimate ~~of~~ the constraints from surface bareness on the variations of ~~dust~~ DOD and overestimate~~ion of~~ the influences from surface wind speed and precipitation in those major dust source regions~~, in addition to the fact that coupled models usually do not capture the observed interannual variations of precipitation, surface wind, and bareness as well~~. CMIP5 model projected change of DOD in the late half of the 21st century (under the RCP 8.5 scenario) with reference to historical condition (1861-2005) also shows greater influence from precipitation and surface wind change than from surface bareness. Overall, multi-model mean projects a change of DOD over land from -3.8% in SON to 3.3% in JJA.

We also provide a projection of future DOD change using a regression model based on local controlling factors such as surface wind, bareness, and precipitation (Pu and Ginoux, 2017). This model can largely capture the interannual variations of MODIS DOD in 2004-2016. The regression model projects a reduction of DOD in the Sahel in all seasons in the late half of the 21st century under the RCP 8.5 scenario, largely due to a decrease of surface bareness. DOD is projected to increase over the southern edge of the Sahara in association with surface wind and precipitation changes except in MAM, when a reduction of DOD over most part of North Africa is projected. DOD is also projected to increase over the central Arabian Peninsula in all seasons and to decrease over northern China from MAM to SON.

Despite large uncertainties associated with both projections, we find some similarities between the two, ~~which may be informative,~~ which adds to the confidence of projected DOD change in these regions, for instance, changes of DOD over North Africa in DJF and JJA, an increase of DOD in the central Arabian Peninsula in all seasons, and a decrease of DOD over northern China from MAM to SON.

Acknowledgements.

This research is supported by NOAA and Princeton University's Cooperative Institute for Climate Science and NASA under grant NNH14ZDA001N-ACMAP. The authors thank Drs. Songmiao Fan and Fabien Paulot for their helpful comments on the early version of this paper. The insightful comments from two anonymous reviewers improved the paper.

PRECL Precipitation data are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their web site at *http://www.esrl.noaa.gov/psd/*. The CALIOP products are downloaded from *https://www-calipso.larc.nasa.gov/tools/data_avail/dpo_read.php?y=2007&m=08&d=10*. AVHRR leaf area index data are available at: https://www.ngdc.noaa.gov/docucomp/page?xml=NOAA/NESDIS/NCDC/Geoportal/iso/xml/C00898.xml&view=getDataView&header=none. ~~ftp://eclipse.ncdc.noaa.gov/pub/cdr/lai-fapar/files/.~~ The ERA-Interim is downloaded from http://www.ecmwf.int/en/research/climate-reanalysis/era-interim. The AERONET coarse mode aerosol optical depth data are downloaded from https://aeronet.gsfc.nasa.gov/. CMIP5 data are downloaded from: https://pcmdi.llnl.gov/projects/esgf-llnl/.

Reference

Abudu, S., Cui, C. L., King, J. P., Moreno, J., and Bawazir, A. S.: Modeling of daily pan evaporation using partial least squares regression, Sci China Technol Sc, 54, 163-174, 10.1007/s11431-010-4205-z, 2011.

Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., Kharin, V. V., Lee, W. G., and Merryfield, W. J.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, Geophys Res Lett, 38, 10.1029/2010gl046270, 2011.

Ashpole, I., and Washington, R.: A new high-resolution central and western Saharan summertime dust source map from automated satellite dust plume tracking, J Geophys Res-Atmos, 118, 6981-6995, 10.1002/jgrd.50554, 2013.

Baddock, M. C., Ginoux, P., Bullard, J. E., and Gill, T. E.: Do MODIS-defined dust sources have a geomorphological signature?, Geophys Res Lett, 43, 2606-2613, 10.1002/2015gl067327, 2016.

Bangert, M., Nenes, A., Vogel, B., Vogel, H., Barahona, D., Karydis, V. A., Kumar, P., Kottmeier, C., and Blahak, U.: Saharan dust event impacts on cloud formation and radiation over Western Europe, Atmos Chem Phys, 12, 4045-4063, 10.5194/acp-12-4045-2012, 2012.

Bellouin, N., Rae, J., Jones, A., Johnson, C., Haywood, J., and Boucher, O.: Aerosol forcing in the Climate Model Intercomparison Project (CMIP5) simulations by HadGEM2-ES and the role of ammonium nitrate, J Geophys Res-Atmos, 116, 10.1029/2011jd016074, 2011.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevag, A., Seland, O., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., and Kristjansson, J. E.: The Norwegian Earth System Model, NorESM1-M - Part 1: Description and basic evaluation of the physical climate, Geosci Model Dev, 6, 687-720, 10.5194/gmd-6-687-2013, 2013.

Boville, B. A., and Gent, P. R.: The NCAR Climate System Model, version one, J Climate, 11, 1115-1130, Doi 10.1175/1520-0442(1998)011<1115:Tncsmv>2.0.Co;2, 1998.

Chen, M. Y., Xie, P. P., Janowiak, J. E., and Arkin, P. A.: Global land precipitation: A 50-yr monthly analysis based on gauge observations, J Hydrometeorol, 3, 249-266, Doi 10.1175/1525-7541(2002)003<0249:Glpaym>2.0.Co;2, 2002.

Claverie, M., Matthews, J. L., Vermote, E. F., and Justice, C. O.: A 30+ Year AVHRR LAI and FAPAR Climate Data Record: Algorithm Description and Validation, Remote Sens-Basel, 8, 10.3390/rs8030263, 2016.

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S.: Development and evaluation of an Earth-System model-HadGEM2, Geosci Model Dev, 4, 1051-1075, 10.5194/gmd-4-1051-2011, 2011.

Cook, B. I., Miller, R. L., and Seager, R.: Dust and sea surface temperature forcing of the 1930s "Dust Bowl" drought, Geophys Res Lett, 35, 10.1029/2008gl033486, 2008.

Cook, B. I., Miller, R. L., and Seager, R.: Amplification of the North American "Dust Bowl" drought through human-induced land degradation, P Natl Acad Sci USA, 106, 4997-5001, 10.1073/pnas.0810200106, 2009.

Cook, B. I., Seager, R., Miller, R. L., and Mason, J. A.: Intensification of North American Megadroughts through Surface and Dust Aerosol Forcing, J Climate, 26, 4414-4430, 10.1175/Jcli-D-12-00022.1, 2013.

Croft, B., Lohmann, U., and von Salzen, K.: Black carbon ageing in the Canadian Centre for Climate modelling and analysis atmospheric general circulation model, Atmos Chem Phys, 5, 1931-1949, 2005.

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., Golaz, J. C., Ginoux, P., Lin, S. J., Schwarzkopf, M. D., Austin, J., Alaka, G., Cooke, W. F., Delworth, T. L., Freidenreich, S. M., Gordon, C. T., Griffies, S. M., Held, I. M., Hurlin, W. J., Klein, S. A., Knutson, T. R., Langenhorst, A. R., Lee, H. C., Lin, Y. L., Magi, B. I., Malyshev, S. L., Milly, P. C. D., Naik, V., Nath, M. J., Pincus, R., Ploshay, J. J., Ramaswamy, V., Seman, C. J., Shevliakova, E., Sirutis, J. J., Stern, W. F., Stouffer, R. J., Wilson, R. J., Winton, M., Wittenberg, A. T., and Zeng, F. R.: The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3, J Climate, 24, 3484-3519, 10.1175/2011jcli3955.1, 2011.

Dunion, J. P., and Velden, C. S.: The impact of the Saharan air layer on Atlantic tropical cyclone activity, B Am Meteorol Soc, 85, 353-+, 10.1175/Bams-85-3-353, 2004.

Evan, A. T., Dunion, J., Foley, J. A., Heidinger, A. K., and Velden, C. S.: New evidence for a relationship between Atlantic tropical cyclone activity and African dust outbreaks, Geophys Res Lett, 33, 10.1029/2006gl026408, 2006.

Evan, A. T., and Mukhopadhyay, S.: African Dust over the Northern Tropical Atlantic: 1955-2008, J Appl Meteorol Clim, 49, 2213-2229, 10.1175/2010jamc2485.1, 2010.

Evan, A. T., Flamant, C., Fiedler, S., and Doherty, O.: An analysis of aeolian dust in climate models, Geophys Res Lett, 41, 5996-6001, 10.1002/2014gl060545, 2014.

Evan, A. T., Flamant, C., Gaetani, M., and Guichard, F.: The past, present and future of African dust, Nature, 531, 493-+, 10.1038/nature17149, 2016.

Evans, S., Ginoux, P., Malyshev, S., and Shevliakova, E.: Climate-vegetation interaction and amplification of Australian dust variability, Geophys Res Lett, 43, 11823-11830, 10.1002/2016gl071016, 2016.

Fecan, F., Marticorena, B., and Bergametti, G.: Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas, Ann Geophys-Atm Hydr, 17, 149-157, DOI 10.1007/s005850050744, 1999.

Fiedler, S., Schepanski, K., Heinold, B., Knippertz, P., and Tegen, I.: Climatology of nocturnal low-level jets over North Africa and implications for modeling mineral dust emission, J Geophys Res-Atmos, 118, 6100-6121, 10.1002/jgrd.50394, 2013.

