Reply to Anonymous Referee #1

We thank the reviewer for the careful review and helpful comments. Our responses are detailed below (reviewer's comments marked in blue and our responses in black.

Comment: The major problem is that you should explain your results, not just describe the figures. How is the CRE effect influenced by LWP and CDNC? Why does the non-local effect exist? Especially, why the CRE maximum occurs over the Mexican gulf. These should be discussed and investigated.

Reply: Following the suggestion, the corresponding paragraph has been reorganized and additional description is added. We now explain how changes in cloud droplet number concentration (CDNC) and liquid water path (LWP) result in the negative SCRE in detail. The non-local effect, that is, the tendency of maximum SCRE to appear over the Gulf of Mexico is related to a more sensitive SCRE response to the larger relative change of CDNC and LWP over Gulf of Mexico compared to the land region. As shown in Fig.8 in the original manuscript, changes in both CDNC and LWP are of comparable magnitudes between Gulf of Mexico and the land region. However, given the smaller background CDNC and LWP over Gulf of Mexico, SCRE is more sensitive to changes in the two items over Gulf of Mexico than in the land region. In the revised paper, we have pointed out this phenomenon (Line 315-316) and provided an explanation (Line 337-341).

It now reads (Line 315-316):

"It's interesting to note that the maximum SCRE tends to center around adjacent Gulf of Mexico rather than the land region."

and (Line 320-348):

"To find out the causes of the fire aerosol SCRE, fire aerosol-induced changes in cloud properties are analyzed. Given the largely insignificant change in cloud fraction (Fig. 8), the negative fire aerosol SCRE in the selected regions is mainly associated with increases in cloud droplet number concentrations (CDNC) and liquid water path (LWP). The increased CDNC due to an increase of CCN from fire aerosols (Fig. 8) leads to smaller droplet sizes, which in turn increase cloud albedo

by enhancing backscattering (Twomey, 1977) and further affect LWP by decreasing precipitation efficiency and allowing more liquid water to accumulate (Albrecht, 1989; Ghan et al., 2012). These changes in warm cloud properties demonstrate important contributions of both aerosol first and second indirect effects to the negative SCRE. Over Southern Mexico, although changes of CDNC and LWP are of comparable magnitudes between Gulf of Mexico and the land region (Fig.8), relative changes of both items are much larger over Gulf of Mexico (Fig. S6) due to the smaller magnitudes of background CDNC and LWP here (Fig. S5), which tends to lead to a more sensitive response of SCRE. That's why the maximum SCRE over Southern Mexico is more centered around Gulf of Mexico. Changes in ice water path (IWP) and ice crystal number concentration (ICNC) can also significantly affect SCRE, but with an opposite sign and mostly in the central U.S. The decreased IWP and ICNC, which are possibly caused by fire aerosol-induced changes in the circulation (Ten Hoeve et al, 2012 and reduced coarse mode dust aerosol concentrations), are responsible for the positive SCRE in the north part of central U.S. In the south part of central U.S., the reduction of IWP and ICNC also results in a positive SCRE, which partly offsets the negative SCRE resulting from changes in warm cloud properties. This explains the weaker total negative SCRE in this region compared to the Southern Mexico region despite the more substantial increase in CDNC and LWP here. "

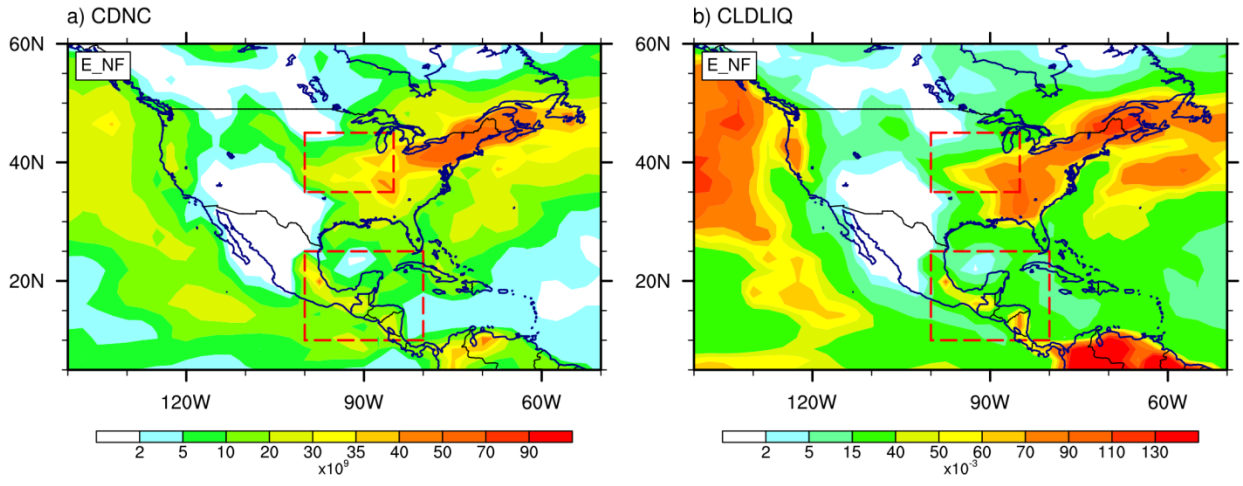

Figure S5. Spatial distributions of 10-day average (Apr. 1-10) ensemble mean a) column-integrated droplet number concentrations (m^{-2}) and b) liquid water path ($g\ m^{-2}$) in the E_NF simulations.

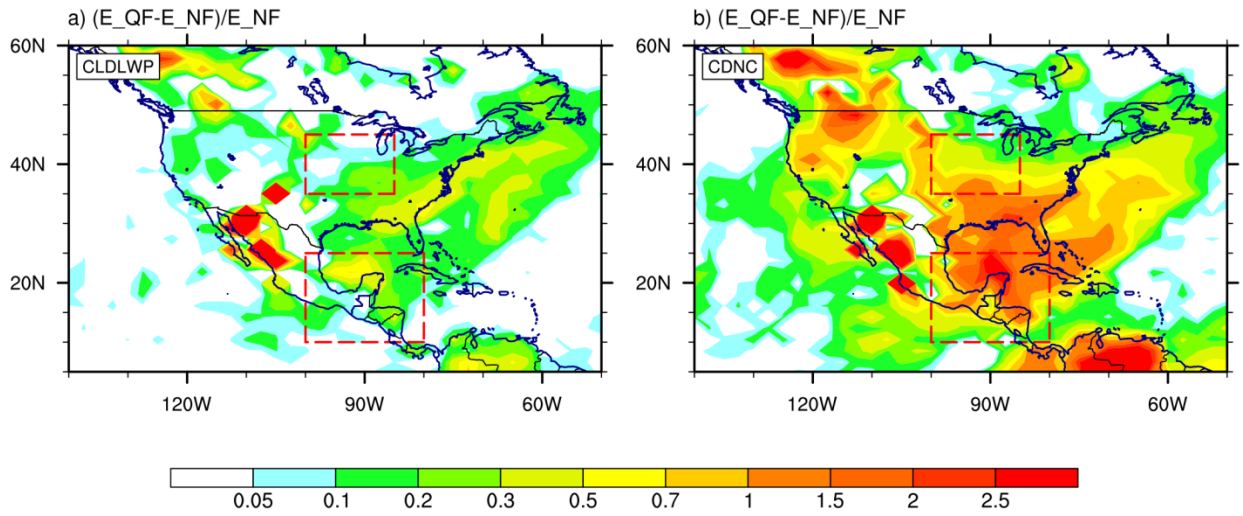

Figure S6. Relative changes of 10-day average ensemble mean cloud properties between the E_NF and E_QF simulations. a) cloud liquid water path, b) column-integrated droplet number concentration

Minor comments:

L60: studied->study

Reply: Done.

L370: Add "are" between "which" and "possibly";

Reply: Done.

L384: Same->The same;

Reply: Done.

L77-79: You should clarify why the indirect effect of fire aerosol deserves study.

Reply: Kaufman et al. (2005) and Zamora et al. (2006) show the short-term indirect effects of fire aerosols are strong based on satellite observations and aircraft measurements (Line 68-72 in the original manuscript). The fire aerosol indirect effect may significantly affect the cloud formation and radiative balance near wildfire burning region. We now explicitly mention the

significant radiative effect of fire aerosol indirect effect in the previous paragraph (Line 65-66) to emphasize this as one motivation of our study. We further pointed out in this paragraph the current lack of model simulations of short-term fire aerosol indirect effects, which is another motivation of our work.

L101-103: What's the difference between nudging horizontal wind and temperature.

Reply: Nudging the horizontal winds will constrain the circulation towards reanalysis, but the thermodynamical features are not directly affected. If temperature is nudged strongly (i.e. use small relaxation time scale) too, the heating/cooling introduced by nudging may affect large scale vertical motion and the parameterized convection. In our study, horizontal wind nudging was applied to constrain the large scale circulations, thus a shorter relaxation time scale of 6 hour is adopted. On the other hand, we only used very weak temperature nudging (much longer relaxation time scale) and perturbed the nudging time scale gently to create ensembles. A much longer relaxation time scale of about 10 days is used. We have clarified this difference in the revised paper. Time scales of wind nudging and temperature nudging are now explicitly provided in the corresponding paragraph. The text reads (Line 92): "*Horizontal winds were nudged towards 6-hourly reanalysis to constrain the large-scale circulation*" and (Line 96): "*we also employed very weak temperature nudging (~10days) in combination with ensembles to quantify the uncertainty. More details of the nudging setup are described in section 2.3.*"

L297-503: How you can get the conclusion that at least 9 members are needed from Fig. 14. You need to quantify the results.

Reply: Thanks for the suggestions. The number 9 in the discussion paper was determined by simple visual comparison. As shown in Fig.14b, discrepancies in the ensemble mean fire aerosol SCRE are substantial when the number of ensemble members (N) is smaller than 8. We agree it is better to determine the minimum required ensemble number in a quantitative way. We now use results from the 20-member ensemble simulations as a reference to evaluate the results from ensemble simulations with varying N. For a specific N, the root mean square error (RMSE) of the ensemble mean SCRE during April 1-10 is chosen to quantify the deviation of the simulated ensemble mean from the reference value. It is calculated as the standard deviation of the differences between the daily ensemble mean SCRE in the N-member simulation and the 20-

member simulation. To get robust results, for each N, we randomly sample N members from the 20 members for 1000 times and evaluated the performance of the 1000 groups (each group has N members) to avoid the influence of limited sampling. Figure 15 shows that both the RMSE of ensemble mean SCRE and the difference of RMSE between the 1000 groups of simulations (for each N) decrease with increasing N. The minimum number of N required is determined when the 90^{th} percentile of RMSE is smaller than a threshold RMSE. Without a good reference, we set the threshold RMSE to 20% (0.566W m^{-2}) of the reference 10-day mean SCRE (-2.83Wm^{-2}). As shown in Fig.15, at least 11 members are needed to meet this criterion. We've refined the conclusion regarding the total number of ensembles needed in the revised paper. The corresponding paragraph now reads (Line 395-408):

"However, discrepancies in the ensemble mean fire aerosol SCRE (Fig. 14b) are substantial when the number of ensemble members is small. The same is true for the ensemble spread of fire aerosol SCRE (Fig.S8).In order to quantify the discrepancies of the simulated SCRE, we chose the ensemble mean SCRE in the 20-member simulation as a reference and use the root mean square errors (RMSE) of the ensemble mean SCRE in the N-member simulation to quantify the deviation of the simulated SCRE from the reference value. It is calculated as the standard deviation of the differences between the daily ensemble mean SCRE in the N-member simulation and the 20-member simulation. For each N, we randomly sampled 1000 times from the 20 members to help reduce the influence from limited sampling. Figure 15 shows that both the RMSE of ensemble mean SCRE and the difference of RMSE between the 1000 groups of simulations (for each N) decrease with increasing N. The minimum number of N required is determined when the 90^{th} percentile of RMSE is smaller than a threshold RMSE. Without a good reference, we set the threshold RMSE to 20% (0.566 W m^{-2}*) of the reference 10-day mean SCRE (-2.83* W m^{-2}*). As shown in Fig.15, at least 11 members are needed to meet this criterion."*

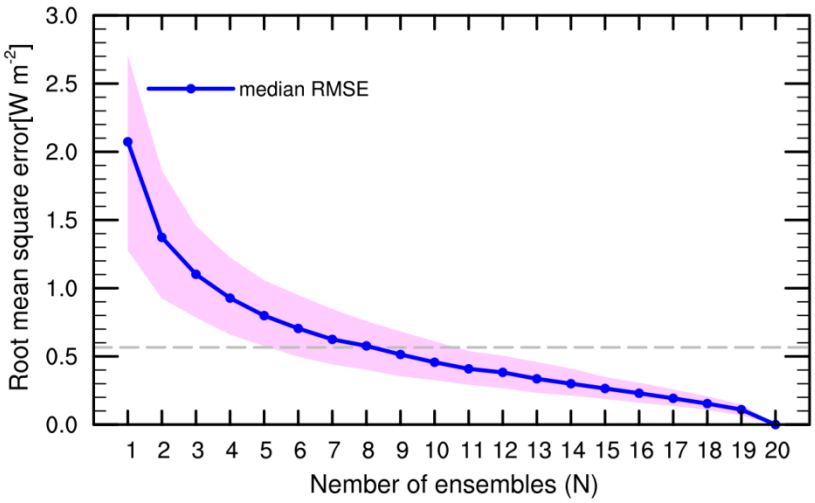

Figure 15 Root mean square errors (RMSE) of the ensemble mean of the regional mean fire aerosol SCRE during April 1-10 over Southern Mexico in simulations with different total number of ensemble members (N). The blue line represents the median RMSE of the 1000 groups (each group has N members/simulations). The grey line represents the threshold RMSE. Shaded area denotes the range between the 10^{th} and 90^{th} percentiles.

Reply to Anonymous Referee #2

We thank the reviewer for the helpful and constructive comments. Our point-by-point responses are provided below (reviewer's comments marked in blue and our responses in black.

This papers studies the direct/indirect aerosol effect from fires using CAM5 with nudged horizontal wind speed and/or nudged temperature. Overall the approach is sound and the paper is well written. Yet it still needs some major clarifications before it is accepted for publication.

General comments:

Since fire emission inventories are critical to this study, please provide a table or a plot to show the BC/OC/SO2 emissions from the 3 different inventories quantitatively. The colorbar in Fig. 2 is difficult to tell how bigger is the QF than the GF3/4. It seems QF emissions are at least 5 times larger than the other two. Please provide a table showing the different fire aerosol forcing components. In the introduction part, 3 different aerosol forcing are mentioned, but only direct forcing and short wave cloud forcing are presented in the result section. Please show long wave forcing and surface albedo forcing as well in the table. The initial condition for the 10-day ensemble runs is generated by S_NF with only u and v being nudged. Temperature is not nudged for the S_NF run. So my question is when you now include slow temperature nudging in the ensemble runs, will they go through some adjustment through the 10-day period? Or in other word, how well is the initial T compared to the T being nudged to on April 1st?

Reply: We have added the Table S1 to show the monthly mean and 10-day mean BC/OC/SO2 emissions from the three inventories. In both regions, the factor difference between the GFED and QFED data is similar no matter whether monthly mean or 10-day mean is used for comparison. Overall, the QFEDv2.4 emissions are larger than GFED v4.1s emissions by a factor of approximately 10 in the Central U.S and a factor of approximately 3 in Southern Mexico. We mentioned in the original manuscript that "*Values of emitted BC in QFED v2.4 are larger than those in GFED v4.1s by a factor of 9.7 in the Central U.S. and a factor of 2.7 in Southern Mexico*", but didn't specify this was based on the 10-day mean emissions. This has been clarified in the revised paper (Line 152-153).

Following the reviewer's suggestion, we have also added Table S3 to show the different components of fire aerosol effects. As shown in Table S3, Shortwave direct radiative effect (SDRE) and shortwave cloud radiative effect (SCRE) are the major contributors to the total fire aerosol effect in both regions. The surface albedo effect is almost negligible. The E_QF simulation shows a significant negative longwave cloud radiative effect (LCRE) in the Central U.S., which is associated with reduced ice crystal number concentrations and ice water path in the north part of this region. We have added a brief description of the negative LCRE (Line 344-345) in the revised paper. Causes of the fire aerosol effect are discussed in detail in Section 3.2.2.

The global pattern of the initial T (or the anomaly) is very similar to those in the reanalysis data but there are some regional differences. It should be noted that the horizontal wind nudging and temperature nudging are applied for different purposes. The horizontal wind nudging is used to constrain large-scale features in simulations with and w/o fire emissions to help better identify radiative forcing signal, while the weak temperature nudging conducted with gently perturbed time scale is used to generate ensemble and investigate the fire forcing uncertainty. Given its large relaxation time scale (10 to 11days compared to the 6-hourly wind nudging), the temperature nudging with small adjustments of temperature has little effect on constraining the temperature field towards reanalysis. We have added the time scale of wind and temperature nudging in Line 92 and Line 96 to clarify the difference.

Table S1 Regional mean emissions of fire aerosols in April, 2009 from three emission inventories (Unit: $x10^{-12}$ kg $m^{-2}s^{-1}$). Numbers in the parentheses show results averaged in April 1-10.

	BC		OC		SO2	
	Central U.S.	Southern Mexico	Central U.S.	Southern Mexico	Central U.S.	Southern Mexico
GFED v3.1	0.25(0.38)	0.69(0.82)	1.82(3.58)	5.60(6.77)	1.35(2.01)	3.69(4.35)
GFED v4.1s	0.23(0.34)	1.17(1.44)	1.75(3.24)	8.80(10.76)	1.21(1.81)	6.25(7.69)
QFED v2.4	2.63(3.29)	3.87(3.87)	23.54(32.25)	36.81(36.58)	14.04(17.59)	20.62(20.65)

Table S3 Regional mean total AOD, fire AOD (differences in AOD between simulations with and without fire) and radiative effects of fire aerosols during April 1-10, 2009 in group B simulations (Unit: W m^{-2}). Total fire aerosol radiative effect is decomposed into shortwave direct radiative effect (SDRE), shortwave cloud radiative effect (SCRE), longwave cloud radiative effect (LCRE) and surface albedo effect (SAE).

	Total AOD	Fire AOD	SDRE	SCRE	LCRE	Total SAE
Central U.S.						
E_NF	0.047					
E_GF3	0.050	0.003	0.02	-0.86	0.04	0.02
E_GF4	0.050	0.003	-0.01	-0.39	0.002	-0.003
E_QF	0.08	0.033	-0.10	-0.56	-0.76	0.12
Southern Mexico						
E_NF	0.135					
E_GF3	0.149	0.014	-0.18	-1.91	-0.21	0.06
E_GF4	0.153	0.018	-0.20	-2.06	-0.23	0.11
E_QF	0.202	0.067	-0.86	-3.02	-0.47	0.14

Specific comments:

Comment: Abstract: why no forcing numbers are provided here. It is expected to see direct forcing and net indirect forcing rather than some changes in the short wave cloud forcing.

Reply: We have added forcing numbers in the revised paper and the text reads (Line 30-32) : *"Strong negative shortwave cloud radiative effect (SCRE) covers almost the entire Southern Mexico with a 10-day regional mean value of -3.02W m^{-2}. Over the Central U.S, comparable positive and negative SCRE of approximately 2W m^{-2} appear in the north and south part of the region respectively. "*

Comment: Line 62-76: Please show what these forcings are. Direct or indirect?

Reply: We have specified direct or indirect effect (Lines 60-67).

Comment: Line 103: Please provide relaxation time for the very weak temperature nudging

Reply: The nudging time scale has been added and it now reads (Line 96): *"Even for short simulations, small perturbations of meteorological states might have large impact on the local aerosol and cloud properties, thus bring uncertainty to the aerosol forcing estimate. Therefore, in our simulations, we also employed very weak temperature nudging (~10days) in combination with ensembles to quantify the uncertainty. More details of the nudging setup are described in Section 2.3."*

Comment: Line 131-132: Please show or elaborate how you convert monthly mean emissions to daily emissions?

Reply: The GFED website (http://www.globalfiredata.org/data.html) provides monthly emission data and daily scalars to convert the monthly emission to daily emission. We have added following descriptions in Section 2.2 (Line 121-123): "*The daily emission data is obtained using daily scalars (http://www.globalfiredata.org/data.html) to distribute monthly emissions over the days and is only available from 2003 onwards."*

Comment: Line 138-142: Does the CAM5 default/background emission already include fire emission? Or did you remove fire emissions from the CAM5's emission files if there is any?