Fiedler, S., Knippertz, P., Woodward, S., Martin, G. M., Bellouin, N., Ross, A. N., Heinold, B., Schepanski, K., Birch, C. E., and Tegen, I.: A process-based evaluation of dust-emitting winds in the CMIP5 simulation of HadGEM2-ES, Clim Dynam, 46, 1107-1130, 10.1007/s00382-015-2635-9, 2016.

Fung, I. Y., Meyn, S. K., Tegen, I., Doney, S. C., John, J. G., and Bishop, J. K. B.: Iron supply and demand in the upper ocean, Global Biogeochem Cy, 14, 281-295, Doi 10.1029/1999gb900059, 2000.

Fuss, S., Canadell, J. G., Peters, G. P., Tavoni, M., Andrew, R. M., Ciais, P., Jackson, R. B., Jones, C. D., Kraxner, F., Nakicenovic, N., Le Quere, C., Raupach, M. R., Sharifi, A., Smith, P., and Yamagata, Y.: COMMENTARY: Betting on negative emissions, Nat Clim Change, 4, 850-853, DOI 10.1038/nclimate2392, 2014.

Gillette, D. A., and Passi, R.: Modeling Dust Emission Caused by Wind Erosion, J Geophys Res-Atmos, 93, 14233-14242, DOI 10.1029/JD093iD11p14233, 1988.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., and Lin, S. J.: Sources and distributions of dust aerosols simulated with the GOCART model, J Geophys Res-Atmos, 106, 20255-20273, Doi 10.1029/2000jd000053, 2001.

Ginoux, P., Garbuzov, D., and Hsu, N. C.: Identification of anthropogenic and natural dust sources using Moderate Resolution Imaging Spectroradiometer (MODIS) Deep Blue level 2 data, J Geophys Res-Atmos, 115, 10.1029/2009jd012398, 2010.

Ginoux, P., Clarisse, L., Clerbaux, C., Coheur, P. F., Dubovik, O., Hsu, N. C., and Van Damme, M.: Mixing of dust and NH3 observed globally over anthropogenic dust sources, Atmos Chem Phys, 12, 7351-7363, 10.5194/acp-12-7351-2012, 2012a.

Ginoux, P., Prospero, J. M., Gill, T. E., Hsu, N. C., and Zhao, M.: Global-Scale Attribution of Anthropogenic and Natural Dust Sources and Their Emission Rates

Based on MODIS Deep Blue Aerosol Products, Rev Geophys, 50, 10.1029/2012rg000388, 2012b.

Houghton, J., Ding, Y., Griggs, D. J., Noguer, M., Linden, P. J. v. d., Dai, X., Maskell, K., and Johnson, C. A.: Climate Change 2001: The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, in, Cambridge Univ. Press, Cambridge, UK, 2001.

Hsu, N. C., Tsay, S. C., King, M. D., and Herman, J. R.: Aerosol properties over bright-reflecting source regions, Ieee T Geosci Remote, 42, 557-569, 10.1109/Tgrs.2004.824067, 2004.

Hsu, N. C., Tsay, S. C., King, M. D., and Herman, J. R.: Deep blue retrievals of Asian aerosol properties during ACE-Asia, Ieee T Geosci Remote, 44, 3180-3195, 10.1109/Tgrs.2006.879540, 2006.

Hsu, N. C., Jeong, M. J., Bettenhausen, C., Sayer, A. M., Hansell, R., Seftor, C. S., Huang, J., and Tsay, S. C.: Enhanced Deep Blue aerosol retrieval algorithm: The second generation, J Geophys Res-Atmos, 118, 9296-9315, 10.1002/jgrd.50712, 2013.

Huneeus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, J., Kinne, S., Bauer, S., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Fillmore, D., Ghan, S., Ginoux, P., Grini, A., Horowitz, L., Koch, D., Krol, M. C., Landing, W., Liu, X., Mahowald, N., Miller, R., Morcrette, J. J., Myhre, G., Penner, J., Perlwitz, J., Stier, P., Takemura, T., and Zender, C. S.: Global dust model intercomparison in AeroCom phase I, Atmos Chem Phys, 11, 7781-7816, 10.5194/acp-11-7781-2011, 2011.

Jickells, T. D., An, Z. S., Andersen, K. K., Baker, A. R., Bergametti, G., Brooks, N., Cao, J. J., Boyd, P. W., Duce, R. A., Hunter, K. A., Kawahata, H., Kubilay, N., laRoche, J., Liss, P. S., Mahowald, N., Prospero, J. M., Ridgwell, A. J., Tegen, I., and Torres, R.: Global iron connections between desert dust, ocean biogeochemistry, and climate, Science, 308, 67-71, DOI 10.1126/science.1105959, 2005.

Jin, Q., Wei, J., Yang, Z. L., Pu, B., and Huang, J.: Consistent response of Indian summer monsoon to Middle East dust in observations and simulations, Atmos Chem Phys, 15, 9897-9915, 10.5194/acp-15-9897-2015, 2015.

Jin, Q. J., Wei, J. F., and Yang, Z. L.: Positive response of Indian summer rainfall to Middle East dust, Geophys Res Lett, 41, 4068-4074, 10.1002/2014gl059980, 2014.

Jin, Q. J., Yang, Z. L., and Wei, J. F.: Seasonal Responses of Indian Summer Monsoon to Dust Aerosols in the Middle East, India, and China, J Climate, 29, 6329-6349, 10.1175/Jcli-D-15-0622.1, 2016.

Kim, M. K., Lau, W. K. M., Kim, K. M., Sang, J., Kim, Y. H., and Lee, W. S.: Amplification of ENSO effects on Indian summer monsoon by absorbing aerosols, Clim Dynam, 46, 2657-2671, 10.1007/s00382-015-2722-y, 2016.

Kok, J. F., Ridley, D. A., Zhou, Q., Miller, R. L., Zhao, C., Heald, C. L., Ward, D. S., Albani, S., and Haustein, K.: Smaller desert dust cooling effect estimated from analysis of dust size and abundance, Nat Geosci, 10, 274-+, 10.1038/Ngeo2912, 2017.

Laurent, B., Marticorena, B., Bergametti, G., and Mei, F.: Modeling mineral dust emissions from Chinese and Mongolian deserts, Global Planet Change, 52, 121-141, 10.1016/j.gloplacha.2006.02.012, 2006.

Levin, Z., Ganor, E., and Gladstein, V.: The effects of desert particles coated with sulfate on rain formation in the eastern Mediterranean, J Appl Meteorol, 35, 1511-1523, Doi 10.1175/1520-0450(1996)035<1511:Teodpc>2.0.Co;2, 1996.

Mahowald, N. M., and Luo, C.: A less dusty future?, Geophys Res Lett, 30, 10.1029/2003gl017880, 2003.

Mahowald, N. M., Kloster, S., Engelstaedter, S., Moore, J. K., Mukhopadhyay, S., McConnell, J. R., Albani, S., Doney, S. C., Bhattacharya, A., Curran, M. A. J., Flanner, M. G., Hoffman, F. M., Lawrence, D. M., Lindsay, K., Mayewski, P. A., Neff, J., Rothenberg, D., Thomas, E., Thornton, P. E., and Zender, C. S.: Observed 20th century desert dust variability: impact on climate and biogeochemistry, Atmos Chem Phys, 10, 10875-10893, 10.5194/acp-10-10875-2010, 2010.

Marticorena, B., and Bergametti, G.: Modeling the Atmospheric Dust Cycle .1. Design of a Soil-Derived Dust Emission Scheme, J Geophys Res-Atmos, 100, 16415-16430, Doi 10.1029/95jd00690, 1995.

Marx, S. K., McGowan, H. A., and Kamber, B. S.: Long-range dust transport from eastern Australia: A proxy for Holocene aridity and ENSO-type climate variability, Earth Planet Sc Lett, 282, 167-177, 10.1016/j.epsl.2009.03.013, 2009.

McConnell, J. R., Aristarain, A. J., Banta, J. R., Edwards, P. R., and Simoes, J. C.: 20th-Century doubling in dust archived in an Antarctic Peninsula ice core parallels climate change and desertification in South America, P Natl Acad Sci USA, 104, 5743-5748, 10.1073/pnas.0607657104, 2007.

Miller, R. L., and Tegen, I.: Climate response to soil dust aerosols, J Climate, 11, 3247-3267, Doi 10.1175/1520-0442(1998)011<3247:Crtsda>2.0.Co;2, 1998.

Miller, R. L., Tegen, I., and Perlwitz, J.: Surface radiative forcing by soil dust aerosols and the hydrologic cycle, J Geophys Res-Atmos, 109, 10.1029/2003jd004085, 2004.

Misra, A., Kanawade, V. P., and Tripathi, S. N.: Quantitative assessment of AOD from 17 CMIP5 models based on satellite-derived AOD over India, Ann Geophys-Germany, 34, 657-671, 10.5194/angeo-34-657-2016, 2016.

Moulin, C., and Chiapello, I.: Impact of human-induced desertification on the intensification of Sahel dust emission and export over the last decades, Geophys Res Lett, 33, 10.1029/2006gl025923, 2006.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing. , in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Nakajima, T., Higurashi, A., Kawamoto, K., and Penner, J. E.: A possible correlation between satellite-derived cloud and aerosol microphysical parameters, Geophys Res Lett, 28, 1171-1174, Doi 10.1029/2000gl012186, 2001.

Neff, J. C., Reynolds, R. L., Belnap, J., and Lamothe, P.: Multi-decadal impacts of grazing on soil physical and biogeochemical properties in southeast Utah, Ecol Appl, 15, 87-95, Doi 10.1890/04-0268, 2005.

Neff, J. C., Ballantyne, A. P., Farmer, G. L., Mahowald, N. M., Conroy, J. L., Landry, C. C., Overpeck, J. T., Painter, T. H., Lawrence, C. R., and Reynolds, R. L.: Increasing eolian dust deposition in the western United States linked to human activity, Nat Geosci, 1, 189-195, 10.1038/ngeo133, 2008.

O'Brien, R. M.: A caution regarding rules of thumb for variance inflation factors, Qual Quant, 41, 673-690, 10.1007/s11135-006-9018-6, 2007.

O'Neill, N. T., Eck, T. F., Smirnov, A., Holben, B. N., and Thulasiraman, S.: Spectral discrimination of coarse and fine mode optical depth, J Geophys Res-Atmos, 108, 10.1029/2002jd002975, 2003.

Prospero, J. M., and Nees, R. T.: Impact of the North African Drought and El-Nino on Mineral Dust in the Barbados Trade Winds, Nature, 320, 735-738, DOI 10.1038/320735a0, 1986.

Prospero, J. M., and Lamb, P. J.: African droughts and dust transport to the Caribbean: Climate change implications, Science, 302, 1024-1027, DOI 10.1126/science.1089915, 2003.

Pu, B., Fu, R., Dickinson, R. E., and Fernando, D. N.: Why do summer droughts in the Southern Great Plains occur in some La Nina years but not others?, J Geophys Res-Atmos, 121, 1120-1137, 10.1002/2015jd023508, 2016.

Pu, B., and Ginoux, P.: The impact of the Pacific Decadal Oscillation on springtime dust activity in Syria, Atmos Chem Phys, 16, 13431-13448, 10.5194/acp-16-13431-2016, 2016.

Pu, B., and Ginoux, P.: Projection of American dustiness in the late 21st century due to climate change, Scientific reports, 7, 10.1038/s41598-017-05431-9, 2017.

Reader, M. C., Fung, I., and McFarlane, N.: The mineral dust aerosol cycle during the last glacial maximum (vol 104, pg 9381, 1999), J Geophys Res-Atmos, 104, 22319-22320, Doi 10.1029/1999jd900434, 1999.

Riahi, K., Rao, S., Krey, V., Cho, C. H., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P.: RCP 8.5-A scenario of comparatively high greenhouse gas emissions, Climatic Change, 109, 33-57, 10.1007/s10584-011-0149-y, 2011.

Rosenfield, J. E., Considine, D. B., Meade, P. E., Bacmeister, J. T., Jackman, C. H., and Schoeberl, M. R.: Stratospheric effects of Mount Pinatubo aerosol studied with a coupled two-dimensional model, J Geophys Res-Atmos, 102, 3649-3670, Doi 10.1029/96jd03820, 1997.

Sanap, S. D., Ayantika, D. C., Pandithurai, G., and Niranjan, K.: Assessment of the aerosol distribution over Indian subcontinent in CMIP5 models, Atmospheric Environment, 87, 123-137, 10.1016/j.atmosenv.2014.01.017, 2014.

Seland, O., Iversen, T., Kirkevag, A., and Storelvmo, T.: Aerosol-climate interactions in the CAM-Oslo atmospheric GCM and investigation of associated basic shortcomings, Tellus A, 60, 459-491, 10.1111/j.1600-0870.2008.00318.x, 2008.

Shao, Y. P., Wyrwoll, K. H., Chappell, A., Huang, J. P., Lin, Z. H., McTainsh, G. H., Mikami, M., Tanaka, T. Y., Wang, X. L., and Yoon, S.: Dust cycle: An emerging core theme in Earth system science, Aeolian Res, 2, 181-204, 10.1016/j.aeolia.2011.02.001, 2011.

Sharma, D., and Miller, R. L.: Revisiting the observed correlation between weekly averaged Indian monsoon precipitation and Arabian Sea aerosol optical depth, Geophys Res Lett, 44, 10006-10016, 10.1002/2017gl074373, 2017.

Shindell, D. T., Lamarque, J. F., Schulz, M., Flanner, M., Jiao, C., Chin, M., Young, P. J., Lee, Y. H., Rotstayn, L., Mahowald, N., Milly, G., Faluvegi, G., Balkanski, Y., Collins, W. J., Conley, A. J., Dalsoren, S., Easter, R., Ghan, S., Horowitz, L., Liu, X., Myhre, G., Nagashima, T., Naik, V., Rumbold, S. T., Skeie, R., Sudo, K., Szopa, S., Takemura, T., Voulgarakis, A., Yoon, J. H., and Lo, F.: Radiative forcing in the ACCMIP historical and future climate simulations, Atmos Chem Phys, 13, 2939-2974, 10.5194/acp-13-2939-2013, 2013.

Solmon, F., Nair, V. S., and Mallet, M.: Increasing Arabian dust activity and the Indian summer monsoon, Atmos Chem Phys, 15, 8051-8064, 10.5194/acp-15-8051-2015, 2015.

Strong, J. D., Vecchi, G. A., and Ginoux, P.: The Climatological Effect of Saharan Dust on Global Tropical Cyclones in a Fully Coupled GCM, Journal of Geophysical Research - Atmospheres, 123, https://doi.org/10.1029/2017JD027808, 2018.

Strong, J. D. O., Vecchi, G. A., and Ginoux, P.: The Response of the Tropical Atlantic and West African Climate to Saharan Dust in a Fully Coupled GCM, J Climate, 28, 7071-7092, 10.1175/Jcli-D-14-00797.1, 2015.

Sun, D. L., Lau, K. M., and Kafatos, M.: Contrasting the 2007 and 2005 hurricane seasons: Evidence of possible impacts of Saharan dry air and dust on tropical cyclone activity in the Atlantic basin, Geophys Res Lett, 35, 10.1029/2008gl034529, 2008.

Takemura, T., Okamoto, H., Maruyama, Y., Numaguti, A., Higurashi, A., and Nakajima, T.: Global three-dimensional simulation of aerosol optical thickness distribution of various origins, J Geophys Res-Atmos, 105, 17853-17873, Doi 10.1029/2000jd900265, 2000.

Tegen, I., Harrison, S. P., Kohfeld, K., Prentice, I. C., Coe, M., and Heimann, M.: Impact of vegetation and preferential source areas on global dust aerosol: Results from a model study, J Geophys Res-Atmos, 107, 10.1029/2001jd000963, 2002.

Tegen, I., Werner, M., Harrison, S. P., and Kohfeld, K. E.: Relative importance of climate and land use in determining present and future global soil dust emission, Geophys Res Lett, 31, 10.1029/2003gl019216, 2004.

Todd, M. C., Washington, R., Raghavan, S., Lizcano, G., and Knippertz, P.: Regional model simulations of the Bodele low-level jet of northern Chad during the Bodele Dust Experiment (BoDEx 2005), J Climate, 21, 995-1012, 10.1175/2007jcli1766.1, 2008.

Vinoj, V., Rasch, P. J., Wang, H. L., Yoon, J. H., Ma, P. L., Landu, K., and Singh, B.: Short-term modulation of Indian summer monsoon rainfall by West Asian dust, Nat Geosci, 7, 308-313, 10.1038/ngeo2107, 2014.

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., and Kawamiya, M.: MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments, Geosci Model Dev, 4, 845-872, 10.5194/gmd-4-845-2011, 2011.

Winker, D. M., Hunt, W., and Hostetler, C.: Status and performance of the CALIOP lidar, Bba Lib, 5575, 8-15, 10.1117/12.571955, 2004.

Winker, D. M., Hunt, W. H., and McGill, M. J.: Initial performance assessment of CALIOP, Geophys Res Lett, 34, 10.1029/2007gl030135, 2007.

Wong, S., and Dessler, A. E.: Suppression of deep convection over the tropical North Atlantic by the Saharan Air Layer, Geophys Res Lett, 32, 10.1029/2004gl022295, 2005.

Wurzler, S., Reisin, T. G., and Levin, Z.: Modification of mineral dust particles by cloud processing and subsequent effects on drop size distributions, J Geophys Res-Atmos, 105, 4501-4512, Doi 10.1029/1999jd900980, 2000.

Zender, C. S., and Kwon, E. Y.: Regional contrasts in dust emission responses to climate, J Geophys Res-Atmos, 110, 10.1029/2004jd005501, 2005.

Zhao, T. L., Gong, S. L., Zhang, X. Y., Blanchet, J. P., McKendry, I. G., and Zhou, Z. J.: A simulated climatology of Asian dust aerosol and its trans-Pacific transport. Part

I: Mean climate and validation, J Climate, 19, 88-103, Doi 10.1175/Jcli3605.1, 2006.