Reply: We didn't use the default climatological monthly mean fire emission in the CAM5 model. We replaced the original fire emission with the daily emission dataset mentioned in this study.

Comment: Line220-226: Please define these forcings with a few sentences rather than refer readers to Ghan 2013.

Reply: We have added a few sentences to elaborate the calculation of fire aerosol effect. Fire aerosol DRE, CRE and surface albedo effect are first defined as changes in total aerosol forcing, cloud forcing and surface albedo forcing between simulations with and without fire emissions. Then, a detailed description of the calculation of total aerosol forcing, cloud forcing and surface albedo forcing in the fire (or no fire) simulation is provided. The text now reads (Line 194-204)

"Similar to Jiang et al. (2016), our calculations are based on the work of Ghan et al. (2012) and Ghan (2013). Fire aerosol DRE, CRE and surface albedo effect are defined as fire induced changes in aerosol forcing, cloud forcing, and surface albedo forcing respectively, and are calculated as the difference of each item between simulations with and without fire emissions (denoted by Δ). In each simulation, aerosol forcing was defined as the difference between all-sky and clean-sky TOA radiative fluxes ($F - F_{clean}$). Cloud forcing was defined as the difference between all-sky and clear sky TOA radiative fluxes under clean-sky conditions ($F_{clean} - F_{clean,clear}$). The rest were related to surface albedo forcing ($F_{clean,clear}$). Thus fire aerosol DRE, CRE, and surface albedo effect were expressed as $\Delta(F - F_{clean})$, $\Delta(F_{clean} - F_{clean,clear})$, and $\Delta F_{clean,clear}$, respectively. More details about the method can be found in section 2 of Ghan (2013). CRE includes contributions of both aerosol indirect effect and aerosol semi-direct effect but was analyzed as a single term (i.e., the sum)."

Comment: Line 238: Please explain what "LEV 2.0 cloud-screened" is.

Reply: The AERONET web page (https://aeronet.gsfc.nasa.gov/) provides both original raw data and processed data at AERONET sites. The LEV 2.0 AOD is the processed data based on a cloud-screening algorithm (Smirnov et al. 2000). This information is now added (Line 214-215): *"LEV 2.0 cloud-screened all points AOD at 500 nm and 675 nm was used to generate hourly AOD at 550 nm, which are the processed data based on a cloud-screening algorithm (Smirnov et al. 2000)"*

Comment: How much does the fire emitted aerosols contribute to the total AOD? It would be helpful to show some estimate of the contributions from fire emitted aerosols and other background aerosols. I realized you presented background AOD and fire AOD later in Fig. 6. But it would be more helpful if you can present some data here when you quote the need to increase the fire emissions by a factor of 1-3. And please explain why increasing the fire emission by a factor 1-3 could then make the simulated AOD large enough to compare with the reanalysis. It seems it is still unlikely to me.

We have added Table S2 showing total AOD, fire AOD and the contributions of fire AOD for reference in the revised paper. The following description is added (Line 245-246): "*Fire aerosols-*

induced AOD increase accounts for 8.1% (S_GF3), 11.2% (S_GF4) and 48.8% (S_QF) of the background AOD (Table S2)."

In the original manuscript, we quoted the need to increase the emissions by a factor of 1-3 mainly in order to show that the underestimation of AOD in simulations driven by GFED data is common. Thus the results in this study are consistent with previous studies. We didn't mean to imply that the scale factor from 1-3 can also be applied in this case. In fact, the scale factors are regionally-specific. In the work of Tosca et al. (2013), they derived the scale factors by computing the ordinary least squares regression between the simulated AOD and the observed AOD (MISR and MODIS) for those months over 1997-2009 that cumulatively contributed to 80% of regional fire emissions. In the selected four regions in their work including South America, South Africa, equatorial Asia and boreal North America, the scale factors range from 1-3. For the selected region and period in this study, larger scaler factors are needed to make simulated AOD in agreement with the NRL and MACC reanalysis data. However, it's worth noting that both reanalysis data shows a positive bias when compared to AERONET data especially in the Central U.S, while a good agreement is found between the simulated AOD in S_QF and AERONET observations (shown in Fig.4 in the original manuscript). Therefore a scale factor that brings the simulated AOD in the GFED forced simulations comparable to that in the QFED forced simulation might be more reasonable. We have performed sensitivity analysis by performing simulations driven by varying magnitudes of fire emissions (0, 1, 2, 5 times the emission). Results show that changes of fire AOD is proportional to changes in fire emissions. Therefore, the conclusion of a scale factor from 1-3 to match observations still holds in Southern Mexico, but a much larger scale factor of about 10 is needed in the Central U.S. We reorganized corresponding sentences in the revised paper to avoid ambiguity (Line 239-241): "*Previous studies have found the underestimation of AOD in simulations with GFED emissions and suggested the need to scale up GFED emissions by a factor of 1-3 to match the observed AOD (Tosca et al., 2013). This is consistent with the large negative bias in the simulation S_GF3 and S_GF4.*" and clarified that (Line 242): "*However, a much larger scale factor might be needed in this case.*"

Table S2 Regional mean total AOD, fire AOD (difference in total AOD between simulations with and without fire) and the contributions of fire AOD (fire AOD divided by total AOD in the S_NF simulation)during April, 2009 in group A simulations.

	Central U.S.			Southern Mexico		
	Total AOD	Fire AOD	Percentage	Total AOD	Fire AOD	Percentage
S_NF	0.066			0.130		
S_GF3	0.068	0.002	3.42%	0.141	0.011	8.10%
S_GF4	0.070	0.004	5.63%	0.145	0.015	11.20%
S_QF	0.099	0.033	49.33%	0.194	0.064	48.84%

Comment: Line 357-358: It is confusing here. Please consider revising.

Reply: The text is revised as follows (Line 317-319): *"In the central U.S, despite moderate fire aerosol SDRE, a positive fire aerosol SCRE exceeding 2 W* m^{-2} *appears in the north part of the region while a comparable negative SCRE appears in the south part of the region."* Discussions on these results are provided in the next paragraph.

Comment: In the simulations with nudged U and V, the circulation is constrained. So it seems the circulation change may be small enough. Then use this to explain to change of ice clouds. I suspect the coarse mode dust number may be smaller and this may contribute to the decrease of produced ice number since the ice nuclei number (dust) is smaller. Need further investigation here.

Reply: We thank the reviewer for the helpful comment. Yes. The coarse mode dust concentration is reduced in the S_QF simulation (Figure S7), thus it is one of the possible causes for the reduced cloud ice amount and cloud ice number concentration. However, note that the location of maximum cloud ice decrease is a bit different from the location of maximum coarse mode dust concentration. In addition, a rough estimate of the change in coarse mode dust number concentration is on the order of $10^{-3}\,\mathrm{kg}^{-1}$ when using dust aerosol density of about 2.5 $\mathrm{g\,cm}^{-3}$and a coarse mode diameter of about $2\,\mu\mathrm{m}$. This means, even assuming that all the

reduced coarse mode dust aerosols are effective ice nuclei, the resulting decrease in cloud ice number concentration from heterogeneous nucleation is still one order smaller than that shown in Figure S7 b).

Although horizontal wind is nudged in this study, changes in the vertical heating profile (e.g. due to BC absorption) would affect the distribution of large-scale vertical velocity. We notice a substantial change in the vertical velocity right below the location with maximum reduced ice cloud. Positive values in Figure S7e) represent downward motions induced by fire aerosols, which lead to a reduced vertical transport of moisture to the upper levels and will suppress the homogeneous ice nucleation there.

We have clarified this in the revised paper and the text now reads (Line 342-344): "*The decreased IWP and ICNC, which are possibly caused by fire aerosol-induced changes in the circulation (Ten Hoeve et al, 2012) and reduced coarse mode dust aerosol concentrations (Fig.S7), are responsible for the positive SCRE in the north part of central U.S*"

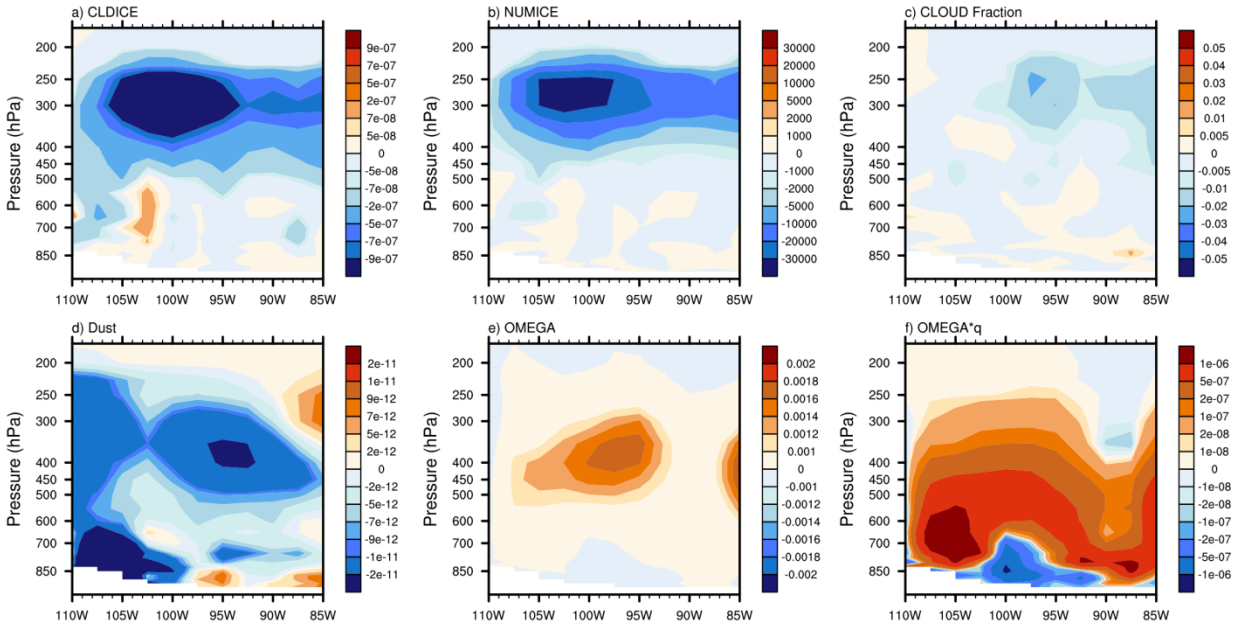

Figure S7. Pressure and longitude distribution of meridional mean (40-45 ° N) difference of 10-day average (April 1 -10) ensemble mean between simulation E_NF and E_QF: a) cloud ice amount ($kg \cdot kg^{-1}$) b) cloud ice number concentration (kg^{-1}) c) cloud fraction (1) d) Coarse mode dust concentration ($kg \cdot kg^{-1}$) e) vertical velocity ($Pa \cdot s^{-1}$) f) vertical moisture transport ($kg \cdot kg^{-1} \cdot Pa \cdot s^{-1}$)

Comment: Line 385**:** make it clear it is SW cloud forcing

Reply: Corresponding sentences have been revised (Line 357-359).

Comment: Line 420: Why quote Fig. 3 here? I think Fig.3 shows results from Group A not from ensemble runs.

Reply: Fig.3 shows the results from both group A (for the whole month) and group B (for the first 10 days) simulations. Results from the ensemble runs are indicated by the range between the maximum and minimum values among ensemble members with darker colors. However, since the simulated AOD is barely distinguishable among members, and between two groups of simulations, the line and the shaded areas almost overlap. The shaded range can be more easily seen in Fig.3 b) by comparing the difference between first 10 days and the rest 20 days. We mentioned in the original manuscript that the shaded areas are very narrow. We have added following sentences in the caption of Fig.3 to further clarify this "*For the single-member simulation and the ensemble simulation driven by same fire emission, the shaded area and the solid line almost overlap given the barely indistinguishable AOD between ensemble members and the corresponding Group A simulation.*"

Comment: Line 430-432: How is the spread calculated for different N? Also how do you select the ensemble member for each different N? I suspect the number 9 required to converge may be different if the ensemble members for different N are constructed differently

Reply:

1) The spread is calculated as the sample standard deviation of ensemble members. For a specific N (number of ensemble members), values of ensemble members are denoted as Xi (i = 1, 2, …N). Ensemble mean is calculated as $\overline{X} = \frac{1}{N}\sum_{i=1}^{N} Xi$. Ensemble spread is calculated as $\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(Xi - \overline{X})^2}$.
2) We performed a group of simulations with a total of 20 members. They are generated by implementing a weak temperature nudging with relaxation time scale ranging from (10 ~11.9 days with an interval of 0.1 day). For a specific N, the first N members are chosen

in calculation, that is, members with relaxation time scale ranging from (10 ~ 10 + 0.1(N-1) days).

3) We thank the reviewer for this insightful comment, which has allowed us to further validate the robustness of our results. A quantitative way is now provided to determine the minimum numbers of ensemble members required in this case study. We now use results from the 20-member ensemble simulations as a reference to evaluate the results from ensemble simulations with varying N. For a specific N, the root mean square error (RMSE) of the ensemble mean SCRE during April 1-10 is used to quantify the deviation of the simulated ensemble mean from the reference value. It is calculated as the standard deviation of the differences between the daily ensemble mean SCRE in the N-member simulation and the 20-member simulation. To get robust results, for each N, we randomly sample N members from the 20 members for 1000 times and evaluated the performance of the 1000 groups (each group has N members) to avoid the influence of limited sampling. Figure 15 shows that both the RMSE of ensemble mean SCRE and the difference of RMSE between the 1000 groups of simulations (for each N) decrease with increasing N. The minimum number of N required is determined when the 90^{th} percentile of RMSE is smaller than a threshold RMSE. Without a good reference, we set the threshold RMSE to 20% (0.566W m^{-2}) of the reference 10-day mean SCRE (-2.83 W m^{-2}). As shown in Fig.15, at least 11 members are needed to meet this criterion. We've refined the conclusion regarding the total number of ensembles needed in the revised paper. Corresponding paragraph has been rewritten and now reads (Line 395-408):

"However, discrepancies in the ensemble mean fire aerosol SCRE (Fig. 14b) are substantial when the number of ensemble members is small. The same is true for the ensemble spread of fire aerosol SCRE (Fig.S8).In order to quantify the discrepancies of the simulated SCRE, we chose the ensemble mean SCRE in the 20-member simulation as a reference and use the root mean square errors (RMSE) of the ensemble mean SCRE in the N-member simulation to quantify the deviation of the simulated SCRE from the reference value. It is calculated as the standard deviation of the differences between the daily ensemble mean SCRE in the N-member simulation and the 20-member simulation. For each N, we randomly sampled 1000 times from the 20 members to help reduce the influence from limited sampling. Figure 15 shows that both the RMSE of ensemble

mean SCRE and the difference of RMSE between the 1000 groups of simulations (for each N) decrease with increasing N. The minimum number of N required is determined when the 90th percentile of RMSE is smaller than a threshold RMSE. Without a good reference, we set the threshold RMSE to 20% (0.566$\mathrm{Wm^{-2}}$*) of the reference 10-day mean SCRE (-2.83*$\mathrm{W\ m^{-2}}$*). As shown in Fig.15, at least 11 members are needed to meet this criterion."*

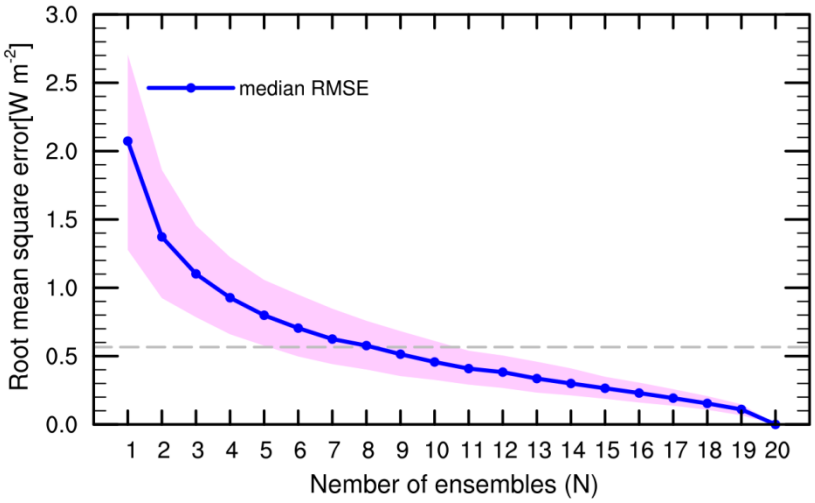

Figure 15 Root mean square errors (RMSE) of the ensemble mean of the regional mean fire aerosol SCRE during April 1-10 over Southern Mexico in simulations with different total number of ensemble members (N). The blue line represents the median RMSE of the 1000 groups (each group has N members/simulations). The grey line represents the threshold RMSE. Shaded area denotes the range between the 10th and 90th percentiles.

Comment: Fig3: Please give full name of TCC.

Reply: Full name of TCC is temporal correlation coefficient. We now provide the full name in the revised paper.

Comment: Fig7: what is the KS test? Please give full name

Reply: Full name of the KS test is Kolmogorov-Smirnov test. We now provide the full name in the revised paper.

Comment: Fig10: is a) total aerosol direct forcing?

Reply: Yes. We have clarified this in the revised paper.

Investigating the short-term effective radiative forcing of fire aerosols using nudged hindcast ensembles

Yawen Liu[1, 2], Kai Zhang[2], Yun Qian[2], Yuhang Wang[3], Yufei Zou[3], Yongjia Song[3], Hui Wan[2], Xiaohong Liu[4], and Xiu-Qun Yang[1]

[1] School of Atmospheric Sciences, Nanjing University, Nanjing, China

[2]Pacific Northwest National Laboratory, Richland, Washington, USA

[3] School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA

[4] Department of Atmospheric Science, University of Wyoming, Laramie, Wyoming, USA

Corresponding to: Yun Qian [Yun.Qian@pnnl.gov]

Abstract

Aerosols from fire emissions can potentially have large impact on clouds and radiation. However, fire aerosol sources are often intermittent and their effect on weather and climate is difficult to quantify. Here we investigated the short-term effective radiative forcing of fire aerosols using the global aerosol-climate model Community Atmosphere Model Version 5 (CAM5). Different from previous studies, we used nudged hindcast ensembles to quantify the forcing uncertainty due to the chaotic response to small perturbations in the atmosphere state. Daily mean emissions from three fire inventories were used to consider the uncertainty in emission strength and injection heights. The simulated aerosol optical depth (AOD) and mass concentrations were evaluated against in-situ measurements and re-analysis data. Overall, the results show the model has reasonably good predicting skills. Short (10-day) nudged ensemble simulations were then performed with and without fire emissions to estimate the effective radiative forcing. Results show fire aerosols have large effects on both liquid and ice clouds over the two selected regions in April 2009. Ensemble mean results show strong negative shortwave cloud radiative effect (SCRE) over almost the entire Southern Mexico, with a 10-day regional mean value of -3.02 W m^{-2}. Over the Central U.S, the SCRE is positive in the north but negative in the south and the regional mean SCRE is small (-0.56 W m^{-2}). For the 10-day average, we found a large ensemble spread of regional mean shortwave cloud radiative effect over Southern Mexico (15.6% of the corresponding ensemble mean) and the Central U.S. (64.3%), despite that the regional mean AOD time series are almost indistinguishable during the 10-day period. Moreover, the ensemble spread is much larger when using daily averages instead of 10-day averages. ~~For the case investigated here, a minimum of 9 ensemble members is necessary to get a reasonable estimate of the ensemble mean and spread of the forcing on individual days.~~ This demonstrates the importance of using a large ensemble of simulations to estimate the short-term aerosol effective radiative forcing.

1. Introduction

Natural and human-induced fires play an important role in the Earth system. Aerosol and gas emissions from biomass burning can change the atmospheric composition and potentially affect the weather and climate. Over 30% of the global total emission of black carbon (BC) comes from open burning of forests, grasslands and agricultural residues (Bond et al. 2013). For organic aerosols, substantial increases of concentrations dominated by organic carbon enhancements are observed in regions with biomass burning events (Zeng et al. 2011; Lin et al. 2013; Brito et al. 2014; Reddington et al. 2014). As a result, biomass burning emissions have a large impact on the global and regional mean aerosol optical depth (Jacobson, 2014).