Table 1 CMIP5 models used in this study. Models tagged with plus signs (+) ~~considered~~ included anthropogenic land use/land cover change in their vegetation prediction.

Table 2 List of regions selected to compare model output with MODIS DOD. Locations of these regions are also plotted in Fig. 1b. Acronyms are used for some regions for short, and are listed in the brackets in the first column. Note that the region names such as Northern China and India are not exactly the same as their geographical definitions but also covers some areas from nearby countries.

Table 3 Changes of DOD in the late half of the 21st century (2051-2100; RCP 8.5 scenario) from the historical condition (1861-2005) projected by CMIP5 multi-model mean (second to fifth columns) and the regression model (sixth to ninth columns) in the nine regions. Changes of DOD are shown in percentage with reference to CMIP5 multi-model historical run. Note that in some regions the projected change by the regression model is quite large (i.e., greater than ± 100%), largely due to the underestimation of CMIP5 historical run in these regions.

Figure 1. Figure 1. Climatology (2004-2016) of Aqua and Terra combined DOD (i.e., MODIS DOD; top panel) and multi-model mean of CMIP5 DOD (bottom) for four seasons. The pattern correlation (centered; calculated after interpolating ~~MOIDS~~ MODIS DOD to CMIP5 DOD grids) between CMIP5 and MODIS DOD are shown in pink in the bottom panel. Blue numbers denote global mean DOD over land. For CMIP5 model results, ± one standard deviation among seven CMIP5 models is also shown. Black boxes in (b) denote nine averaging regions (Table 2). Here we only added these boxes in (b) instead of every plot to keep the figure clean. Note that CMIP5 multi-model mean is masked by MODIS DOD for comparison. Dotted area in (e)-(h) shows where multi-model mean is greater than one inter-model standard deviation.

Figure 2. Zonal mean DOD from MODIS (thick red), CMIP5 multi-model mean (thick black), and each individual model (other colorful lines).

Figure 3. Seasonal cycle of DOD in nine regions (Table 2) averaged over 2004-2016. Thick red lines denote MODIS DOD, thick black lines denote CMIP5 multi-model mean, and other colorful lines denote individual model output. The annual means from MODIS DOD (Obs; red) and multi-model mean (Ens; black) are shown~~-~~ in each panel. Note that in (i) MODIS DOD (red line) is scaled ten times to better display the season cycle.

Figure 4. Spatial statistics comparing DOD from CMIP5 models with that from MODIS in nine regions. Label on the X-axis shows individual models (1-7) and multi-model mean (8). Y-axis shows the ratio of pattern standard deviations between model climatology (2004-2016) and that of MODIS, which reveals the relative amplitude of the simulated DOD versus satellite DOD. The color denotes pattern correlation (centered) between each model and MODIS DOD in each region.

Figure 5. Correlations (color) between regional averaged time series from CMIP5 DOD and MODIS DOD from 2004 to 2016 for four seasons. Numbers in the X-axis denotes each model (1-7) and multi-model mean (8). Correlations significant at the 90% confidence level are marked by a star and significance at the 95% confidence level by two stars.

Figure 6. Regression coefficients calculated by regressing DOD in each season onto standardized precipitation (purple), bareness (orange), and surface wind speed (green) from 2004 to 2016. Coefficients obtained using MODIS DOD and observed controlling factors (interpolated to a 2° by 2.5° grid) and those using CMIP5 multi-model mean DOD and controlling factors are shown in the left and right columns, respectively. The color of the shading denotes the largest coefficient in absolute value among the three, while the saturation of the color shows the magnitude of the coefficient (from 0 to 0.02). Only regression coefficients significant at the 90% confidence level (Bootstrap test) are shown. Missing values are shaded in grey. To highlight coefficients near the source regions, a mask of LAI ≤ 0.5 is applied.

Figure 7. Projected changes of DOD in the late half of the 21st century (under the RCP 8.5 scenario) from that in the historical level (1861-2005) by CMIP5 multi-model mean for four seasons. The percentage change of global mean (over land) DOD ± one inter-model standard deviation is shown at the bottom of each plot. Areas with sign agreement among the models reaches 71.4% (i.e., at least five out of seven models have the same sign as the multi-model mean) are dotted. ~~one inter-pamong the models reaches 71.4% (i.e., at least five out seven models have the same sign as the multi-model mean) are dotted.~~

Figure 8. Projected difference of (a)-(d) precipitation (mm day-1), (e)-(h) bareness, and (i)-(l) 10 m wind (m s-1) between the late half of the 21st century (2051-2100; RCP 8.5 scenario) and historical level (1861-2005) from multi-model mean of seven CMIP5 models. Areas with sign agreement among the models reaches 71.4% (i.e., at least five out of seven models have the same sign as the multi-model mean) are dotted.

Figure 9. Projected change of DOD in the late half of the 21st century under the RCP 8.5 scenario by the regression model. The results are calculated using the regression coefficients obtained from observations during 2004-2016 (see methodology) and projected changes of precipitation, bareness, and surface wind from seven~~16~~ CMIP5 models. Dotted areas are regions with sign agreement among the regression projections (using output of each of the seven models) above 71.4% (i.e., at least five out of seven regression projections have the same sign as the multi-model mean projection). ~~Dotted areas are regions with sign agreement among the models above 62.5% (i.e., at least 10 out 16 models have the same sign as the multi-model mean).~~ To highlight DOD variations near the source regions, a mask of LAI ≤ 0.5 (from present-day climatology) is applied.

Figure 10. (a)-(d) Projected change of DOD in the late half of the 21st century under the RCP 8.5 scenario by the regression model and output from seven CMIP5 models (same as Fig. 9), and contributions from each component, (e)-(h) precipitation, (j)-(i) bareness, and (m)-(p) surface wind speed. Dotted areas are regions with sign agreement among the models above ~~62.5~~71.4%. To highlight DOD variations near the source regions, a mask of LAI ≤ 0.5 (from present-day climatology) is applied.

Table 1 CMIP5 models used in this study. Models tagged with plus signs (+) ~~considered~~ included anthropogenic land use/land cover change in their vegetation prediction.

Model	**lat/lon resolution**	**Dust emission implementation~~scheme~~**	**Dynamic Vegetation**	**Model reference**
CanESM2	2.8°×2.8°	Reader et al. (1999); Croft et al. (2005)	N+	Arora et al. (2011)
GFDL-CM3	2.0°×2.5°	Ginoux et al. (2001)	Y+	Donner et al. (2011)
HadGEM2-CC	1.2°×1.8°	Marticorena and Bergametti (1995)	Y+	Collins et al. (2011)
HadGEM2-ES	1.2°×1.8°	Marticorena and Bergametti (1995)	Y+	Collins et al. (2011)
MIROC-ESM	2.8°×2.8°	Takemura et al. (2000)	Y+	Watanabe et al. (2011)
MIROC-ESM-CHEM	2.8°×2.8°	Takemura et al. (2000)	Y+	Watanabe et al. (2011)
NorESM1-M	1.9°×2.5°	Seland et al. (2008)	N+	Bentsen et al. (2013)

Table 2 List of regions selected to compare model output with MODIS DOD. Locations of these regions are also plotted in Fig. 1b. Acronyms are used for some regions for short, and are listed in the brackets in the first column. Note that the region names such as n~~N~~orthern China and India are not exactly the same as their geographical definitions but also cover~~s~~ some areas from nearby countries.

Region	Domain
North Africa (N. Africa)	5°-50°N, 18°W-35°E
Middle East	12°-50°N, 35°-60°E
Northern China (N. China)	35°-50°N, 70°-110°E
North America (N. America)	25°-50°N, 95°-125°W
India	5°-35°N, 60°-90°E
Southeastern Asia (SE. Asia)	9°-35°N, 90°-121°E
South Africa (S. Africa)	15°-35°S, 10°-50°N
South America (S. America)	0°-55°S, 60°-83°W
Australia	10°-40°S, 112°-155°E

Table 3 Changes of DOD in the late half of the 21st century (2051-2100; RCP 8.5 scenario) from the historical condition (1861-2005) projected by CMIP5 multi-model mean (second to fifth columns) and the regression model (sixth to ninth columns) in nine regions. Changes of DOD are shown in percentage with reference to CMIP5 multi-model historical run. Note that in some regions the projected change by the regression model is quite large (i.e., greater than ± 100%), largely due to the underestimation of CMIP5 historical run in these regions.