Through interactions with radiation and cloud, fire aerosols can significantly affect the long-term Earth's energy budget. Previous studies have investigated the global and regional radiative forcing of fire aerosols using long climatological simulations or satellite retrievals. For example, Ward et al. (2012) investigated the radiative forcing of global fires in pre-industrial, present day, and future periods. For the present-day condition, they estimated a direct aerosol effect (or radiative forcing through aerosol–radiation interactions as defined in IPCC AR5, RFari; see section 2.4) of +0.1W m^{-2} and an indirect effect (radiative forcing through aerosol–cloud interactions, RFaci) of -1.0W m^{-2} . Using a newer model, Jiang et al. (2016) found similar RFari but slightly smaller RFaci (-0.70W m^{-2}). Sena et al.

(2013) assessed the direct impact of biomass burning aerosols over the Amazon basin using satellite data. Over the 10-year ~~studied~~ study period, the estimated radiative forcing is about -5.6W m^{-2} .

On short timescales, fire aerosols have even larger radiative impacts. Observed maximum daily direct aerosol radiative effects can reach -20W m^{-2} at TOA locally in Amazonia during biomass burning seasons (Sena et al., 2013). Very large direct effects of fire aerosols were observed during extreme fire events over Central Russia (Tarasova et al. 2004; Chubarova et al. 2008; Chubarova et al. 2012). Instantaneous direct radiative effects of emitted aerosols reached -167 W m^{-2} and monthly mean direct radiative effects reached about -65 W m^{-2} in the 2010 Russia wildfires (Chubarova et al. 2012). Kolusu et al. (2015) investigated direct radiative effect of biomass burning aerosols over tropical Southern America. By quantifying results from the first and second day of 2-day single-member forecasts in September 2012, they found the modeled biomass burning aerosols reduced all-sky net radiation by 8 W m^{-2} at TOA and 15 W m^{-2} at surface Fire aerosol indirect effect may also significantly affect the cloud formation and radiative balance on short time scales. Using satellite data and a radiative transfer model, Kaufman et al. (2005) found an indirect radiative effect of -9.5W m^{-2} due to smoke aerosol-induced cloud changes over Southeast Atlantic for the 3 months studied. Smoke-derived cloud albedo effect on local shortwave radiative forcing is estimated to be between -2 and -4 W m^{-2} in a day case study of aircraft-measured indirect cloud effects (Zamora et al., 2016). ~~Kolusu et al. (2015) investigated direct radiative effect of biomass burning aerosols over tropical Southern America. By quantifying results from the first and second day of 2-day single-member forecasts in September 2012, they found the modeled biomass burning aerosols reduced all-sky net radiation by 8 W m^{-2} at TOA and 15 W m^{-2} at surface~~.

Previous modeling studies on the short-term fire aerosol effects mainly focused on aerosol direct effects (e.g., Keil and Haywood, 2003; Chen et al., 2014; Kolusu et al., 2015), and only a couple of studies investigated the indirect effects of fire aerosols (Lu et al. 2013). In addition, to estimate the aerosol indirect effect, long simulations (multi-years, >5 years preferred) are often needed to remove the noise, because aerosol life cycle and cloud properties are affected by strong natural variability on different timescales (Bony et al. 2006; Kooperman et al. 2012). To solve the problem, alternative methods have been proposed to help extract signals with shorter simulations. For example, nudging (also called Newton relaxation method) can help reduce uncertainties associated with natural variability by constraining certain meteorological fields towards prescribed conditions. A robust estimate of global anthropogenic aerosol indirect effects can be obtained on substantially shorter timescales (1-2 years) by implementing nudging to constrain simulations with pre-industrial and present-day aerosol emissions toward identical circulation and meteorology (Kooperman et al. 2012). When nudged towards re-analysis data, Zhang et al. (2014) found constraining only the horizontal winds is a preferred strategy to estimate the aerosol indirect effect since it provides well-constrained meteorology without strongly perturbing the model's mean climate state. Another example is the use of representative ensembles of short simulations to replace a typical long integration. Wan et al. (2014) explored the feasibility of this method and showed that 3-day ensembles of 20 to 50 members are able to reveal the main signals revealed by traditional 5-year simulations.

In this study, we performed month-long and 10-day nudged CAM5 simulations to investigate the effects of fire aerosols on radiation and cloud processes on short time scales (less than two weeks). Horizontal winds were nudged towards 6-hourly reanalysis to constrain the large-scale circulation and to allow for more accurate model evaluations

against observations. We also used daily mean emissions from three fire inventories to consider the uncertainty in emission strength and injection heights. Even for short simulations, small perturbations of meteorological states might have large impact on the local aerosol and cloud properties, thus bring uncertainty to the aerosol forcing estimate. Therefore, in our simulations, we also employed very weak temperature nudging (~10days) in combination with ensembles to quantify the uncertainty. More details of the nudging setup are described in section2.3.

The rest of the paper is organized as follows. Sect. 2 describes the model and data used in this study. It also introduces how the ensembles are generated in the short nudged simulations and explains how the fire aerosol forcing is estimated. Results and discussions are presented in Sect. 3 and conclusions are summarized in Sect. 4

2. Model, Method and Data

2.1 Model description

In this study, we used the Community Atmosphere Model (CAM) version 5.3 with the finite volume dynamical core at 1.9° (latitude) × 2.5° (longitude) horizontal resolution with 30 vertical layers. The aerosol life cycle is represented using the modal aerosol module MAM3 (Liu et al., 2012). CAM5 links the simulated aerosol fields with cloud and radiation through interactions of the aerosol module with the cloud microphysics and radiative transfer parameterizations. The two-moment bulk cloud microphysics scheme from Morrison and Gettelman (2008) is used to track mass mixing ratios and number concentrations of cloud droplets and ice crystals in stratiform clouds. Representation of shallow convection is based on the work of Park and Bretherton (2009). The deep convection parameterization was developed by Zhang and McFarlane (1995) and later revised by Richter and Rasch (2008) and Neale et al. (2008). Longwave and shortwave radiative transfer are calculated with the Rapid Radiative Transfer Model for GCMs (RRTMG, Malwer et al. 1997; Iacono et al. 2008).

2.2 Fire Emission Inventories

Three fire emission inventories were used in this study. Two of them are widely used bottom-up inventories— Global Fire Emissions Database version 3.1 (GFED v3.1, van der Werf et al., 2010; https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1191) and GFED v4.1s (Giglio et al. 2013; Randerson et al. 2012; https://daac.ornl.gov/VEGETATION/guides/fire_emissions_v4.html). Another one is a top-down emission inventory—Quick Fire Emissions Dataset version 2.4 (QFED v2.4). GFED v3.1 and GFED v4.1s provide global monthly emissions at 0.25×0.25 degree spatial resolution from 1997 through the present. Daily emission data ~~are~~can be obtained by disaggregating monthly emissions based on daily temporal variability in fire emissions derived from MODIS measurements of active fires (Mu et al. 2011). The daily emission data is obtained using daily scalars (http://www.globalfiredata.org/data.html) to distribute monthly emissions over the days and is only available from 2003 onwards. The more recent version GFED v4.1s improves by including small fires based on active fire detections outside the burned area maps (Randerson et al., 2012). QFED v2.4 estimates global fire emissions using the Moderate Resolution Imaging Spectroradiometer (MODIS) measurements of fire radiative power and generates daily products at 0.1×0.1 degree resolution.

To drive CAM5 simulations, fire emission data were regridded to the model resolution and distributed vertically. For the GFED v3.1 and QFED v2.4 emission data we adopted the same injection heights (from surface to 6 km) as used in the standard CAM5 model. While for GFEDv4.1s, in this study the injection heights were estimated using a fire plume model and scaled to the 6-hourly interval.

The fire emission inventories were first analyzed to select appropriate time periods and regions for our study before being used to drive model simulations. Fig.1 shows the multi-year mean biomass burning emissions from GFED v4.1 over North America. The emission manifests significant seasonality with large dry matter consumption during March to April and June to September. The summer and autumn burning covers Pacific Northwest and part of Canada and is mainly associated with forest fires, while the spring burning occurs in more densely populated regions like Mexico and central and eastern United States with a large contribution of agricultural fires in croplands (Korontzi et al., 2006; Magi et al., 2012). Similar features are also captured in GFED v3.1 and QFED v2.4 with differences in the magnitude. We chose to analyze the simulated fire aerosol effect in April, the peak month of spring burning, when there are extreme fire activities over Mexico (10 N to 25N, 100W to 80W) and occasionally large fires in the Central U.S. (35 N to 45N, 100W to 85W). For the U.S., extended fire period is rare, making it necessary to perform short-term evaluation. Fire aerosols formed from these two regions are often transported to the Eastern and Southeastern U.S., where they mix with aerosols from anthropogenic sources and potentially have significant impact on clouds and radiation over these areas. Time series of regional mean fire emissions in April during 2003-2014 shows that relatively large fires occur in both regions in 2009 (Fig.S1). Values of fire emissions in 2009 are larger than the multi-year April mean by a factor of 1.9 in the Central U.S. and 1.5 in Southern Mexico. Thus, in the following model simulations, we focused on analyzing the aerosol properties and radiative effects over the two selected regions (denoted by the red boxes in Fig.1) in April, 2009.

Fire emitted BC from different emission inventories in April, 2009 is shown is Fig.2. Although GFED v4.1s includes the contributions of small fires (Randerson et al., 2012), the emitted BC in GFED v4.1 shows no substantial increase compared to GFED v3.1during the selected period. Only an increase by 1.75 is seen over Southern Mexico. In the Central U.S., the BC emission is even slightly weaker in GFED v4.1. QFED v2.4 shows a much larger BC emission than the GFED inventories. Monthly mean ~~Values~~ -values of emitted BC in QFED v2.4 are larger than those in GFED v4.1s by a factor of ~~9.7~~ 11.4 in the Central U.S. and a factor of ~~2.7~~ 3.3 in Southern Mexico.

2.3 Simulations

Two groups of simulations were conducted (Table1) using the same greenhouse gas concentrations, sea surface conditions and anthropogenic emissions of aerosols and precursors. Each group includes four simulations, performed either without fire emission or with daily fire emissions from one of the three fire emission inventories introduced in section 2.2. The emitted species include BC, OC, and SO2. Horizontal winds were nudged to 6-hourly ERA-Interim reanalysis (Dee et al., 2011) as described in Zhang et al. (2014) in both groups.

Simulations in Group A are month-long single-member nudged simulations. These simulations were performed to provide longer time series for model evaluation and generate initial condition files for simulations in Group B. They

started from January 1, 2009 and were integrated for four months with 3-month spin-up. Initial condition files were generated on April 1 at 00 UTC for simulations in group B.

Simulations in group B are 10-day ensemble simulations. Unlike the traditional way of perturbing initial conditions, in this study we constructed the ensembles by implementing a very weak temperature nudging and perturbing the nudging time scale. This is because under the influence of horizontal-wind nudging, ensemble differences generated by perturbing initial conditions would fade away during the integration. In contrast, our method can consider the influence of small temperature perturbations during the entire simulation period, as nudging is applied at every time step. On the other hand, the large-scale circulation patterns simulated in the different ensemble members are very similar (not shown), so the noises caused by the chaotic system can be constrained and the effective fire aerosol forcing signal can be easily identified.

Each ensemble in group B includes 10 members. The only difference between the members is the relaxation time scale of temperature, which varies from 10 to 11 days at an interval of 0.1 day. All simulations started on April 1, 2009 and were integrated for 10 days. For each simulation (e.g. E_QF), the initial condition was generated by combining the meteorological fields from initial condition outputs in the S_NF simulation with aerosol and precursor concentrations from initial condition outputs in the single-member simulation forced by the corresponding fire emission (S_QF).

2.4 Calculation of fire aerosol RF

The IPCC AR5 report provides a more useful characterization of aerosol forcing by allowing for rapid tropospheric adjustments (Boucher et al., 2013) compared to the original definition of aerosol forcing. It quantifies aerosol radiative effects in terms of Effective Radiative Forcing from aerosol-radiation interactions (ERFari) and Effective Radiative Forcing from aerosol-cloud interactions (ERFaci). ERFari refers to the combined effect of instantaneous radiative forcing from direct scattering and absorption of sunlight (aerosol direct effect) and related subsequent rapid adjustments of atmospheric state variables and cloudiness (aerosol semi-direct effect). ERFaci refers to the indirect forcing resulting from aerosol induced changes in cloud albedo (first albedo effect) and subsequent changes in cloud lifetime as rapid adjustments (second aerosol indirect effect) via microphysical interactions.

To allow for a straightforward comparison with previous studies in the literature, we followed the IPCC concept of including rapid adjustments (effective aerosol radiative forcing), but continued to decompose the aerosol effect in the conventional terms as aerosol direct radiative effect (DRE), aerosol cloud radiative effect (CRE) and surface albedo effect. Note that as nudging timescale determines the degree to which model physics are constrained (Kooperman et al., 2012), the use of a 6-hour relaxation time scale for horizontal wind nudging means only very fast adjustments are considered in the simulations.

Similar to Jiang et al. (2016), our calculations ~~of fire aerosol DRE, CRE and surface albedo effect~~ are based on the work of Ghan et al. (2012) and Ghan (2013). Fire aerosol DRE, CRE and surface albedo effect are defined as fire induced changes in aerosol forcing, cloud forcing, and surface albedo forcing respectively, and are calculated as the difference of each item ~~They were calculated as the radiative flux differences~~ between simulations with and without fire emissions (denoted by Δ). In each simulation, aerosol ~~(direct)~~ forcing was defined as the difference between all-

sky and clean-sky TOA radiative fluxes ($F - F_{clean}$). ~~Aerosol induced c~~Cloud forcing ~~change~~ was defined as the difference between all-sky and clear sky TOA radiative fluxes under clean-sky conditions ($F_{clean} - F_{clean,clear}$). The rest were related to surface albedo forcing ($F_{clean,clear}$). Thus fire aerosol DRE, CRE, and surface albedo effect were expressed as $\Delta(F - F_{clean})$, $\Delta(F_{clean} - F_{clean,clear})$, and $\Delta F_{clean,clear}$, respectively. More details about the method can be found in section 2 of Ghan (2013). CRE includes contributions of both aerosol indirect effect and aerosol semi-direct effect but was analyzed as a single term (i.e., the sum).

2.5 Observational Data

In this study, we used two sets of AOD reanalysis and the AERONET data (Holben et al. 1998) to evaluate the modeled AOD. The two AOD reanalysis datasets are the Naval Research Laboratory (NRL) reanalysis (Rubin et al. 2015) and the Monitoring Atmospheric Composition and Climate (MACC) reanalysis (Eskes et al. 2015). Both are generated by assimilating AOD retrievals from MODIS (Zhang et al., 2008; Benedetti et al., 2009) with forecast fields. The NRL reanalysis provides 6-hourly AOD at 1°horizontal resolution. The MACC dataset provides 3-hourly AOD at 1.125°horizontal resolution. Daily averages in April, 2009 were used for model evaluation in this study. AERONET retrievals of AOD from April 1 to April 30 in 2009 were used for model evaluation. Two sites are available in the selected regions: Cart_Site (36°N, 97°W) and Mexico_City (19°N, 99°W). LEV 2.0 cloud-screened all points AOD at 500 nm and 675 nm was used to generate hourly AOD at 550nm, which are the processed data based on a cloud-screening algorithm (Smirnov et al. 2000).

In addition, the simulated BC and primary organic matter (POM) concentrations were compared with observations from the Interagency Monitoring of Protected Visual Environments (IMPROVE) (Malm et al. 2004). IMPROVE aerosol data are only available over the Central U.S. A total of fifteen sites were selected and marked in Fig 2, which include the sites west of 94°W near the source region (asterisks) and sites east of 94°W in the downwind region (dots). Observed organic carbon concentrations were multiplied by 1.4 for comparison with simulated POM. Detailed descriptions about the data and sites are available at http://vista.cira.colostate.edu/improve/. The IMPROVE network collect 24-hour aerosol data on every third day. Daily averages during April, 2009 are compared on IMPROVE observation days only.

3. Results

In this part, the model performance is first evaluated based on the simulations in group A. Next, we present the simulated short-term effective fire aerosol forcing on 10-day and daily timescales based on the results from group B simulations. We will demonstrate the importance of using ensemble simulations in estimating the short-term aerosol effective forcing and give a quantitative estimate of how many ensemble members are needed for the case selected in this study.

3.1 Model Evaluation

Model simulated AOD are evaluated against the NRL and MACC reanalysis data (Fig. 3). The simulated temporal variation of regional mean AOD over the central U.S. is consistent with that in the reanalysis, but the magnitudes of simulated AOD are lower (Fig. 3). A better agreement is found between the model and the NRL data, despite the horizontal winds in the simulation are nudged towards a reanalysis that is very similar to the data used to derive MACC. Temporal correlation coefficients (TCC) between the modeled AOD and the NRL reanalysis are 0.87 and 0.82 for S_QF and S_GF4 simulations, respectively, but are lower (0.67 and 0.78) between the modeled AOD and the MACC reanalysis data. The corresponding root mean square errors rise from 0.13 (S_QF) and 0.1 (S_GF4) to 0.23 and 0.21. Generally, AOD is underestimated by a factor of 2-4 in all simulations compared to the reanalysis, especially in simulations with GFED emissions. Previous studies have found the underestimation of AOD in simulations with GFED emissions and suggested the need to scale up GFED emissions by a factor of 1-3 to match the observed AOD (Tosca et al., 2013). This is consistent with the large negative bias in the simulation S_GF3 and S_GF4. However, a much larger scale factor might be needed in this case. Simulated AOD in these two simulations are almost indistinguishable due to the small difference in the total fire emission in the region.

Over Mexico, different simulations produce similar temporal variations in AOD, but the magnitude is smaller in the GFED simulations. Fire aerosol-induced AOD increase accounts for 8.1% (S_GF3), 11.2% (S_GF4) and 48.8% (S_QF) of the background AOD (Table S2). Large discrepancies are found between model results and reanalysis data during Apr. 17-20. An increase of AOD is captured by both reanalysis datasets, while model results display a decrease of AOD compared to earlier days in the simulation period. Note that the two sets of reanalysis data also have some differences occasionally. For example, during Apr. 10-12, NRL data displays an increase of AOD, while MACC data show the opposite. These discrepancies may partly result from the large internal variability in this tropical region, where the simulated atmosphere state and its influence on aerosol transport are more likely to disagree between the model and the reanalysis. Generally speaking, the model forced with different fire emissions is capable of capturing daily variation of AOD in both regions, especially during Apr. 1-10. This period was selected for further investigation of the short-term fire aerosol effect.

Model simulated AOD are also evaluated against AERONET retrievals (Fig. 4). At Cart Site (36°N, 97°W), with the QFED emission (S_QF) the model performs well in simulating both the temporal variation (TCC=0.62) and magnitude of AOD. Simulations with GFED emissions also reproduce the temporal evolution well (TCC = 0.58 for S_GF3 and 0.55 for S_GF4), but with significantly low bias (mean bias by a factor of 2). The simulated difference in AOD magnitude is similar to that found by Zhang et al. (2014) over the northern sub-Saharan African. Using the QFEDv2.4 fire emission, the simulated regional mean AOD is a factor of 1.5 higher than that using the GFEDv3.1 emission in their study. Relatively good performance of S_QF is also seen over Mexico. The simulated time evolution agrees well with AERONET retrievals except for small discrepancies (e.g. during Apr.17 -19). A better agreement with the AERONET retrievals is found for the NRL data than MACC reanalysis at both sites. Consistent with the evaluation using reanalysis, the simulated temporal evolution of AOD during Apr. 1-10 agrees well with both reanalysis data and AERONET retrievals in selected regions. This gives us further confidence in choosing this period for further investigation.