Region	**CMIP5**				**Regression model**			
	DJF	**MAM**	**JJA**	**SON**	**DJF**	**MAM**	**JJA**	**SON**
N. Africa	-3.8	-3.6	2.4	-16.3	-0.8	-17.7	11.1	-10.3
Middle East	7.8	4.5	6.4	1.5	9.8	-16.0	-5.4	-8.4
N. China	-33.5	-11.4	-9.8	-14.4	312.3	-238.6	-51.2	-30.0
N. America	42.6	26.8	13.2	-6.4	-38.5	-90.0	9.3	-42.4
India	-5.1	0.2	-1.0	-9.9	-27.6	-8.2	-2.9	-32.3
SE. Asia	-45.7	-16.5	-13.5	-17.1	-34.8	1.6	4.2	96.3
S. Africa	24.0	6.1	38.5	54.4	22.3	59.3	231.8	78.3
S. America	35.7	27.4	51.8	36.0	14.8	56.1	78.3	154.6
Australia	-3.2	-3.2	15.3	17.0	2.7	0.4	0.7	3.7

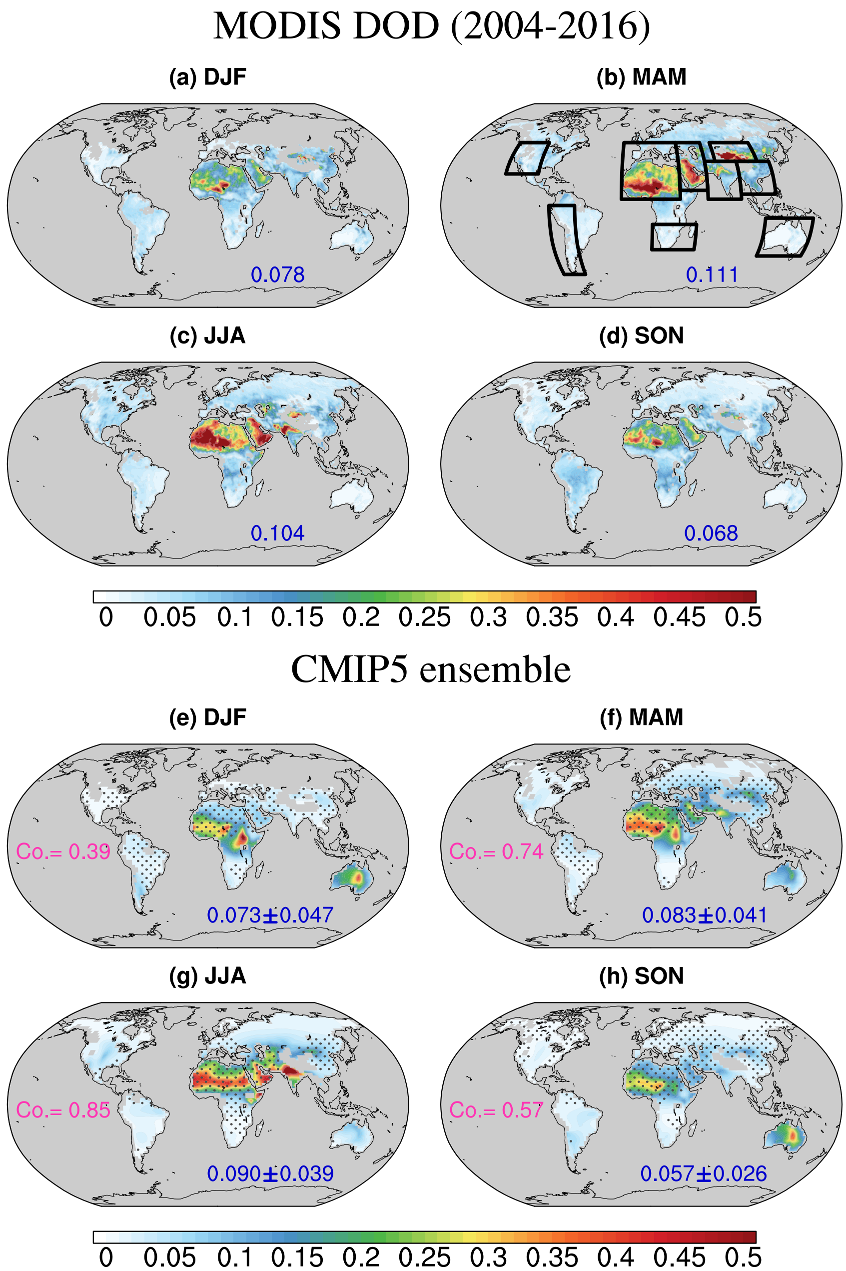


Figure 1. Climatology (2004-2016) of Aqua and Terra combined DOD (i.e., MODIS DOD; top panel) and multi-model mean of CMIP5 DOD (bottom) for four seasons. The pattern correlation (centered; calculated after interpolating ~~MOIDS~~ MODIS DOD to CMIP5 DOD grids) between CMIP5 and MODIS DOD are shown in pink in the bottom panel. Blue numbers denote global mean DOD over land. For CMIP5 model results, ± one standard deviation among seven CMIP5 models is also shown. Black boxes in (b) denote nine averaging regions (Table 2). Here we only added these boxes in (b) instead of every plot to keep the figure clean. Note that CMIP5 multi-model mean is masked by MODIS DOD for comparison. Dotted area in (e)-(h) shows where multi-model mean is greater than one inter-model standard deviation.

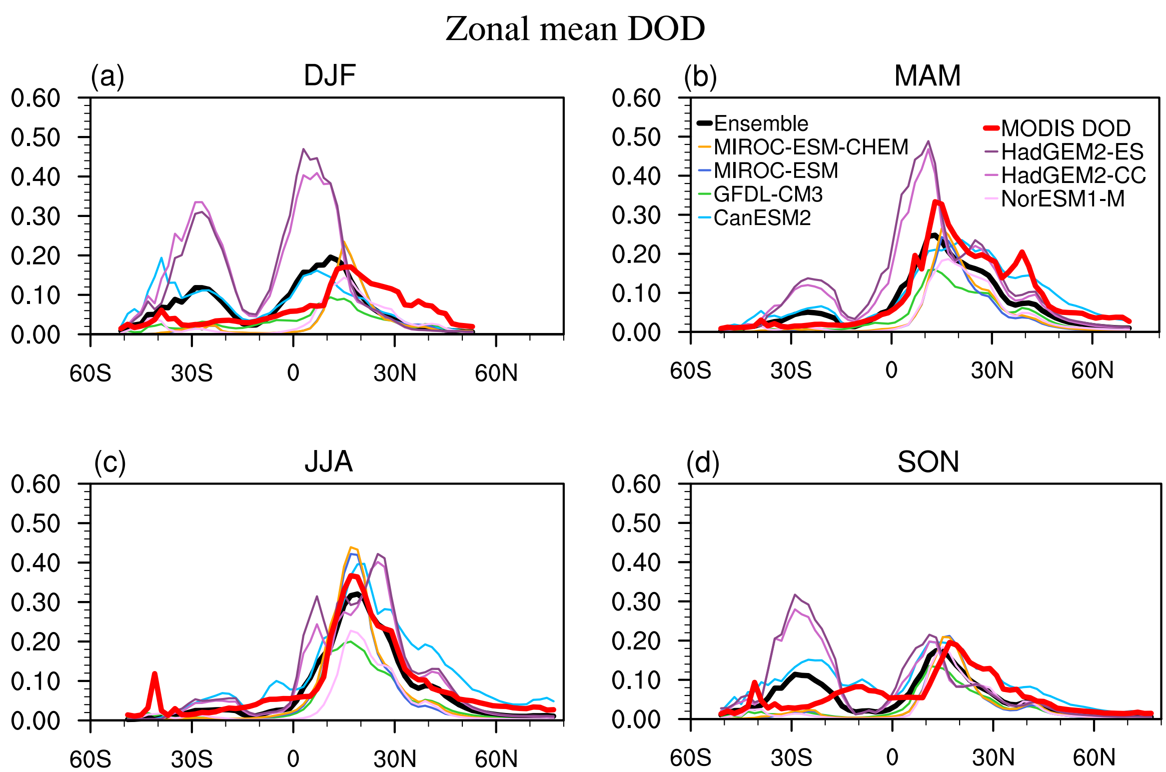


Figure 2. Zonal mean DOD from MODIS (thick red), CMIP5 multi-model mean (thick black), and each individual model (other colorful lines).