The model is further evaluated against the IMPROVE data for BC and POM mass concentrations (Fig. 5). In the downwind region, the simulated mass concentrations in simulation S_QF lie within a factor of 2 of the observed values at most sites. However, the magnitude is generally underestimated in simulations with the GFED emissions (S_GF3 and S_GF4), especially in S_GF3. BC and POM concentrations in the downwind regions are affected by transport of aerosols from Southern Mexico (Fig. S3). A larger amount of fire emission in Southern Mexico would result in a higher BC (POM) concentration in the downwind region. This explains the slightly higher concentrations in the simulation S_GF4 than S_GF3, as BC and POM emissions over Southern Mexico are higher in GFED v4.1 due to the inclusion of small fires (Randerson et al., 2012). The good agreement between S_QF and observations suggests that the QFED data have a reasonable total emission rate. However, in the source region, the simulation S_QF displays large positive bias with a large majority of the values fall out of the a-factor-of-2 band. Given the reasonable total emission rate in QFED and a good agreement of AOD with AERONET retrievals at Cart_site, this might result from the discrepancies in the vertical distribution the fire emissions. Fire-emitted BC and POM in simulations S_QF and S_GF3 reach maximum values in the lowest level and decrease sharply to the next level, while low-level fire emissions in S_GF4 distribute in a more uniform way (Fig. S4). As the sampling was done on the lowest model level at most sites to compare with the IMPROVE data, this explains the strong overestimation in S_QF. Although the same impact from vertical distribution of fire emission also appears in simulation S_GF3, it is partly offset by its negative bias in the total emission rate.

3.2 10-day Mean Results

Given the good model performance during Apr 1-10, we proceed to analyze the short-term effects of fire aerosols during this period with nudged ensemble simulations. We define “fire AOD” as the AOD difference between the simulations with and without fire emissions.

3.2.1 Fire Aerosol Distribution

Fig. 6 shows the spatial distributions of 10-day average ensemble mean fire AOD. For reference, the total AOD in the simulation without fire emissions is shown in Fig. S2. During the period, regional mean AOD increases by 6.4% (E_GF3), 6.4% (E_GF4) and 70.2% (E_QF) in the central U.S. and 10.4% (E_GF3), 13.3% (E_GF4), and 49.6% (E_QF) in Southern Mexico when fire emissions are included. In E_QF, high fire AOD covers almost the entire selected region and extends further north. Maximum values of fire AOD stay above 0.2 around the Yucatan Peninsula. Over the Central U.S, significant fire AOD ranging between 0.04 and 0.1 appears in the southwest part of the selected region. Apart from the significant AOD difference in selected regions, large fire AOD also appears near the eastern coast as a result of local fire emission and the eastward transport of fire aerosols from both regions. Overall, the modeled fire AOD is much smaller in simulations with GFED emissions.

3.2.2 Fire Aerosol Radiative Effect

As described in Sect. 2.4, fire aerosol radiative effect can be decomposed into three items including fire aerosol DRE, fire aerosol CRE and fire aerosol surface albedo effect (Table S3). Fig.7 shows the spatial distributions of

shortwave direct effect (SDRE) and shortwave cloud radiative effect (SCRE). They are major contributors to the total fire aerosol forcing in the selected regions. For reference, total aerosol forcing and total shortwave cloud forcing in the simulation without fire emissions are shown in Fig. S2. The spatial distribution of SDRE and SCRE are similar for the three cases, but with different magnitudes and statistical significant regions for simulations with QFED and GFED fire emissions. In the Central U.S., fire aerosol SDRE is negligible in GFED forced simulations due to small fire AOD. Although the fire AOD is larger in simulation E_QF, the compensation between warming effect of fire BC and cooling effect of fire POM still results a weak forcing of about -0.1W m^{-2}. Over southern Mexico, all simulations produce significant cooling by fire aerosol SCRE with maximum values three times as large as those of corresponding SDRE. For both SDRE and SCRE, the largest fire aerosol effects appear in the E_QF simulation while the E_GF3 yields the weakest forcing, which is consistent with the modeled fire AOD in these simulations.

In the following analysis, we will focus on the results from the E_QF simulation. Both SDRE and SCRE spread outside the two selected regions and extend eastward reaching coast regions. A stronger fire aerosol effect is seen in the Southern Mexico region. Strong SDRE appears over the Yucatan Peninsula where fire AOD peaks (Fig. 6). Regional mean 10-day average of SDRE and SCRE reach ~~-1.66~~-0.86 W m^{-2} and -3.02W m^{-2} respectively. It's interesting to note that the maximum SCRE tends to center around adjacent Gulf of Mexico rather than the land region. In the central U.S, despite moderate fire aerosol SDRE, ~~SCRE near fire source region is weaker than -4 W m^{-2}, which is comparable to that in the extended regions~~ a positive SCRE exceeding 2 W m^{-2} appears in the north part of the region while a comparable negative SCRE appears in the south part of the region~~.~~

To find out the causes of the fire aerosol SCRE, fire aerosol-induced changes in cloud properties are analyzed. Given the largely insignificant change in cloud fraction (Fig. 8), the negative fire aerosol SCRE in both regions ~~are~~ is mainly ~~induced by changes~~ associated with increases in liquid water path (LWP) and droplet number concentrations (CDNC). ~~Changes in ice water path (IWP) and ice crystal number concentration (ICNC) can also significantly affect SCRE, but with an opposite sign and mostly in the central U.S. Fire aerosol SCRE in the central U.S is associated with significant increases in both column integrated droplet number concentration (smaller droplet effective radius) and LWP, indicating important contributions of both the aerosol first and second indirect effects. Increased CDNC enhances cloud albedo by decreasing droplet sizes (Twomey, 1977) and allows more liquid water to accumulate by decreasing precipitation efficiency (Albrecht, 1989; Ghan et al., 2012). Note that although LWP and CDNC over southern Mexico change in a smaller magnitude than those in central U.S., fire aerosol SCRE is stronger over Southern Mexico. This is mainly due to the reductions in IWP and ICNC over the Central U.S. These changes, which possibly caused by fire aerosol induced changes in the circulation (Ten Hoeve et al, 2012), lead to a positive SCRE that partly offsets the negative SCRE caused by changes in warm clouds.~~ The increased CDNC due to an increase of CCN from fire aerosols (Fig. 8) leads to smaller droplet sizes, which in turn increase cloud albedo by enhancing backscattering (Twomey, 1977) and further affect LWP by decreasing precipitation efficiency and allowing more liquid water to accumulate (Albrecht, 1989; Ghan et al., 2012). These changes in warm cloud properties demonstrate important contributions of both aerosol first and second indirect effects to the negative SCRE. Over Southern Mexico, although changes of CDNC and LWP are of comparable magnitudes between Gulf of Mexico and the land region (Fig.8), relative changes of both items are much larger over Gulf of Mexico (Fig.S5)

due to the smaller magnitudes of background CDNC and LWP here (Fig. S6), which tend to lead to a more sensitive response of SCRE. That's why the maximum SCRE over Southern Mexico is more centered around Gulf of Mexico. Changes in ice water path (IWP) and ice crystal number concentration (ICNC) can also significantly affect SCRE, but with an opposite sign and mostly in the central U.S. The decreased IWP and ICNC, which are possibly caused by fire aerosol-induced changes in the circulation (Ten Hoeve et al, 2012) and reduced coarse mode dust aerosol concentrations (Fig.S7), are responsible for the positive SCRE and the negative longwave cloud radiative effect (Table S3) in the north part of central U.S. In the south part of central U.S., the reduction of IWP and ICNC also results in a positive SCRE, which partly offsets the negative SCRE resulting from changes in warm cloud properties. This explains the weaker total negative SCRE in this region compared to the Southern Mexico region despite the more substantial increase in CDNC and LWP here. In the northeast of the extended coastal regions, a more significant change of LWP comparable to that in the central U.S appears, while a more significant change of CDNC comparable to that in Southern Mexico occurs in the southwest. The combined effect leads to the total fire aerosol effect in the extended regions.

The ensemble method provides another effective way to distinguish fire aerosol radiative effect by comparing the radiative forcing distribution of ensemble members between simulation with and without fire emission. A significant difference in the distribution of total aerosol (cloud) forcing indicates a significant fire aerosol direct (cloud) effect. As shown in Fig. 9, a shift towards stronger magnitude occurs to the total aerosol forcing when fire aerosols are considered. Simulation E_QF has a larger percentage of grid cells with SDRE below -4.2W m^{-2}, while more grid cells exceed -4.2W m^{-2} in E_NF, which indicates a significant negative fire aerosol direct effect. The ~~S~~same shift also appears to the total shortwave cloud forcing with more grid cells having shortwave cloud forcing below -30W m^{-2} in the simulation E_QF. Regional mean total aerosol and shortwave cloud forcing in southern Mexico become more negative (-0.86 and -3.02 W m^{-2}) with fire aerosols.

Fig. 10 illustrates ensemble behavior of 10-day average regional mean total aerosol and cloud forcing from all simulations as well as resulted fire aerosol SDRE and SCRE. The GFED forced simulations not only resemble in ensemble mean, but also have small difference in ensemble member distribution. Although members in the E_QF simulation capture stronger aerosol forcing, thus stronger fire aerosol SDRE than those in E_GF3 and E_GF4, the ensemble spread (as indicated by the maximum and minimum values) in the three simulations is similar. Moreover, the E_QF simulation yields a smaller spread of SCRF compared with the GFED forced simulations despite a stronger ensemble mean SCRF. In each fire simulation, ensemble mean fire aerosol SCRE has a much larger magnitude than SDRE. So is the corresponding ensemble spread. Taking results from E_QF simulation as an example, ensemble spread of SCRE reaches 0.47 W m^{-2}, accounting for 15.6% of the corresponding ensemble mean, while ensemble spread of SDRE is 0.03W m^{-2} accounting for 3.5% of the corresponding ensemble mean.

3.3 Daily RF

The fire aerosol effect is also investigated for individual days. The spatial distributions of SDRF and SCRF on April 7 are shown in Fig 11, when relatively high fire emissions appear in both regions. Negative fire aerosol SDRE appears in the central U.S. biomass-burning region indicating the dominant role of POM scattering. Fire aerosol SDRE over

Southern Mexico shows a contrast of warming effect in land region and cooling effect in adjacent ocean despite similar aerosol loading in the two regions. However, they do have nearly equal clear-sky BC absorption and POM scattering (Fig. 12). Difference in low-level cloud distributions between two regions leads to different signs of the simulated all-sky SDRE. Over land, when clouds appear under elevated aerosol layers, more solar radiation is reflected back to space and this leads to amplified BC absorption and more positive direct aerosol forcing (Keil and Haywood, 2003; Zhang et al., 2016; Jiang et al., 2016). In contrast, neither absorption nor scattering changes significantly from clear-sky to all-sky condition over adjacent areas over the ocean, since the small cloud fraction is small. Same enhanced absorption of above-cloud aerosols is also found over the west Atlantic Ocean. Fire aerosols produce remarkable negative SCRE up to -16W m^{-2} over Southern Mexico land in response to the increase in CDNC and LWP.

3.4 Discussion about Simulation Strategy

Fig. 13 shows the daily variation of the regional mean total (direct) aerosol forcing and cloud forcing. Both the ensemble mean and spread are investigated here. The total aerosol forcing exhibits considerable diversity across ensemble members within each simulation even though the simulated AOD is nearly indistinguishable (Fig. 3). Taking results from simulation E_QF as an example, maximum values of difference between members exceed 0.4 W m^{-2} for aerosol forcing and 5W m^{-2} for cloud forcing, which are approximate 10% of the corresponding ensemble mean values. The large spread of total aerosol forcing and cloud forcing will lead to uncertainties in the estimation of fire aerosol effect. This points out the importance of conducting ensemble simulations in order to get a more comprehensive estimate of daily fire aerosol effect. The minimum ensemble size required for this case is investigated in terms of the ensemble mean and spread estimate. Simulated ensemble mean fire aerosol SDRE remains nearly unchanged regardless of the ensemble size (Fig. 14a). However, discrepancies in the ensemble mean fire aerosol SCRF (Fig. 14b) are substantial when the number of ensemble members is smaller than 8. The same is true for the ensemble spread of fire aerosol SCRF (Fig. S8~~5~~). ~~Overall, the time evolution and magnitude of ensemble mean and spread tend to converge when the number of ensemble members reaches about 9 for different days we investigated here.~~ In order to quantify the discrepancies of the simulated SCRE, we chose the ensemble mean SCRE in the 20-member simulation as a reference and use the root mean square errors (RMSE) of the ensemble mean SCRE in the N-member simulation to quantify the deviation of the simulated SCRE from the reference value. It is calculated as the standard deviation of the differences between the daily ensemble mean SCRE in the N-member simulation and the 20-member simulation. For each N, we randomly sampled 1000 times from the 20 members to help reduce the influence from limited sampling. Figure 15 shows that both the RMSE of ensemble mean SCRE and the difference of RMSE between the 1000 groups of simulations (for each N) decrease with increasing N. The minimum number of N required is determined when the 90th percentile of RMSE is smaller than a threshold RMSE. Without a good reference, we set the threshold RMSE to 20% (0.566Wm^{-2}) of the reference 10-day mean SCRE (-2.83Wm^{-2}). As shown in Fig.15, at least 11 members are needed to meet this criterion.

Fire aerosol sources are often intermittent and height-dependent and there is a need to estimate the short-term effective aerosol forcing. Although nudging helps to constrain large-scale features, the simulated cloud properties (e.g. cloud fraction and LWP) and their response to aerosol changes can still be sensitive to small perturbations in the

atmospheric state. Therefore, for investigating the short-term aerosol effect, a single simulation might not be sufficient to tell whether the aerosol effect is significant. The use of ensembles provides an effective way to estimate the uncertainty. Previous investigations of short-term fire aerosol effect are mainly based on single-member simulations (Wu et al., 2011; Sena et al., 2013; Kolusu et al., 2015). While this might be less a problem for SDRE, one should be more careful when investigating the aerosol indirect effect and conduct ensemble simulations to see whether the estimated fire aerosol effects are robust.

4. Summary

In this study, we investigated the short-term effect of fire aerosols on cloud and radiation using CAM5 simulations. Month-long single-member simulations and 10-day ensemble simulations were conducted in April, 2009. In order to help extract signals on short time scales, we used nudging to constrain horizontal winds in all simulations. Our investigation focused on Southern Mexico where there were constant intensive fire activities and the Central U.S. with occasionally large fires. Apart from the local effect, fire emissions from the two regions are shown to affect downwind coastal regions through transport.

Modeled AOD and mass concentrations (BC and POM) were evaluated against observations. In general, all simulations with fire emissions reproduce the observed temporal variation of daily mean AOD well, although the simulated magnitude is smaller. The model performance is better when QFEDv2.4 is used, which has larger fire emissions. Modeled regional mean AOD values in simulations using two versions of GFED fire emission data are barely distinguishable, despite the inclusion of small fires and changed injection heights in GFEDv4.1 used in this study. Both ~~of them~~ simulate about a factor of 1.5 smaller AOD than that in the simulation using the QFED fire emissions. At sites in the downwind region, the modeled BC and POM mass concentrations in the simulation with QFEDv2.4 emission (S_QF) agree well with the IMPROVE data. In contrast, simulations with the other two fire emission datasets (S_GF3 and S_GF4) have a low bias. The simulated AOD in the source region in S_QF also agrees well with the AERONET data (Cart_Site). If there is no large compensating error in the model, QFEDv2.4 seems more reasonable in terms of the total (vertically-integrated) emission rate. On the other hand, S_QF strongly overestimates BC and POM concentrations in the source region. Considering that the source-region AOD and the downwind surface mass concentrations are well simulated, the overestimation suggests the actual emission peak might appear at higher levels compared to the height-dependent injection rates applied in the S_QF simulation.

Based on the evaluation, we chose the first 10 days as the simulation period and focused on the simulation with QFEDv2.4 fire emission in our ensemble nudged simulations. In our method, the nudged ensembles are generated by adding a very weak temperature nudging along with horizontal-wind nudging and perturbing the nudging time scale of temperature gently. In this way, small temperature perturbations are added to the simulation at each time step, while the large-scale circulation features are very similar between individual members. We first investigated the 10-day mean effective fire aerosol forcing. Decomposition of total aerosol radiative forcing shows that fire aerosol effects in the two selected regions are dominated by the ~~shortwave cloud radiative effect~~ SCRE. All fire simulations show similar spatial distribution of SDRE and SCRE, but with different magnitudes and statistically significant regions. The similarity in the spatial distribution is expected since the three emission datasets differ mainly in the emission

magnitude and no much in spatial distribution in the focus regions of this study. Fire aerosol effects in simulations with GFED emissions (E_GF3 and E_GF4) are weaker than that with QFEDv2.4 emission (E_QF) by a factor of 1.5 for SCRE and a factor of more than 4 for SDRE. ~~Generally speaking,~~ Overall, the difference in simulated AOD and fire aerosol indirect radiative effects between simulations is smaller compared to the difference between fire emissions, consistent with the findings in sub-Saharan African biomass-burning region (Zhang et al. 2014).

Fire aerosols produce a negative direct effect of -0.1 $W\,m^{-2}$ in the Central U.S. and -0.86 $W\,m^{-2}$ in Southern Mexico in E_QF during the 10-day period. Within each region, negative fire aerosol SDRE peaks where fire AOD reaches maximum. Unlike the limited area affected by significant fire aerosol SDRE, fire aerosol SCRE from selected regions spreads eastward and northward, affecting remote coast regions. Ensemble mean results show strong SCRE over almost the entire Southern Mexico, with a 10-day regional mean value of -3.02 $W\,m^{-2}$. Over the central U.S, the SCRE is positive in the north and negative in the south and the regional mean is small (-0.56 $W\,m^{-2}$). Maximum SCRE stays below -4 $W\,m^{-2}$ in the (south) central U.S. and -10 $W\,m^{-2}$ in Southern Mexico in response to significantly increased LWP and CDNC. Decreases of IWP and ICNC also contribute to fire aerosol SCRE in the Central U.S. but with an opposite sign. The offset effect of the positive forcing induced by changes in cloud ice properties explains the smaller SCRE in the central U.S. despite the larger changes in cloud droplet properties.

We also investigated fire aerosol effects on the daily time scale, where the variation in the simulated fire aerosol effect can be large among the ensemble members. The large ensemble spread of total aerosol and cloud forcing indicates large uncertainties in estimating daily fire aerosol effects, despite similar AOD across ensemble members. Further investigations show that the simulated ensemble mean and spread with less than 7 members differs considerably to those with more members. ~~A minimum of 9 members is necessary to achieve a steady estimate of the magnitude and temporal variation of SCRE in this case.~~ Our results suggest that for short-term simulations of aerosol and cloud processes, even small perturbations might result in large difference across members despite constrained large scale features. In order to obtain a robust estimate of the effective fire aerosol forcing during a short period, it is important to conduct ensemble simulations with sufficient ensemble members.

Acknowledgments

We thank two anonymous reviewers for their careful reviews and suggestions that helped to greatly improve the analyses and discussion presented in this paper. This study was supported primarily by the U.S. Department of Energy (DOE)'s office of Science as part of the Regional and Global Climate Modeling Program (NSF-DOE-USDA EaSM2). The work was also supported by the National Natural Science Foundation of China (NSFC) under Grants No. 41621005 and 41330420, the National Key Basic Research Program (973 Program) of China under Grant No. 2010CB428504, and the Jiangsu Collaborative Innovation Center of Climate. The Pacific Northwest National Laboratory (PNNL) is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830.

Computations were performed using resources of the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory and PNNL Institutional computing. All model results are available from the corresponding author upon request.

References:

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227-1231, 1989.

Benedetti, A., Morcrette, J. J., Boucher, O., Dethof, A., Engelen, R., Fisher, M., Flentje, H., Huneeus, N., Jones, L., and Kaiser, J.: Aerosol analysis and forecast in the European centre for medium-range weather forecasts integrated forecast system: 2. Data assimilation, Journal of Geophysical Research: Atmospheres, 114, 2009.

Bond, T. C., Doherty, S. J., Fahey, D., Forster, P., Berntsen, T., DeAngelo, B., Flanner, M., Ghan, S., Kärcher, B., and Koch, D.: Bounding the role of black carbon in the climate system: A scientific assessment, Journal of Geophysical Research: Atmospheres, 118, 5380-5552, 2013.

Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J.-L., Hall, A., Hallegatte, S., Holland, M. M., and Ingram, W.: How well do we understand and evaluate climate change feedback processes?, Journal of Climate, 19, 3445-3482, 2006.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., and Lohmann, U.: Clouds and aerosols, 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, Cambridge University Press, 571-657, 2013.

Brito, J., Rizzo, L. V., Morgan, W. T., Coe, H., Johnson, B., Haywood, J., Longo, K., Freitas, S., Andreae, M. O., and Artaxo, P.: Ground-based aerosol characterization during the South American Biomass Burning Analysis (SAMBBA) field experiment, Atmospheric Chemistry and Physics, 14, 12069-12083, 2014.

Chen, D., Liu, Z., Schwartz, C. S., Lin, H.-C., Cetola, J. D., Gu, Y., and Xue, L.: The impact of aerosol optical depth assimilation on aerosol forecasts and radiative effects during a wild fire event over the United States, Geoscientific Model Development, 7, 2709-2715, 2014.

Chubarova, N., Nezval, Y., Sviridenkov, I., Smirnov, A., and Slutsker, I.: Smoke aerosol and its radiative effects during extreme fire event over Central Russia in summer 2010, Atmospheric Measurement Techniques, 5, 557-568, 2012.

Chubarova, N. Y., Prilepsky, N. G., Rublev, A. N., and Riebau, A. R.: A Mega-Fire event in central Russia: fire weather, radiative, and optical properties of the atmosphere, and consequences for subboreal forest plants, Developments in environmental science, 8, 247-264, 2008.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., and Bauer, P.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the royal meteorological society, 137, 553-597, 2011.

Eskes, H., Huijnen, V., Arola, A., Benedictow, A., Blechschmidt, A.-M., Botek, E., Boucher, O., Bouarar, I., Chabrillat, S., and Cuevas, E.: Validation of reactive gases and aerosols in the MACC global analysis and forecast system, Geoscientific model development, 8, 3523-3543, 2015.

513 Ghan, S. J., Liu, X., Easter, R. C., Zaveri, R., Rasch, P. J., Yoon, J.-H., and Eaton, B.: Toward a minimal representation
514 of aerosols in climate models: Comparative decomposition of aerosol direct, semidirect, and indirect radiative forcing,
515 Journal of Climate, 25, 6461-6476, 2012.

516 Ghan, S. J.: Technical Note: Estimating aerosol effects on cloud radiative forcing, Atmos. Chem. Phys., 13, 9971-
517 9974, doi:10.5194/acp-13-9971-2013, 2013.

518 Giglio, L., Randerson, J. T., and van der Werf, G. R. (2013), Analysis of daily, monthly, and annual burned area using
519 the fourth-generation global fire emissions database (GFED4) J. Geophys. Res. Biogeosci., 118, 317–328,
520 doi:10.1002/jgrg.20042.

521 Holben, B. N., Eck, T., Slutsker, I., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., and
522 Nakajima, T.: AERONET—A federated instrument network and data archive for aerosol characterization, Remote
523 sensing of environment, 66, 1-16, 1998.

524 Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing
525 by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical
526 Research: Atmospheres, 113, 2008.

527 Jacobson, M. Z.: Effects of biomass burning on climate, accounting for heat and moisture fluxes, black and brown
528 carbon, and cloud absorption effects, Journal of Geophysical Research: Atmospheres, 119, 8980-9002, 2014.

529 Jiang, Y., Lu, Z., Liu, X., Qian, Y., Zhang, K., Wang, Y., and Yang, X.-Q.: Impacts of global open-fire aerosols on
530 direct radiative, cloud and surface-albedo effects simulated with CAM5, Atmospheric Chemistry and Physics
531 (Online), 16, 2016.

532 Kaufman, Y. J., Koren, I., Remer, L. A., Rosenfeld, D., and Rudich, Y.: The effect of smoke, dust, and pollution
533 aerosol on shallow cloud development over the Atlantic Ocean, Proceedings of the National Academy of Sciences of
534 the United States of America, 102, 11207-11212, 2005.

535 Keil, A., and Haywood, J. M.: Solar radiative forcing by biomass burning aerosol particles during SAFARI 2000: A
536 case study based on measured aerosol and cloud properties, Journal of Geophysical Research: Atmospheres, 108,
537 2003.

538 Kolusu, S., Marsham, J., Mulcahy, J., Johnson, B., Dunning, C., Bush, M., and Spracklen, D.: Impacts of Amazonia
539 biomass burning aerosols assessed from short-range weather forecasts, Atmospheric Chemistry and Physics, 15,
540 12251-12266, 2015.

541 Kooperman, G. J., Pritchard, M. S., Ghan, S. J., Wang, M., Somerville, R. C., and Russell, L. M.: Constraining the
542 influence of natural variability to improve estimates of global aerosol indirect effects in a nudged version of the
543 Community Atmosphere Model 5, Journal of Geophysical Research: Atmospheres, 117, 2012.

544 Korontzi, S., McCarty, J., Loboda, T., Kumar, S., and Justice, C.: Global distribution of agricultural fires in croplands
545 from 3 years of Moderate Resolution Imaging Spectroradiometer (MODIS) data, Global Biogeochemical Cycles, 20,
546 2006.

Lin, N.-H., Tsay, S.-C., Maring, H. B., Yen, M.-C., Sheu, G.-R., Wang, S.-H., Chi, K. H., Chuang, M.-T., Ou-Yang, C.-F., and Fu, J. S.: An overview of regional experiments on biomass burning aerosols and related pollutants in Southeast Asia: From BASE-ASIA and the Dongsha Experiment to 7-SEAS, Atmospheric Environment, 78, 1-19, 2013.

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., and Vitt, F.: Toward a minimal representation of aerosols in climate models: Description and evaluation in the Community Atmosphere Model CAM5, Geoscientific Model Development, 5, 709, 2012.

Liu, X.: Impacts of global open-fire aerosols on direct radiative, cloud and surface-albedo effects simulated with CAM5, Atmos. Chem. Phys, 1680, 7324, 2016.

Lu, Z., and Sokolik, I. N.: The effect of smoke emission amount on changes in cloud properties and precipitation: A case study of Canadian boreal wildfires of 2007, Journal of Geophysical Research: Atmospheres, 118, 2013.

Magi, B., Rabin, S., Shevliakova, E., and Pacala, S.: Separating agricultural and non-agricultural fire seasonality at regional scales, Biogeosciences, 9, 3003, 2012.

Malm, W. C., Schichtel, B. A., Pitchford, M. L., Ashbaugh, L. L., and Eldred, R. A.: Spatial and monthly trends in speciated fine particle concentration in the United States, Journal of Geophysical Research: Atmospheres, 109, 2004.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, Journal of Geophysical Research: Atmospheres, 102, 16663-16682, 1997.

Morrison, H., and Gettelman, A.: A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, version 3 (CAM3). Part I: Description and numerical tests, Journal of Climate, 21, 3642-3659, 2008.

Mu, M., Randerson, J., van der Werf, G., Giglio, L., Kasibhatla, P., Morton, D., Collatz, G., DeFries, R., Hyer, E., and Prins, E.: Daily and hourly variability in global fire emissions and consequences for atmospheric model predictions of carbon monoxide, 2011.

Neale, R. B., Richter, J. H., and Jochum, M.: The impact of convection on ENSO: From a delayed oscillator to a series of events, Journal of Climate, 21, 5904-5924, 2008.

Park, S., and Bretherton, C. S.: The University of Washington shallow convection and moist turbulence schemes and their impact on climate simulations with the Community Atmosphere Model, Journal of Climate, 22, 3449-3469, 2009.

Randerson, J., Chen, Y., Werf, G., Rogers, B., and Morton, D.: Global burned area and biomass burning emissions from small fires, Journal of Geophysical Research: Biogeosciences, 117, 2012.

Reddington, C., Yoshioka, M., Balasubramanian, R., Ridley, D., Toh, Y., Arnold, S., and Spracklen, D.: Contribution of vegetation and peat fires to particulate air pollution in Southeast Asia, Environmental Research Letters, 9, 094006, 2014.

Richter, J. H., and Rasch, P. J.: Effects of convective momentum transport on the atmospheric circulation in the Community Atmosphere Model, version 3, Journal of Climate, 21, 1487-1499, 2008.

582 Rubin, J. I., Reid, J. S., Hansen, J. A., Anderson, J. L., Hoar, T. J., Reynolds, C. A., Sessions, W. R., and Westphal,
583 D. L.: Development of the Ensemble Navy Aerosol Analysis Prediction System (ENAAPS) and its application of the
584 Data Assimilation Research Testbed (DART) in support of aerosol forecasting, Atmospheric Chemistry and Physics,
585 16, 3927, 2016.

586 Sena, E., Artaxo, P., and Correia, A.: Spatial variability of the direct radiative forcing of biomass burning aerosols and
587 the effects of land use change in Amazonia, Atmospheric Chemistry and Physics, 13, 1261-1275, 2013.

588 Smirnov, A., Holben, B., Eck, T., Dubovik, O., and Slutsker, I.: Cloud-screening and quality control algorithms for
589 the AERONET database, Remote Sensing of Environment, 73, 337-349, 2000.
590

591 Stier, P., Schutgens, N., Bellouin, N., Bian, H., Boucher, O., Chin, M., Ghan, S., Huneeus, N., Kinne, S., and Lin, G.:
592 Host model uncertainties in aerosol radiative forcing estimates: results from the AeroCom Prescribed intercomparison
593 study, Atmospheric Chemistry and Physics, 13, 3245-3270, 2013.

594 Tarasova, T., Gorchakova, I., Sviridenkov, M., Anikin, P., and Romashova, E.: Estimation of the radiative forcing of
595 smoke aerosol from radiation measurements at the Zvenigorod scientific station in the summer of 2002, Izvestiya
596 Atmospheric and Oceanic Physics, 40, 454-463, 2004.

597 (Dee et al., 2011), J. E., Jacobson, M. Z., and Remer, L. A.: Comparing results from a physical model with satellite
598 and in situ observations to determine whether biomass burning aerosols over the Amazon brighten or burn off clouds,
599 Journal of Geophysical Research: Atmospheres, 117, 2012.

600 Tosca, M., Randerson, J., and Zender, C.: Global impact of smoke aerosols from landscape fires on climate and the
601 Hadley circulation, Atmospheric Chemistry and Physics, 13, 5227-5241, 2013.

602 Twomey, S.: The influence of pollution on the shortwave albedo of clouds, Journal of the atmospheric sciences, 34,
603 1149-1152, 1977.

604 van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R.
605 S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest,
606 agricultural, and peat fires (1997–2009), Atmos. Chem. Phys., 10, 11707-11735, doi:10.5194/acp-10-11707-2010,
607 2010.

608 Ward, D., Kloster, S., Mahowald, N., Rogers, B., Randerson, J., and Hess, P.: The changing radiative forcing of fires:
609 global model estimates for past, present and future, Atmospheric Chemistry and Physics, 12, 2012.

610 Wan, H., Rasch, P. J., Zhang, K., Qian, Y., Yan, H., and Zhao, C.: Short ensembles: an efficient method for discerning
611 climate-relevant sensitivities in atmospheric general circulation models, Geoscientific Model Development, 7, 1961-
612 1977, 2014.

613 Wu, L., Su, H., and Jiang, J. H.: Regional simulations of deep convection and biomass burning over South America:
614 2. Biomass burning aerosol effects on clouds and precipitation, Journal of Geophysical Research: Atmospheres, 116,
615 2011.

Zamora, L. M., Kahn, R., Cubison, M. J., Diskin, G., Jimenez, J., Kondo, Y., McFarquhar, G., Nenes, A., Thornhill, K., and Wisthaler, A.: Aircraft-measured indirect cloud effects from biomass burning smoke in the Arctic and subarctic, Atmospheric Chemistry and Physics, 16, 715-738, 2016.

Zhang, F., Wang, J., Ichoku, C., Hyer, E. J., Yang, Z., Ge, C., Su, S., Zhang, X., Kondragunta, S., and Kaiser, J. W.: Sensitivity of mesoscale modeling of smoke direct radiative effect to the emission inventory: a case study in northern sub-Saharan African region, Environmental Research Letters, 9, 075002, 2014.

Zhang, G. J., and McFarlane, N. A.: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model, Atmosphere-ocean, 33, 407-446, 1995.

Table 1. List of CAM5 simulations.

Name	Fire emission	Simulation period	Member	Nudging
Group A: Single member simulations				
S_NF	No	January 1- April 30, 2009	1	Horizontal winds (6h)
S_GF3	GFED v3			
S_GF4	GFED v4.1			
S_QF	QFED v2.4			
Group B: Ensemble simulations				
E_NF	No	April 1 - April 10, 2009	10	Horizontal winds (6h) and temperature (~10d)*
E_GF3	GFED v3			
E_GF4	GFED v4.1			
E_QF	QFED v2.4			

* See section 2.3 for details about ensembles

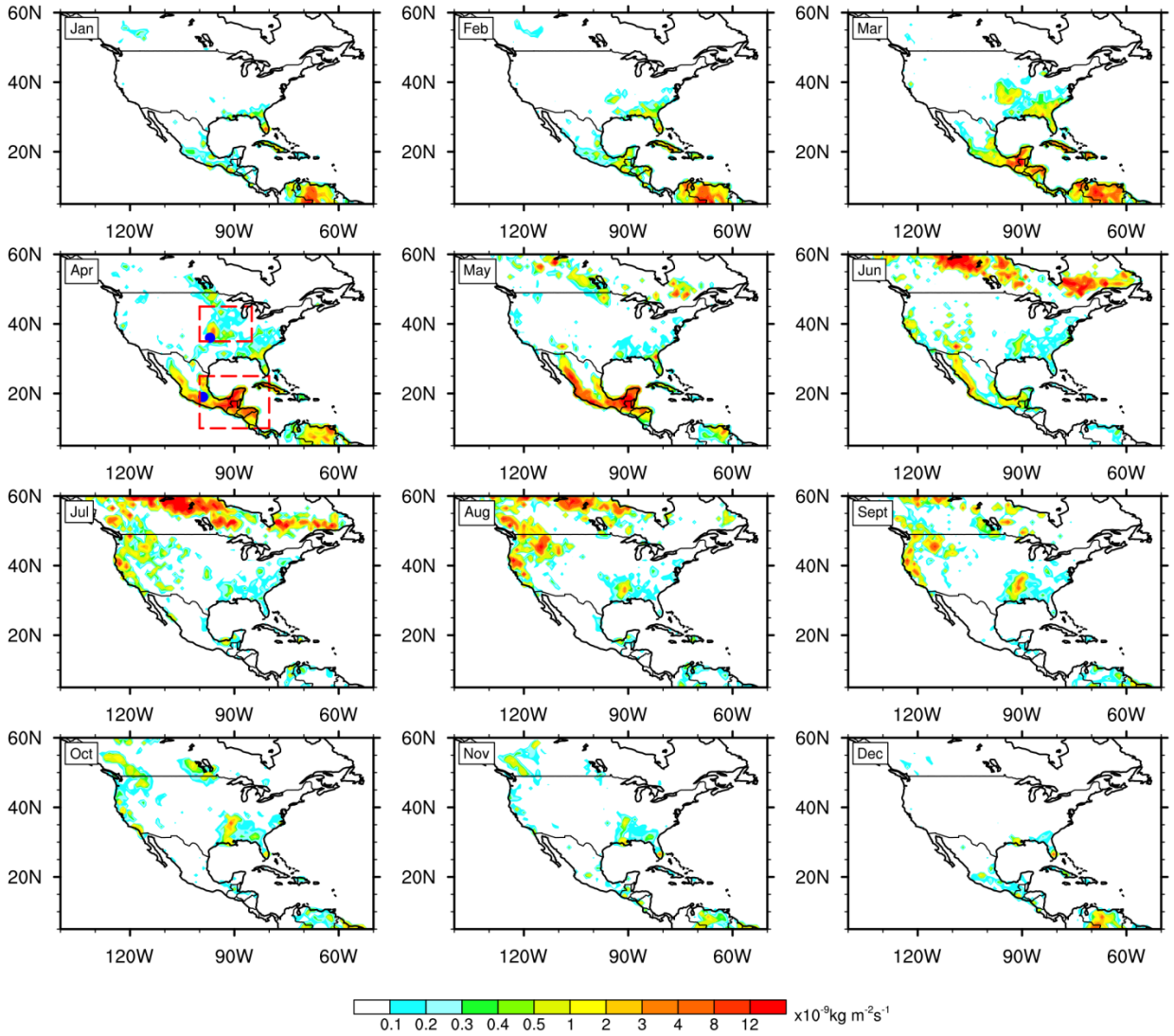

Figure 1. Spatial distributions of multi-year monthly mean biomass burning consumed dry matter over North America during 2003-2014 from GFEDv4.1. Boxes denote selected regions: central U.S (35 - 45°N, 85 - 100°W) and Southern Mexico (10 - 25°N, 80 - 100°W). Dots denote locations of AERONET sites: Cart_Site (36°N, 97°W) and Mexico_City (19°N, 99°W)

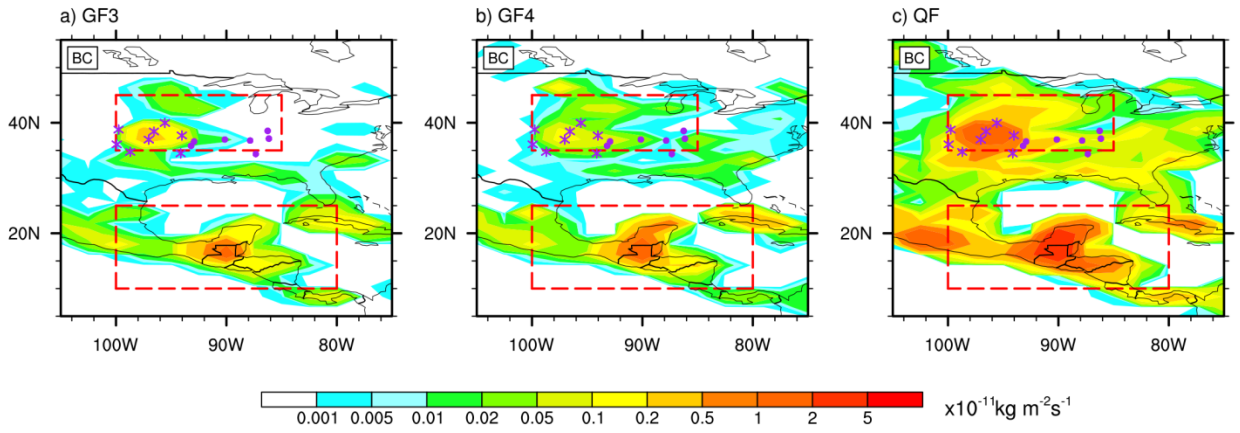

Figure 2. Spatial distributions of monthly mean BC emissions from three emission inventories in April, 2009. IMPROVE data sites are shown as asterisks for sites near the source region and as dots for sites in the region downwind of the fire source.

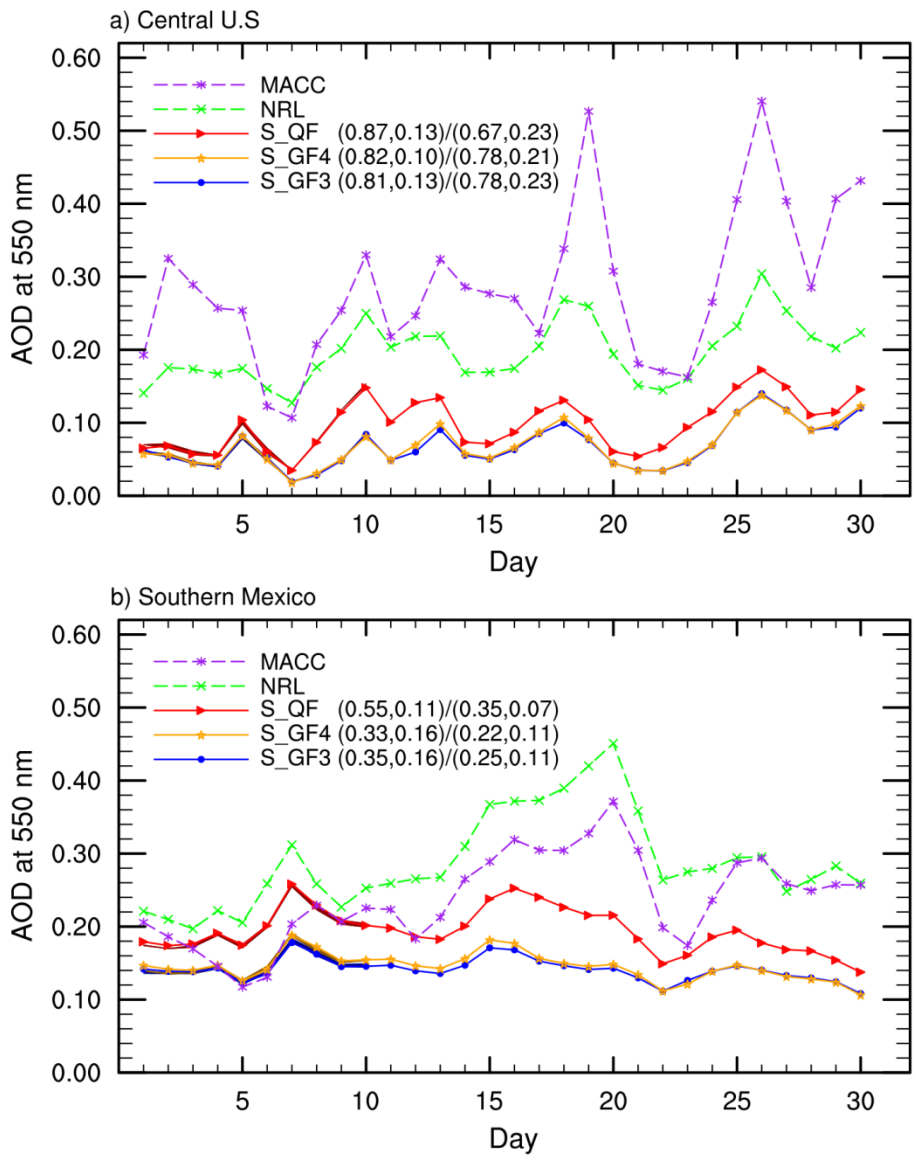

Figure 3. Time series of daily regional mean AOD in April, 2009 in simulations and reanalysis data. Numbers in parenthesis denote time correlation coefficient (TCC) and root mean square error (RMSE) between each simulation in group A and reanalysis data (left: NRL; right: MACC). Individual lines indicate group A simulations. Shaded areas (very narrow) in slightly darker colors during April 1-10 illustrate maximum and minimum values of daily mean AOD among ensemble members in group B simulations. For the single-member simulation and the ensemble simulation driven by same fire emission, the shaded area and the solid line almost overlap, given the barely indistinguishable AOD between ensemble members and the corresponding Group A simulation.