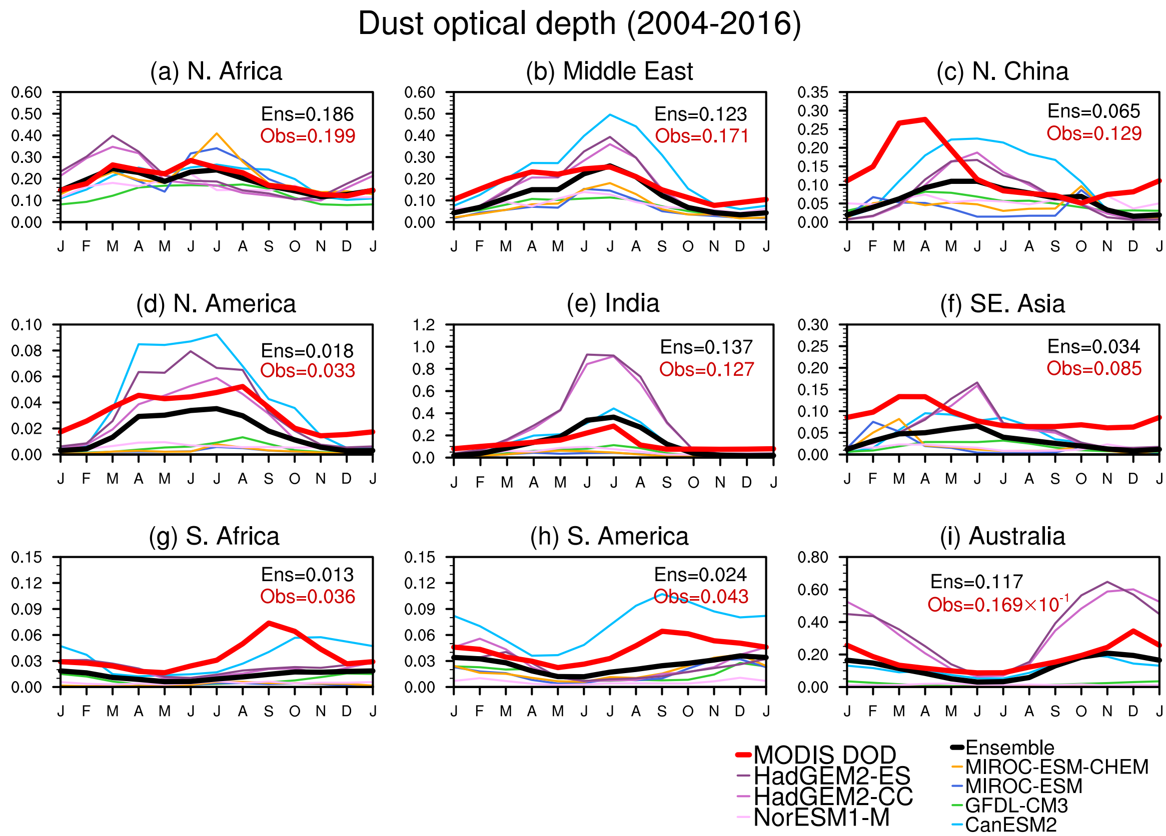


Figure 3. Seasonal cycle of DOD in nine regions (Table 2) averaged over 2004-2016. Thick red lines denote MODIS DOD, thick black lines denote CMIP5 multi-model mean, and other colorful lines denote individual model output. The annual means from MODIS DOD (Obs; red) and multi-model mean (Ens; black) are also listed in each panel. Note that in (i) MODIS DOD (red line) is scaled ten times to better display the season cycle.

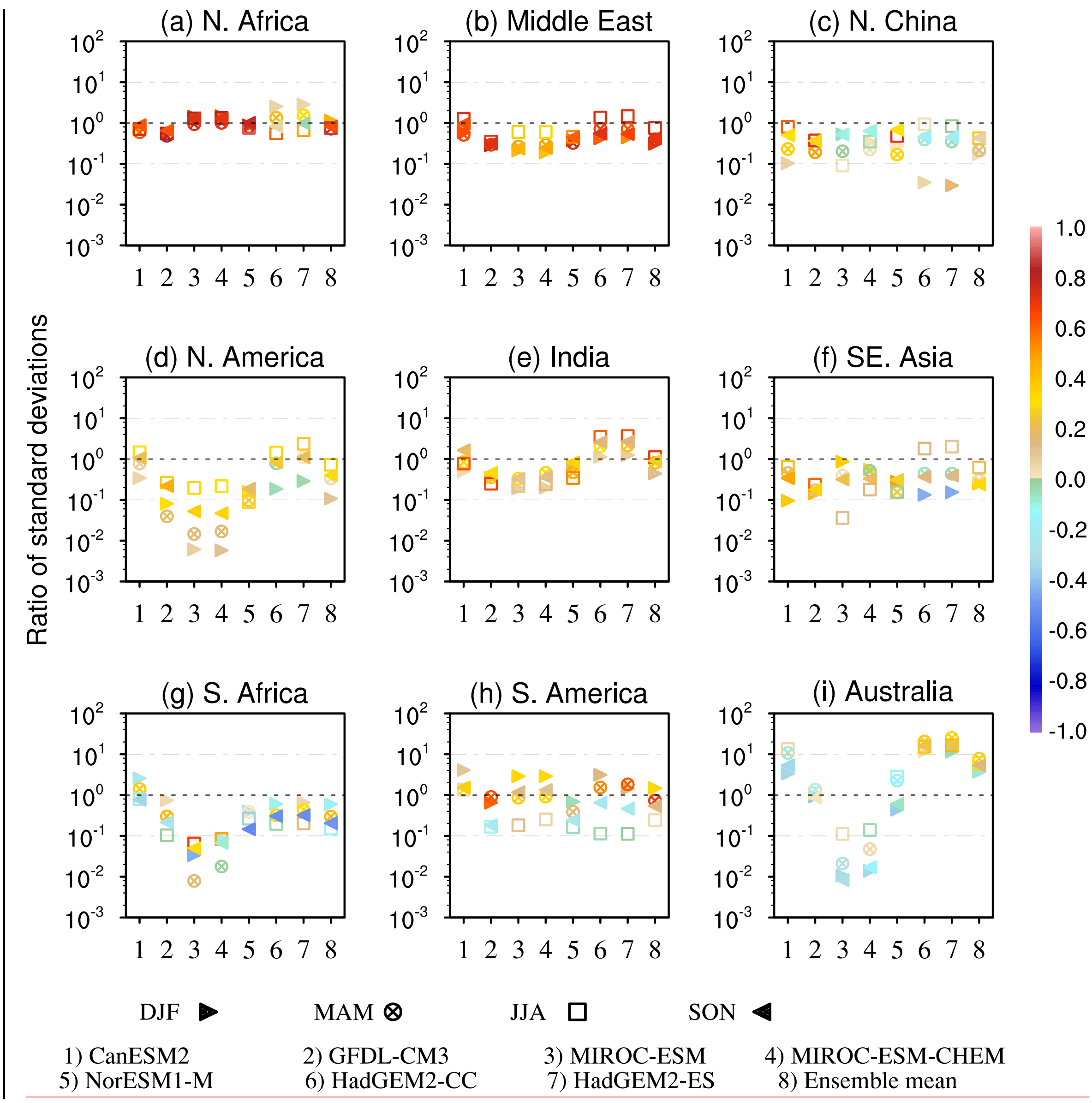


Figure 4. Spatial statistics comparing DOD from CMIP5 models with that from MODIS in nine regions. Label on the X-axis shows individual models (1-7) and multi-model mean (8). Y-axis shows the ratio of pattern standard deviations between model climatology (2004-2016) and that of MODIS, which reveals the relative amplitude of the simulated DOD versus satellite DOD. The color denotes pattern correlation (centered) between each model and MODIS DOD in each region.

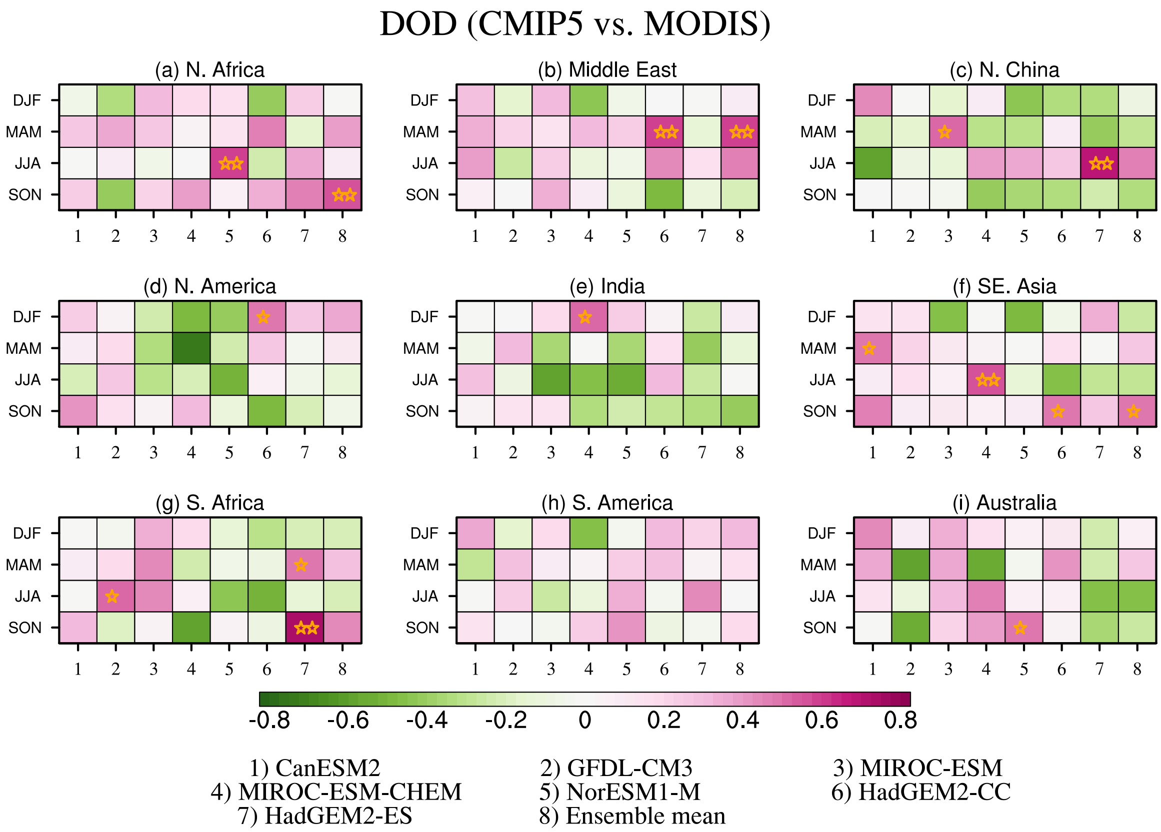


Figure 5. Correlations (color) between regional averaged time series from CMIP5 DOD and MODIS DOD from 2004 to 2016 for four seasons. Numbers in the X-axis denotes each model (1-7) and multi-model mean (8). Correlations significant at the 90% confidence level are marked by a star and significance at the 95% confidence level by two stars.