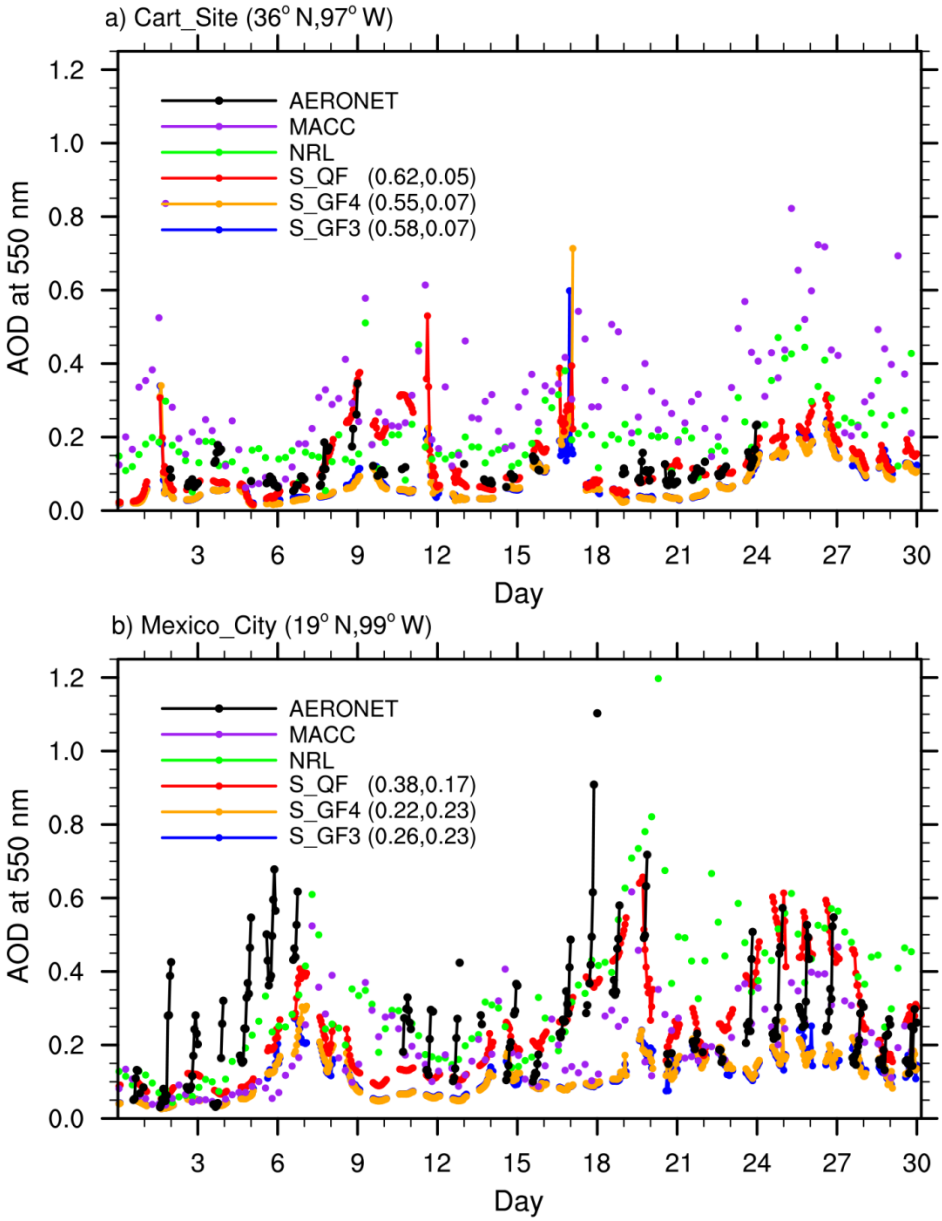

Figure 4. Time series of hourly regional mean AOD in April, 2009 from group A simulations, reanalysis data and AERONET retrievals at AERONET sites. Numbers in parenthesis denote TCC (left) and RMSE (right) between each simulation and AERONET AOD.

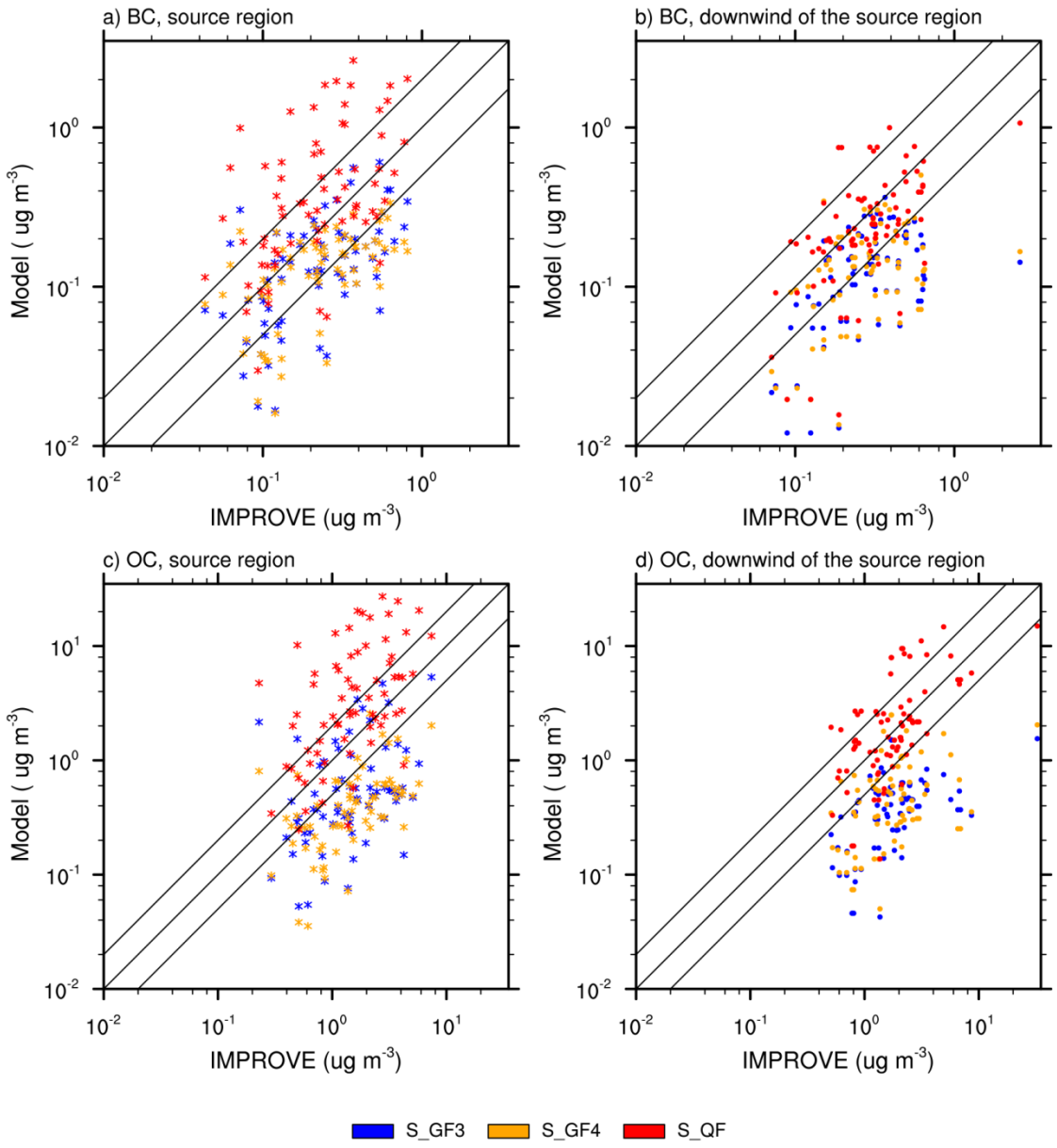

Figure 5. Evaluation of simulated BC (up) and POM (bottom) concentrations in group A simulations against the IMPROVE data at sites near the source and downwind the source region. Locations of these sites are marked with the same symbol in Fig. 2.

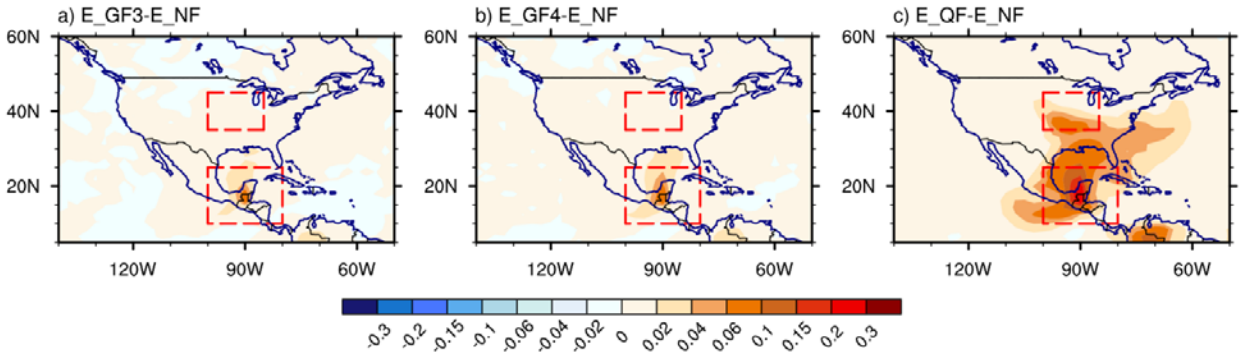

Figure 6. Spatial distributions of 10-day average (Apr. 1-10) ensemble mean AOD differences between simulations with (E_GF3, E_GF4, and E_QF) and without fire emission (E_NF).

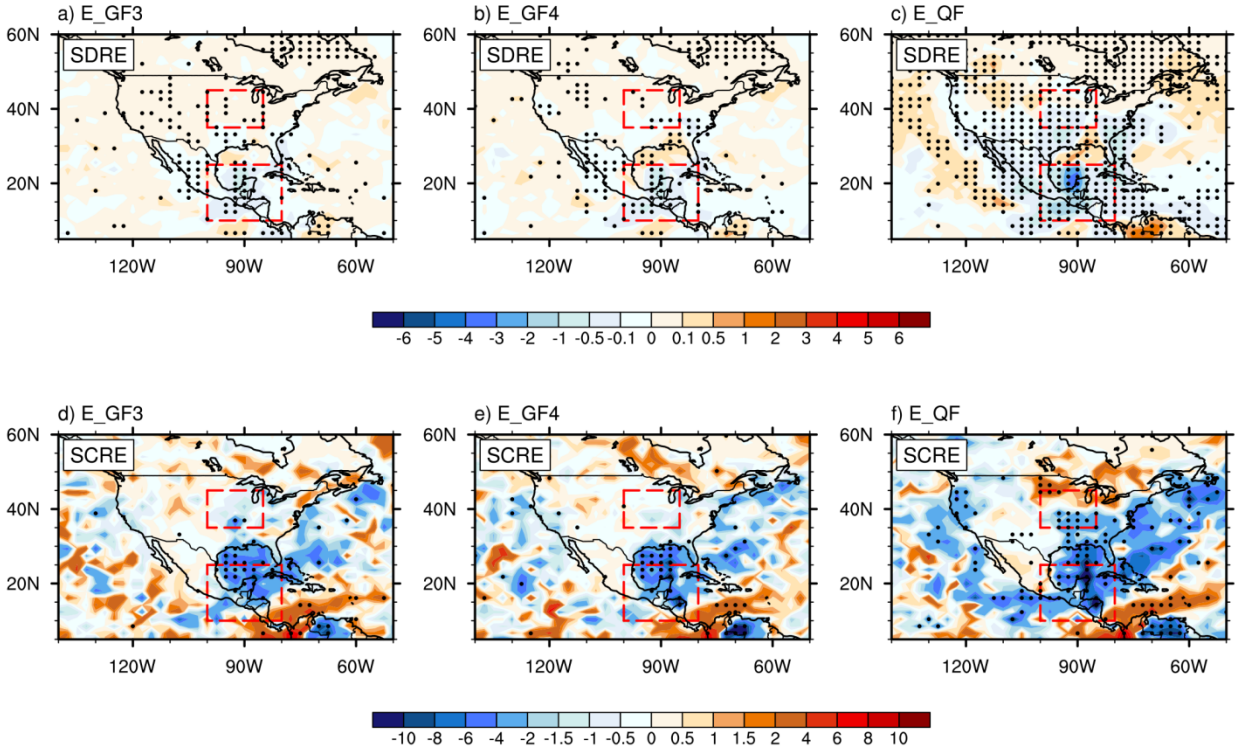

Figure 7. Spatial distributions of 10-day average (Apr. 1-10) ensemble mean fire aerosol shortwave direct radiative effect (SDRE) and shortwave cloud radiative effect (SCRE) ($W\ m^{-2}$) in group B simulations. Dots denote regions where SDRE is statistically significant at the 95% confidence level based on the Kolmogorov-Smirnov (KS) test.

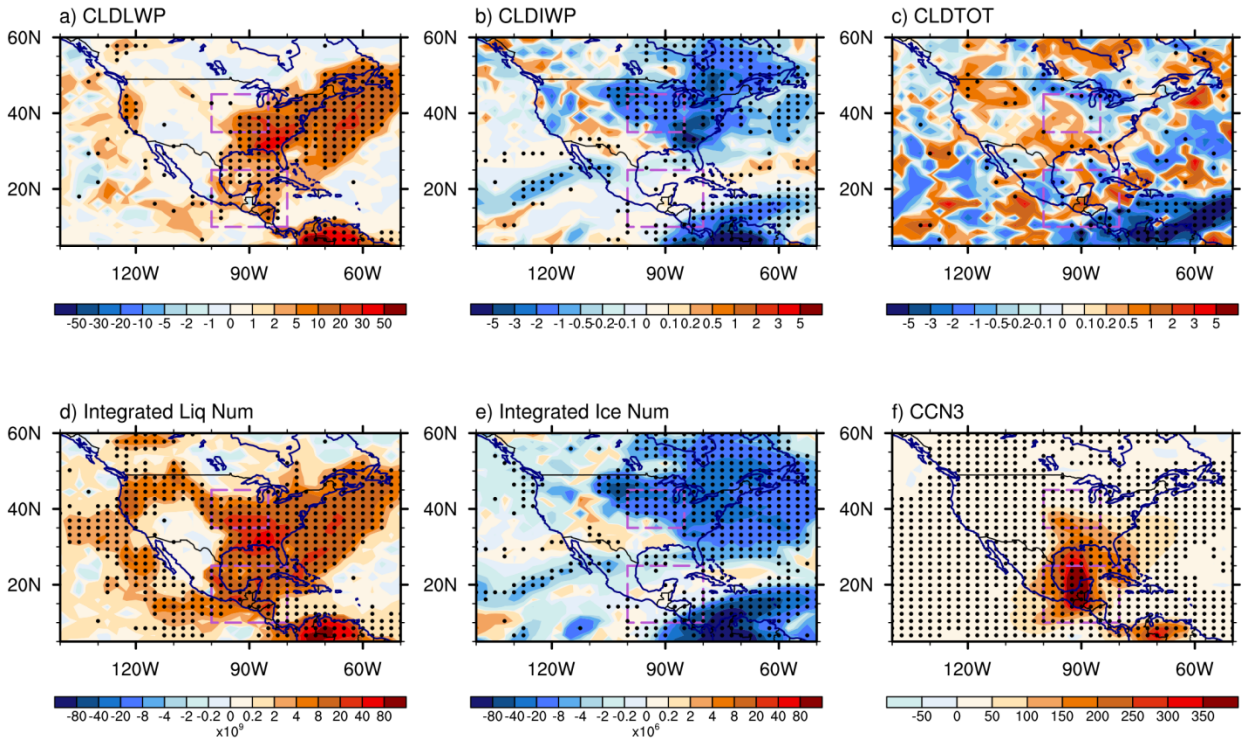

Figure 8. Difference of 10-day average (Apr.1-10) ensemble mean between simulations E_NF and E_QF: a) cloud liquid water path ($g\,m^{-2}$), b) cloud ice water path ($g\,m^{-2}$), c) total cloud fraction (%), d) column-integrated droplet number concentration (m^{-2}), e) column-integrated ice number concentration (m^{-2}), and f) cloud condensation nuclei at 0.1% supersaturation near 900 hPa. Dots denote regions where the difference is statistically significant at the 95% confidence level based on the KS test.

.

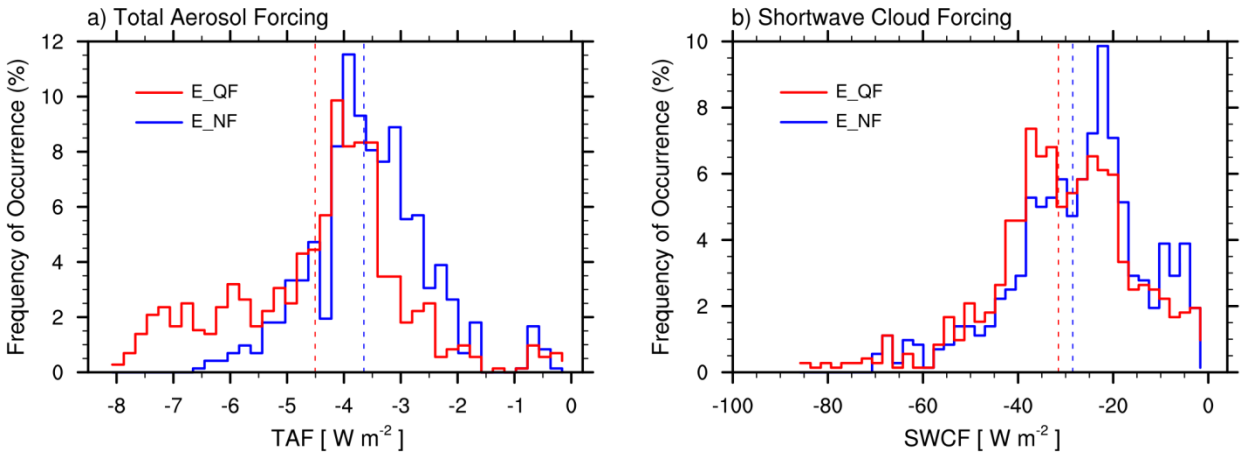

Figure 9. Probability distributions of 10-day average (Apr.1-10) a) total aerosol forcing and b) cloud forcing over Southern Mexico in simulations E_NF and E_QF sampled from grid values of ensemble members (72x10 for each case). Dashed lines indicate the mean of the distribution.