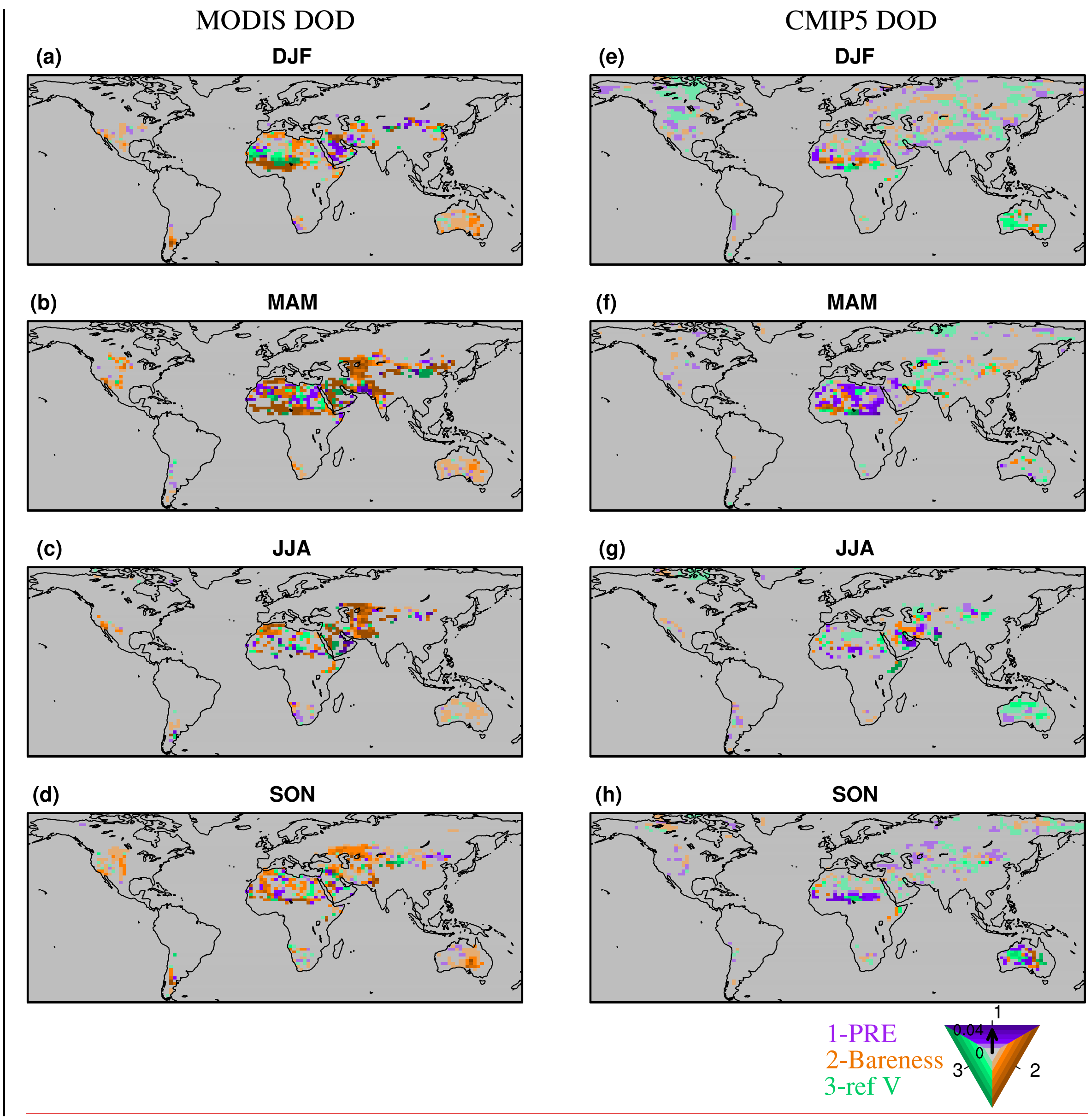


Figure 6. Regression coefficients calculated by regressing DOD in each season onto standardized precipitation (purple), bareness (orange), and surface wind speed (green) from 2004 to 2016. Coefficients obtained using MODIS DOD and observed controlling factors (interpolated to a 2° by 2.5° grid) and those using CMIP5 multi-model mean DOD and controlling factors are shown in the left and right columns, respectively. The color of the shading denotes the largest coefficient in absolute value among the three, while the saturation of the color shows the magnitude of the coefficient (from 0 to 0.04~~2~~). Only regression coefficients significant at the 90% confidence level (Bootstrap test) are shown. Missing values are shaded in grey. To highlight coefficients near dust source regions, a mask of LAI ≤ 0.5 is applied.

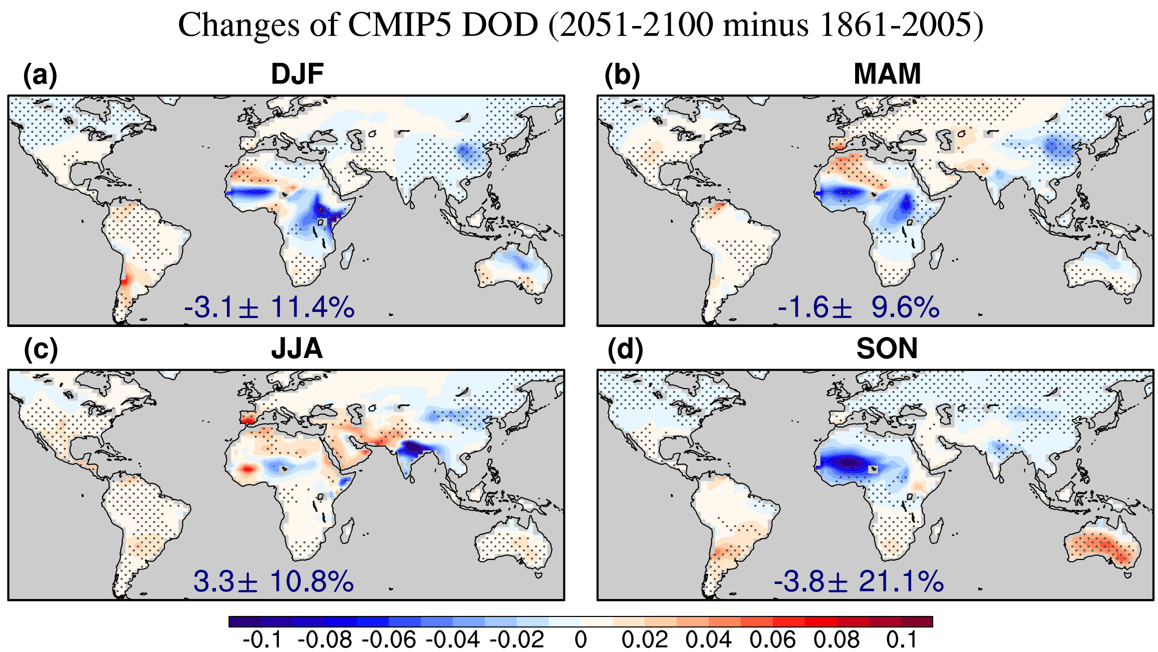


Figure 7. Projected changes of DOD in the late half of the 21st century (under the RCP 8.5 scenario) from that in the historical level (1861-2005) by CMIP5 multi-model mean for four seasons. The percentage change of global mean (over land) DOD ± one inter-model standard deviation is shown at the bottom of each plot. Areas with sign agreement among the models reaches 71.4% (i.e., at least five out of seven models have the same sign as the multi-model mean) are dotted.

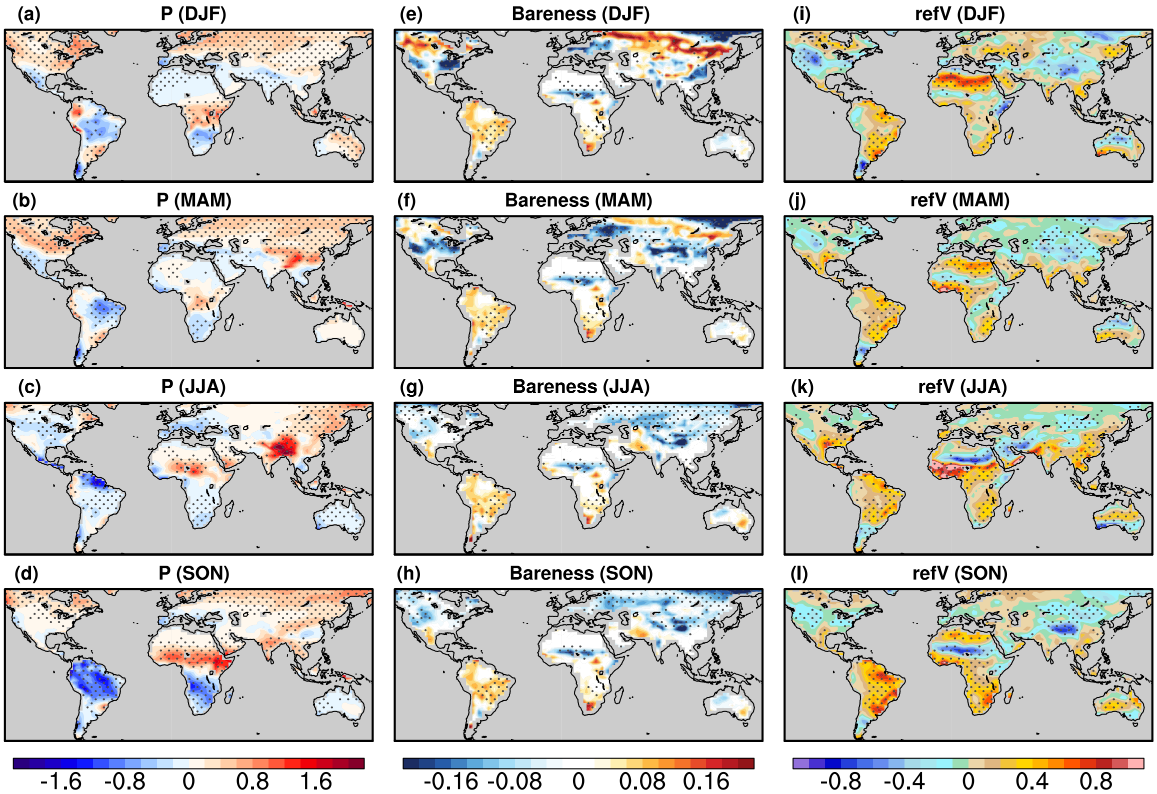


Figure 8. Projected difference of (a)-(d) precipitation (mm day^{-1}), (e)-(h) bareness, and (i)-(l) 10 m wind (m s^{-1}) between the late half of the 21st century (2051-2100; RCP 8.5 scenario) and historical level (1861-2005) from multi-model mean of seven CMIP5 models. Areas with sign agreement among the models reaches 71.4% (i.e., at least five out of seven models have the same sign as the multi-model mean) are dotted.

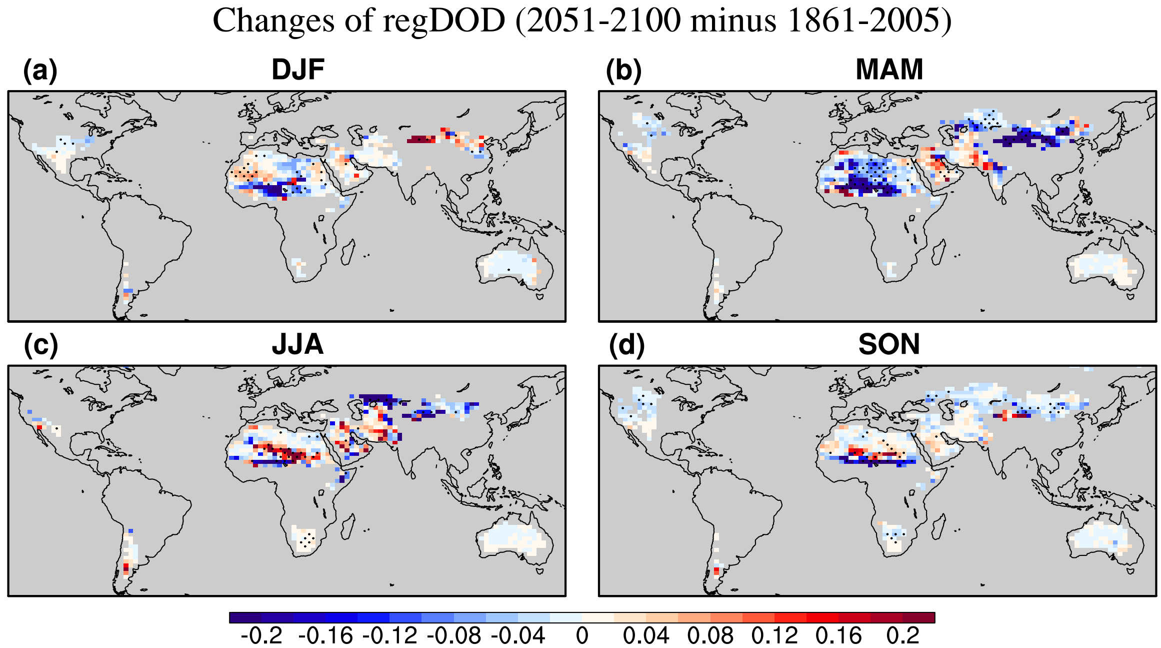


Figure 9. Projected change of DOD in the late half of the 21st century under the RCP 8.5 scenario by the regression model. The results are calculated using the regression coefficients obtained from observations during 2004-2016 (see methodology) and projected changes of precipitation, bareness, and surface wind from ~~16~~ seven CMIP5 models. Dotted areas are regions with sign agreement among the ~~models~~ regression projections (using output of each of the seven models) above ~~62.5~~71.4% (i.e., at least ~~10~~ five out of seven~~16~~ ~~models~~ regression projections have the same sign as the multi-model mean projection). To highlight DOD variations near the source regions, a mask of LAI ≤ 0.5 (from present-day climatology) is applied.

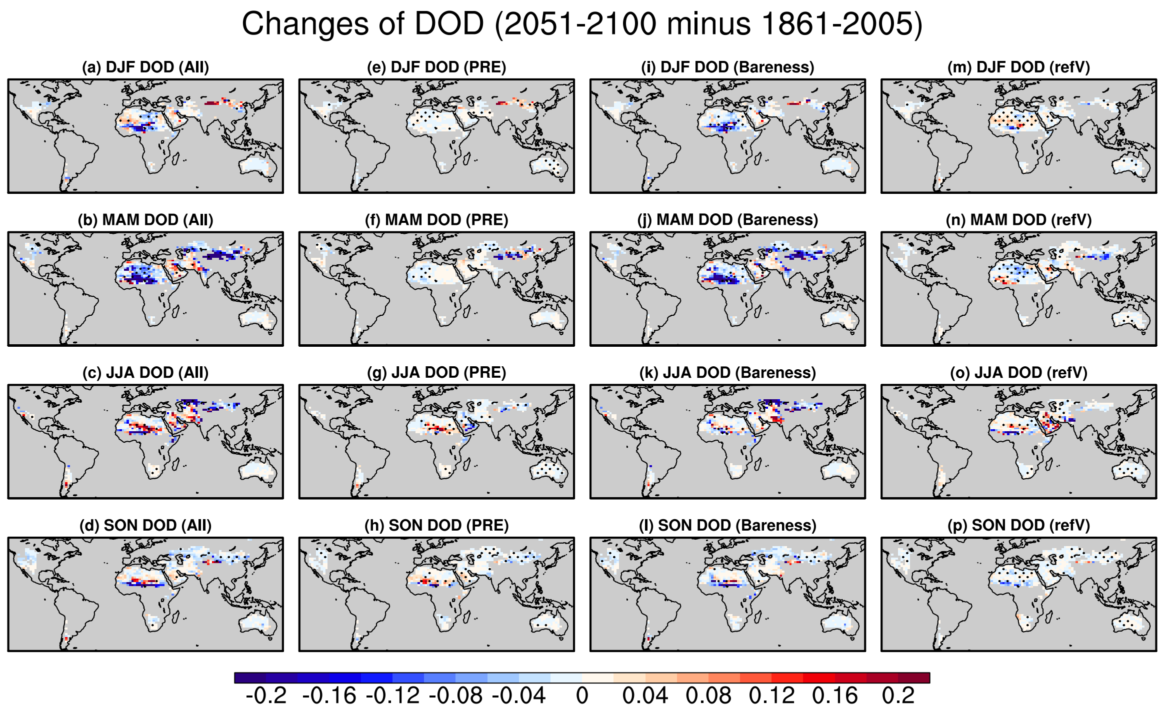


Figure 10. (a)-(d) Projected change of DOD in the late half of the 21st century under the RCP 8.5 scenario by the regression model and output from seven CMIP5 models (same as Fig. 9), and contributions from each component, (e)-(h) precipitation, (j)-(i) bareness, and (m)-(p) surface wind speed. Dotted areas are regions with sign agreement among the ~~models~~ projections above ~~62.5~~71.4%. To highlight DOD variations near the source regions, a mask of LAI ≤ 0.5 (from present-day climatology) is applied.