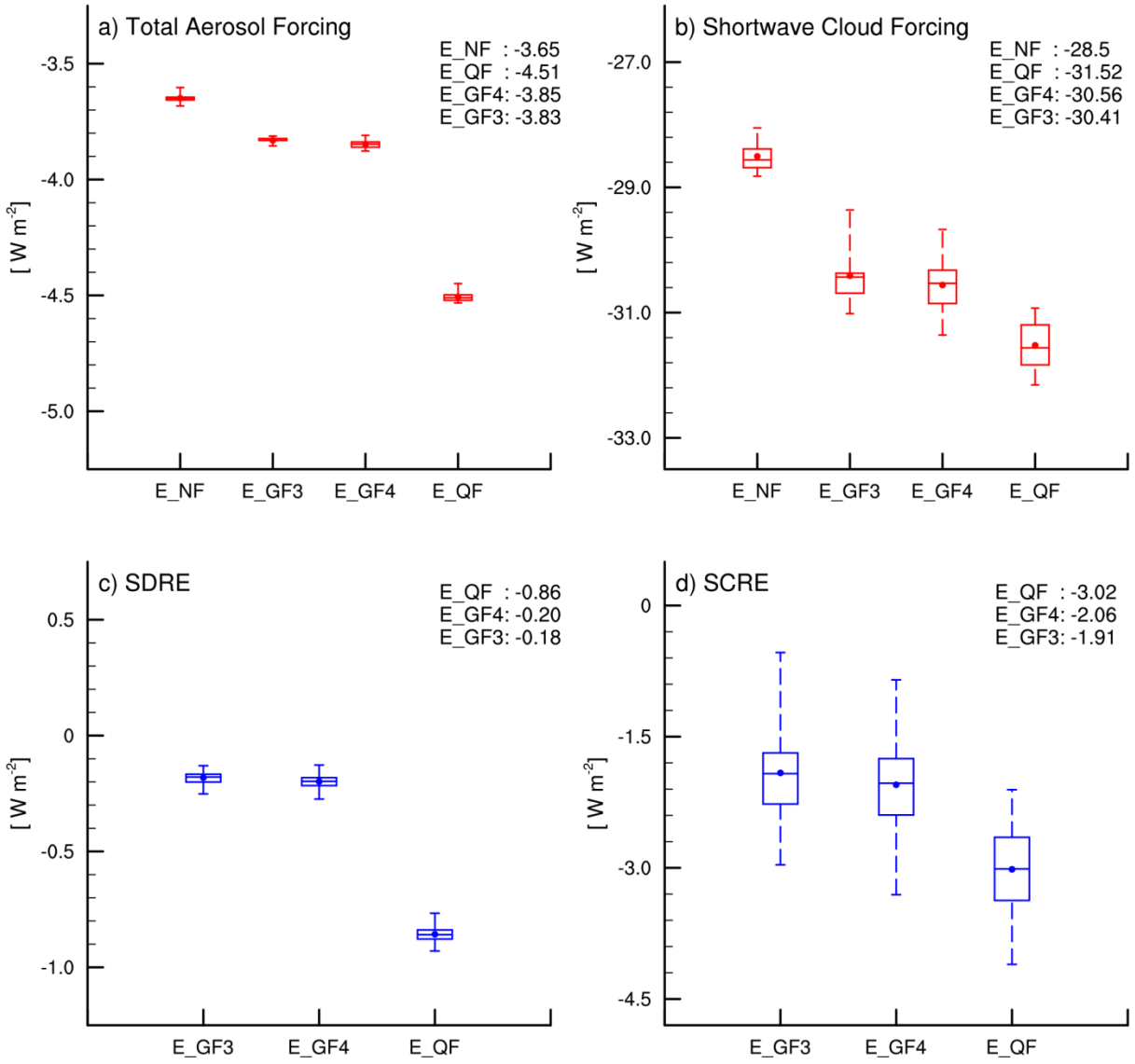

Figure 10. 10-day average (Apr. 1-10) regional mean a) total aerosol direct forcing, b) total shortwave cloud forcing and fire aerosol, c) SDRE, and d) SCRE in Southern Mexico in group B simulations. Box denotes the 25th and 75th percentiles. Bars outside the box indicate minimum and maximum. Bar within the box denotes the 50th percentile. Total aerosol and cloud forcing are sampled from different ensemble members (10 for each case). Fire aerosol SDRE and SCRF are sampled by calculating the difference between members in simulations E_QF (E_GF3/E_GF4) and E_NF (10x10 for each case).

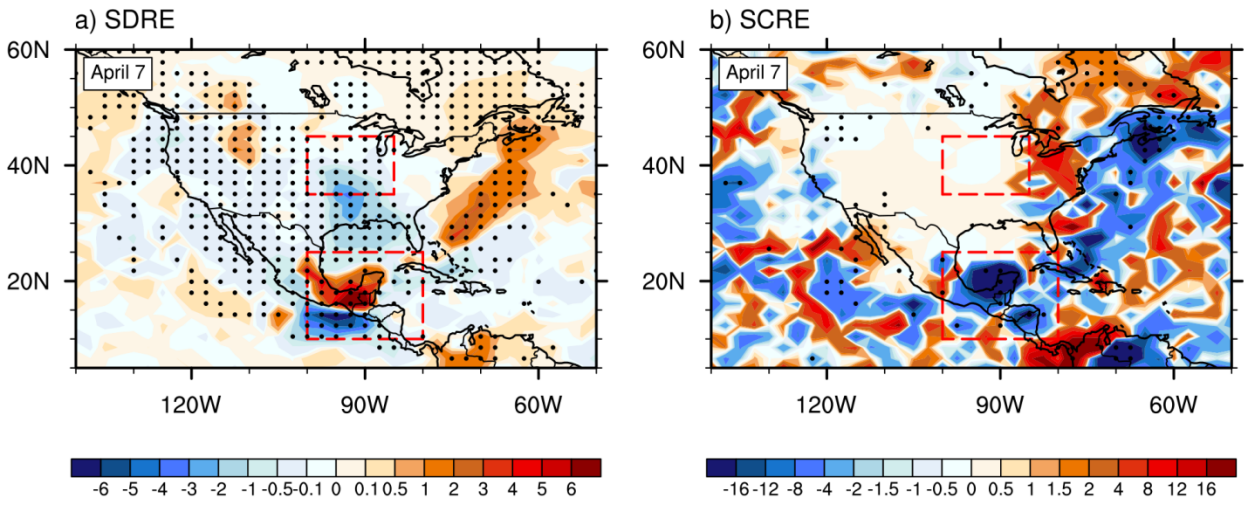

Figure 11. Spatial distributions of ensemble mean fire aerosol a) SDRE and b) SCRE ($W\ m^{-2}$) on April 7 in the E_QF simulation. Dots denote grids where fire aerosol effect is statistically significant at the 95% confidence level based on the KS test.

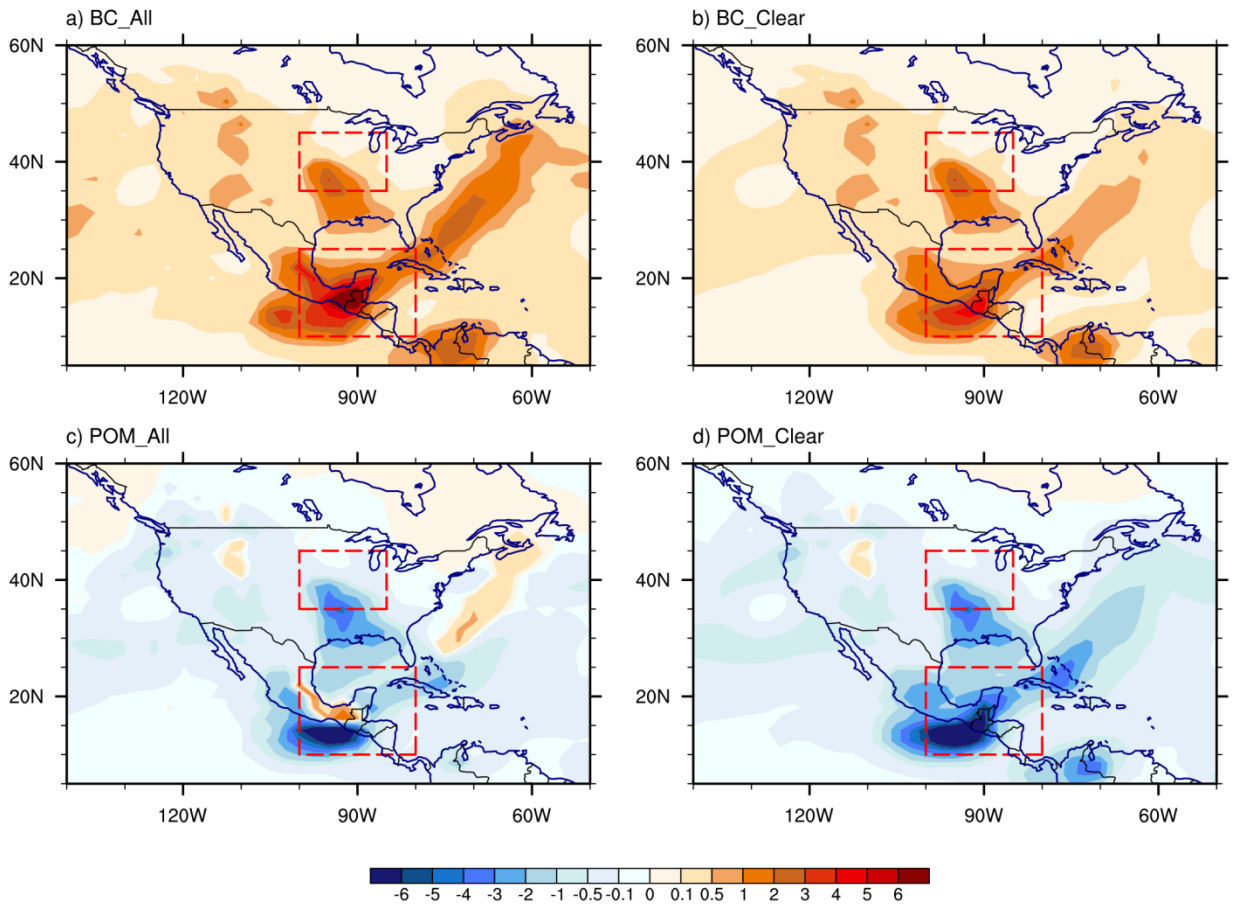

Figure 12. Spatial distributions of fire BC SDRE and fire POM SDRE ($W\ m^{-2}$) on all-sky and clear-sky conditions on April 7 in the E_QF simulation.

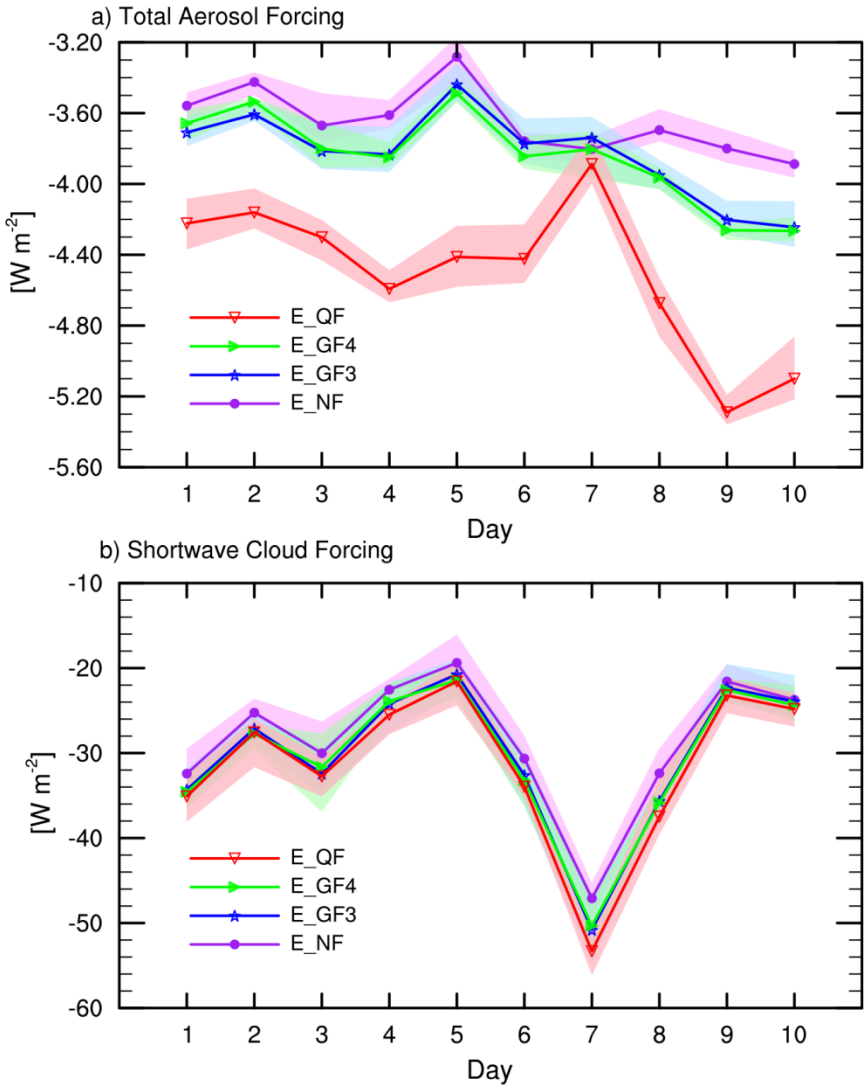

Figure 13. Time series of daily regional mean total a) aerosol forcing and b) cloud forcing in Southern Mexico during Apr.1-10, 2009 in group B simulations. Individual lines indicate ensemble mean values. Shaded areas illustrate the ensemble spread (from minimum to maximum).

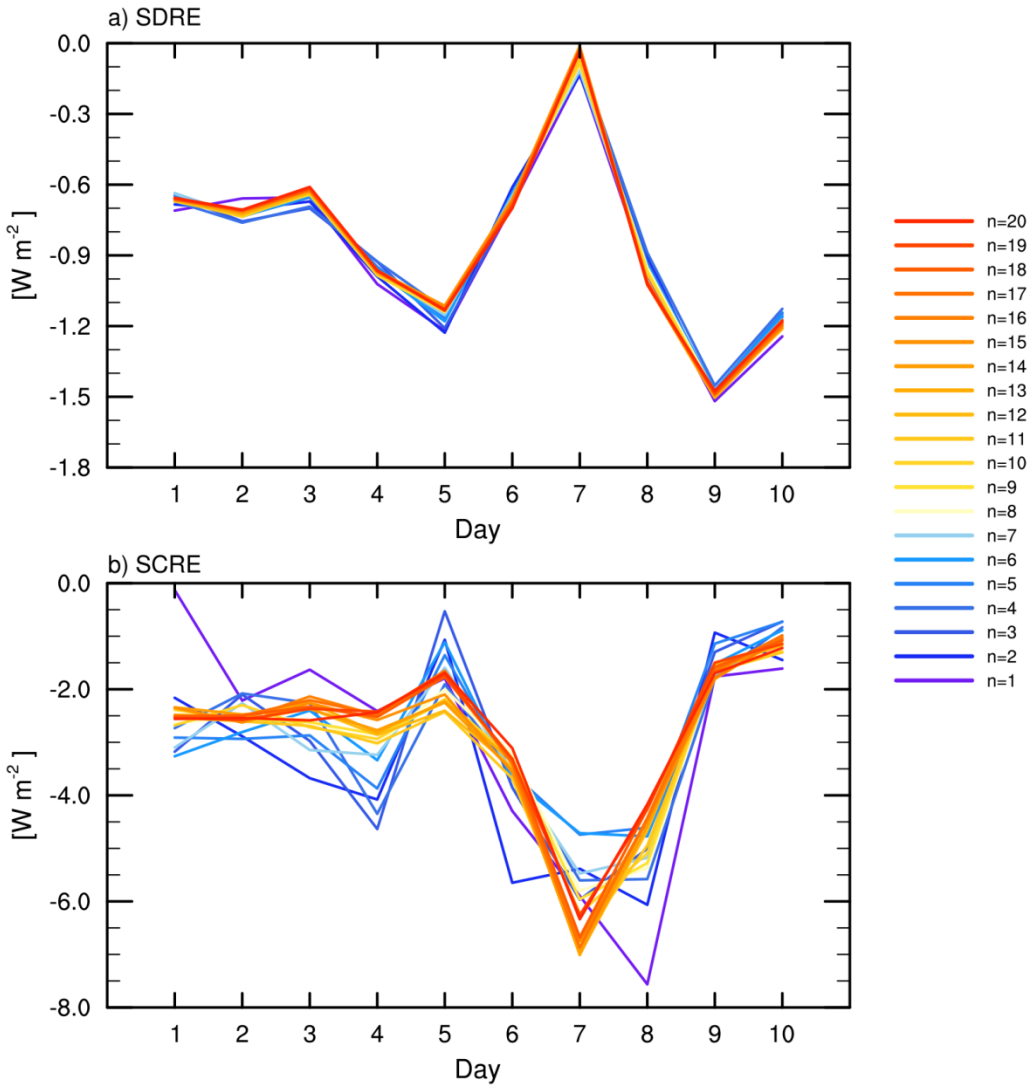

Figure 14. Time series of daily ensemble mean fire aerosol a) SDRE and b) SCRE averaged over Southern Mexico during Apr. 1-10, 2009 in QFED forced ensemble simulations with varying the total number of member numbers (n=1-20).

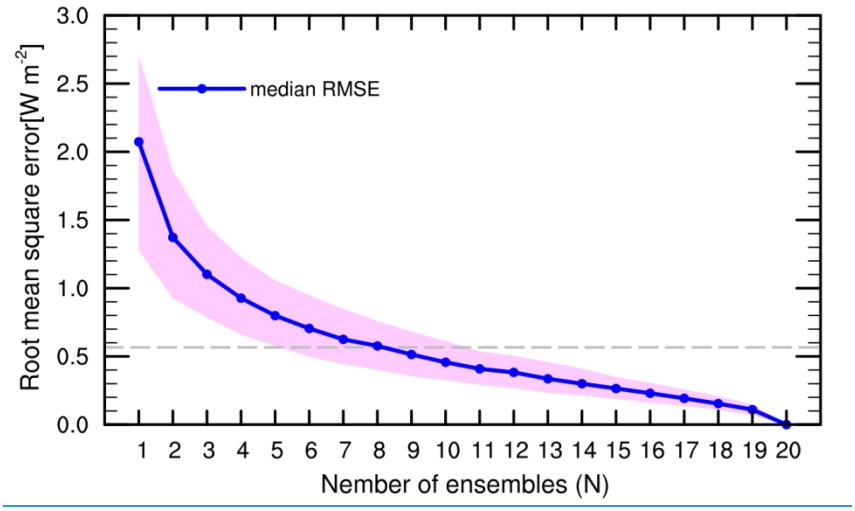

Figure 15 Root mean square errors (RMSE) of the ensemble mean of the regional mean fire aerosol SCRE during April 1-10 over Southern Mexico in simulations with different total number of ensemble members (N). The blue line represents the median RMSE of the 1000 groups (each group has N members/simulations). The grey line represents the threshold RMSE. Shaded area denotes the range between the 10^{th} and 90^{th} percentiles.

Supplemental Materials

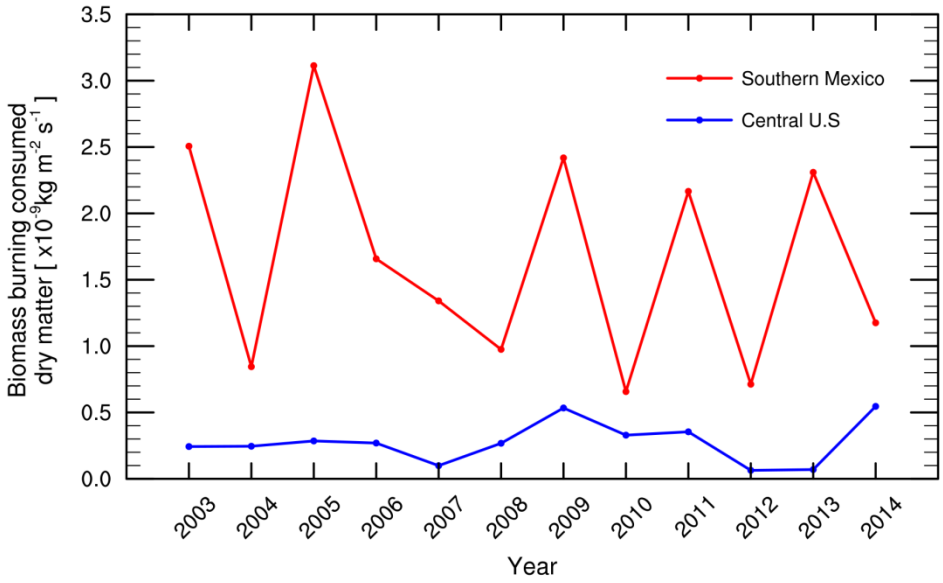

Figure S1. Time series of regional mean biomass burning consumed dry matter during April in central U.S (blue) and Mexico (red) from GFED v4.1.

Table S1 Regional mean emissions of fire aerosols in April, 2009 from three emission inventories (Unit: x10^{-12} kg $m^{-2}s^{-1}$). Numbers in the parentheses show results averaged in April 1-10.

	BC		OC		SO2	
	Central U.S.	Southern Mexico	Central U.S.	Southern Mexico	Central U.S.	Southern Mexico
GFED v3.1	0.25(0.38)	0.69(0.82)	1.82(3.58)	5.60(6.77)	1.35(2.01)	3.69(4.35)
GFED v4.1s	0.23(0.34)	1.17(1.44)	1.75(3.24)	8.80(10.76)	1.21(1.81)	6.25(7.69)
QFED v2.4	2.63(3.29)	3.87(3.87)	23.54(32.25)	36.81(36.58)	14.04(17.59)	20.62(20.65)

Table S2 Regional mean total AOD, fire AOD (difference in total AOD between simulations with and without fire) and the contributions of fire AOD (fire AOD divided by total AOD in the S_NF simulation)during April, 2009 in group A simulations.

	Central U.S.			Southern Mexico		
	Total AOD	Fire AOD	Percentage	Total AOD	Fire AOD	Percentage
S_NF	0.066			0.130		
S_GF3	0.068	0.002	3.42%	0.141	0.011	8.10%
S_GF4	0.070	0.004	5.63%	0.145	0.015	11.20%
S_QF	0.099	0.033	49.33%	0.194	0.064	48.84%

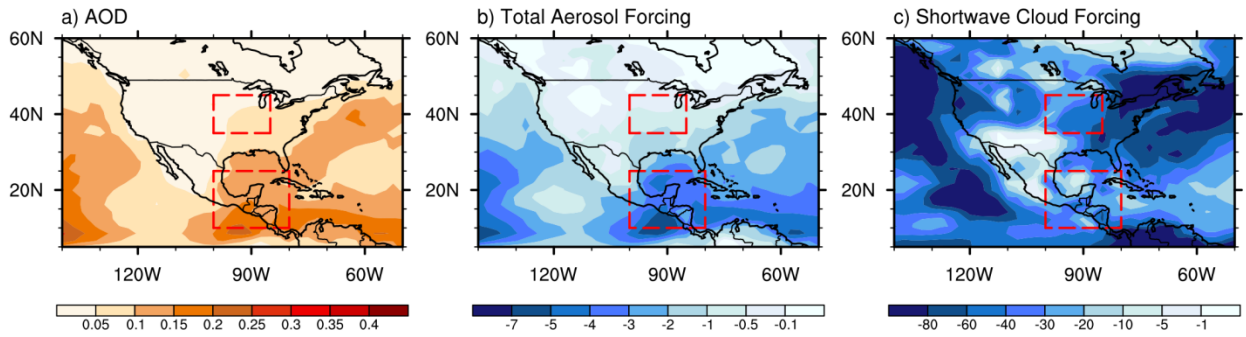

Figure S2. Spatial distributions of 10-day average (Apr. 1-10) ensemble mean a) AOD, b) total aerosol forcing and c) total shortwave cloud forcing($W\ m^{-2}$) in the simulation without fire emissions (E_NF).

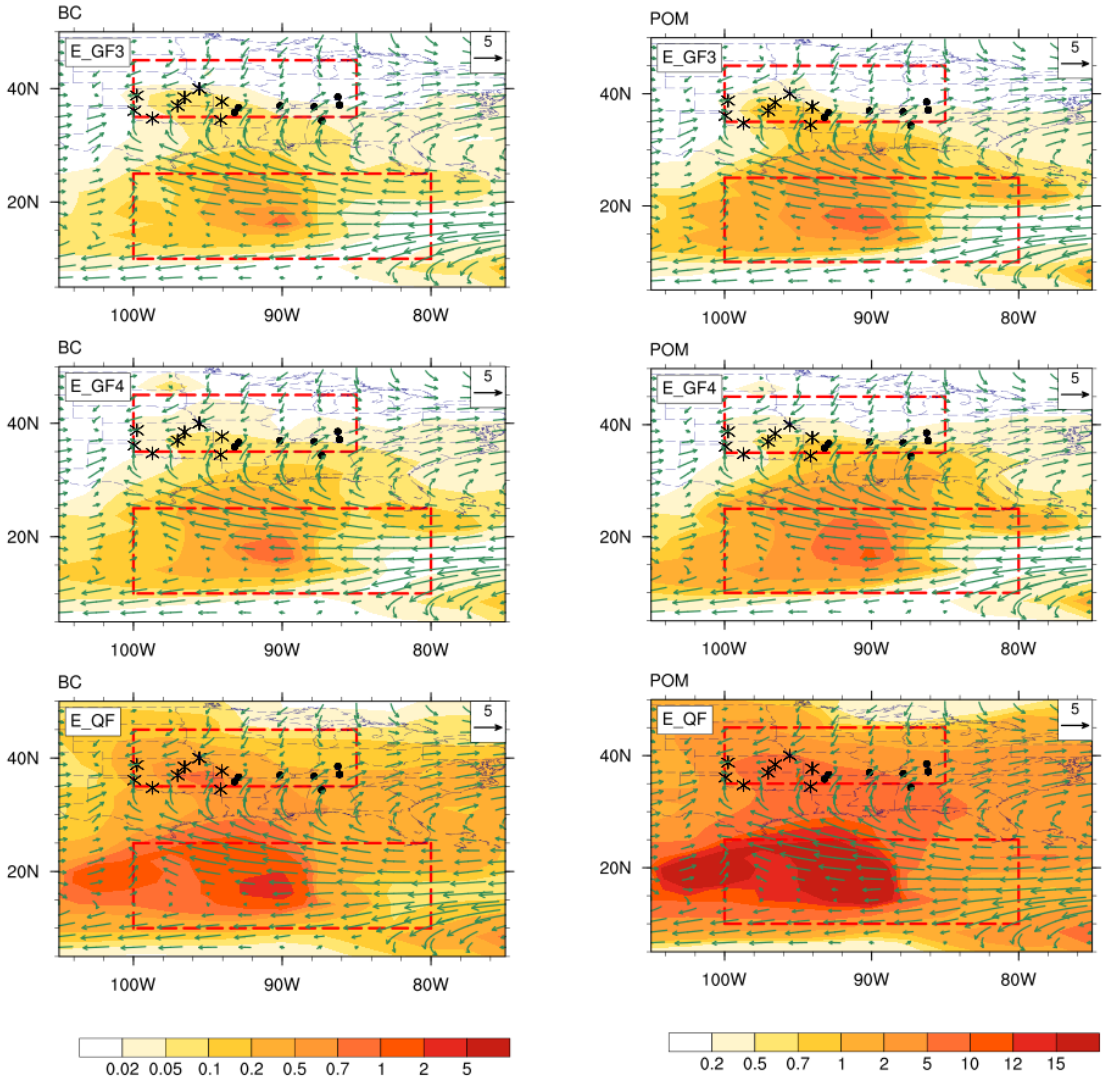

Figure S3. Spatial distributions of April mean fire BC and fire POM burden (shaded) on IMPROVE observation days in group B simulations (E_GF3/E_GF4/E_QF – E_NF). Vectors denote horizontal winds near 850hPa in group B fire simulations (E_GF3/E_GF4/E_NF). IMPROVE data sites are marked with asterisks for sites near the source region and with dots for sites in the downwind region.

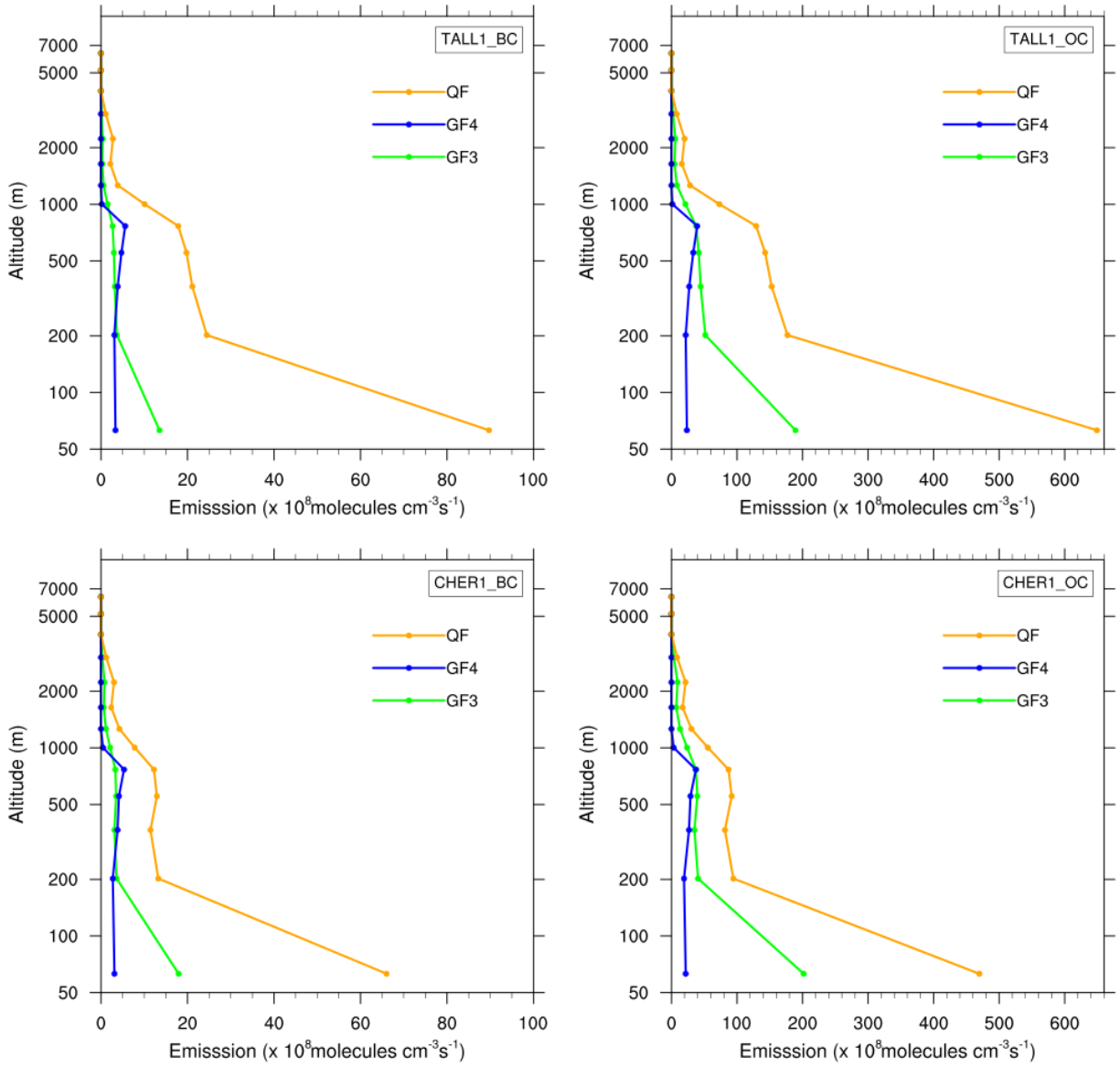

Figure S4. Vertical profiles of fire emissions of BC and OC used in simulations at sites TALL1 (38.43°N, 96.56°W) and CHER1 (38.77°N, 99.76°W).

Table S3 Regional mean total AOD, fire AOD (differences in AOD between simulations with and without fire) and radiative effects of fire aerosols during April 1-10, 2009 in group B simulations (Unit: $W\ m^{-2}$). Total fire aerosol radiative effect is decomposed into shortwave direct radiative effect (SDRE), shortwave cloud radiative effect (SCRE), longwave cloud radiative effect (LCRE) and surface albedo effect (SAE).

	Total AOD	Fire AOD	SDRE	SCRE	LCRE	Total SAE
Central U.S.						
E_NF	0.047					
E_GF3	0.050	0.003	0.02	-0.86	0.04	0.02
E_GF4	0.050	0.003	-0.01	-0.39	0.002	-0.003
E_QF	0.08	0.033	-0.10	-0.56	-0.76	0.12
Southern Mexico						
E_NF	0.135					
E_GF3	0.149	0.014	-0.18	-1.91	-0.21	0.06
E_GF4	0.153	0.018	-0.20	-2.06	-0.23	0.11
E_QF	0.202	0.067	-0.86	-3.02	-0.47	0.14

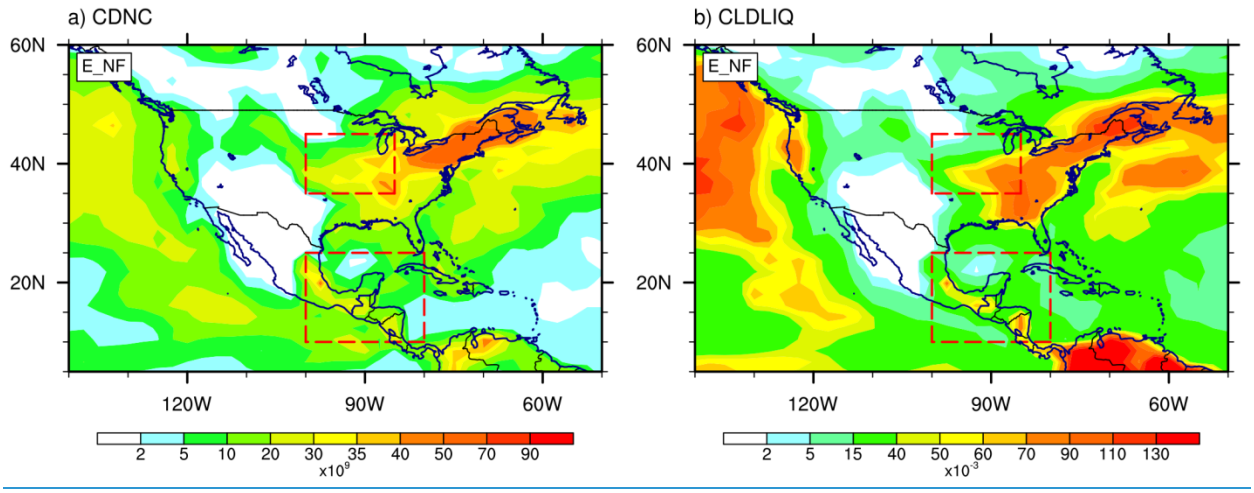

Figure S5. Spatial distributions of 10-day average (Apr. 1-10) ensemble mean a) column-integrated droplet number concentrations (m^{-2}) and b) liquid water path (g m^{-2}) in the E_NF simulations.

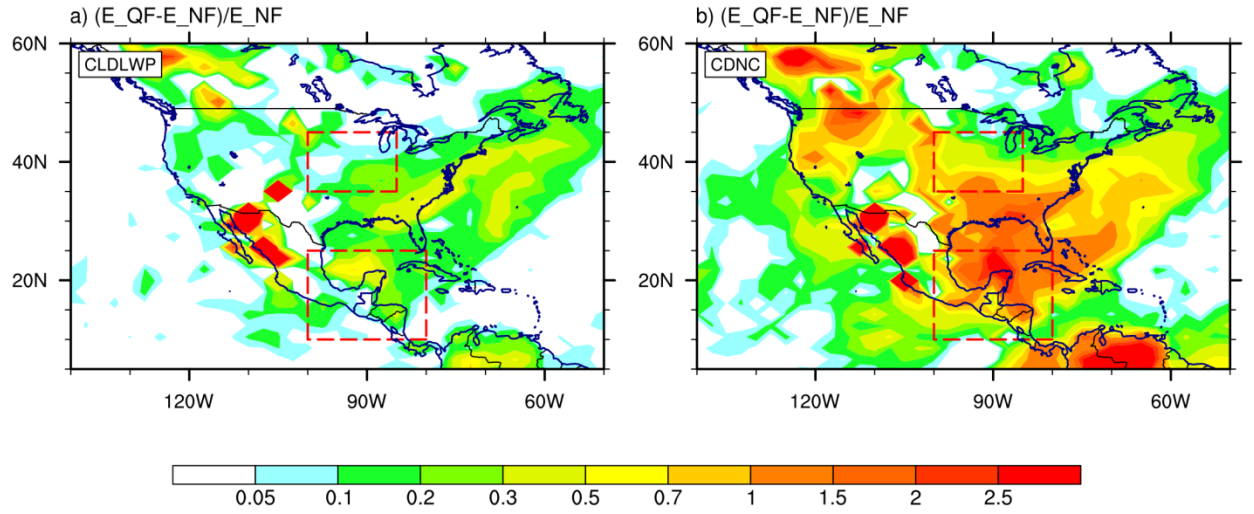

Figure S6. Relative changes of 10-day average ensemble mean cloud properties between the E_NF and E_QF simulations. a) cloud liquid water path, b) column-integrated droplet number concentration

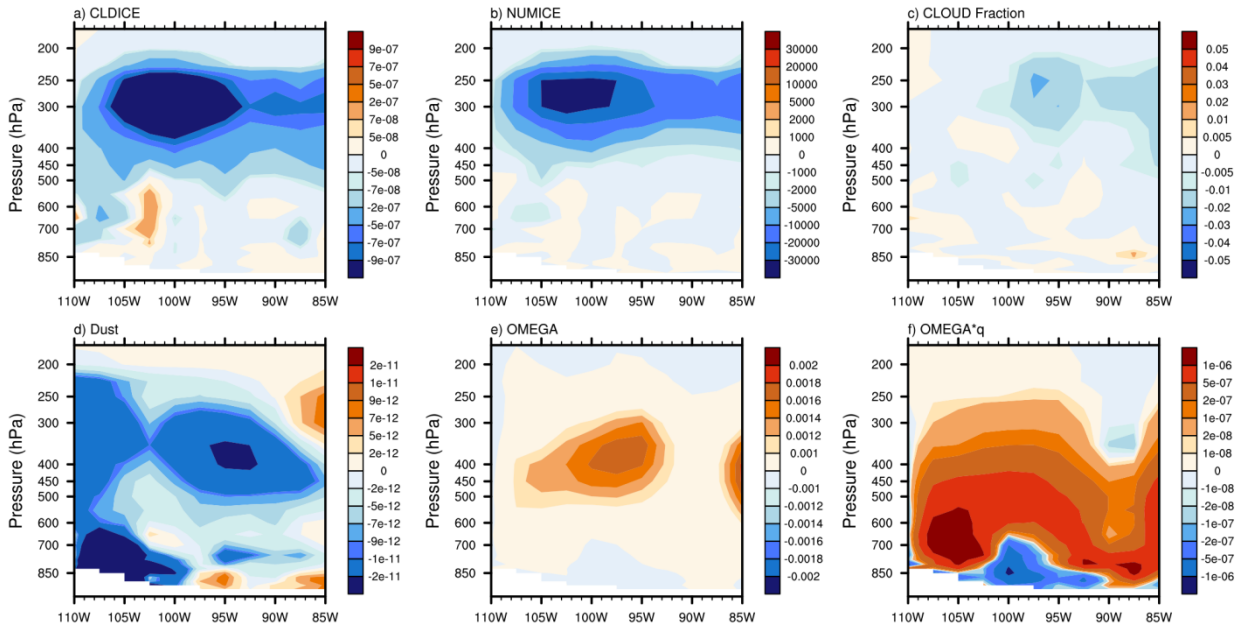

Figure S7. Pressure and longitude distribution of meridional mean (40-45 ° N) difference of 10-day average (April 1 -10) ensemble mean between simulation E_NF and E_QF: a) cloud ice amount ($kg \cdot kg^{-1}$) b) cloud ice number concentration (kg^{-1}) c) cloud fraction (1) d) Coarse mode dust concentration ($kg \cdot kg^{-1}$) e) vertical velocity ($Pa \cdot s^{-1}$) f) vertical moisture transport ($kg \cdot kg^{-1} \cdot Pa \cdot s^{-1}$)

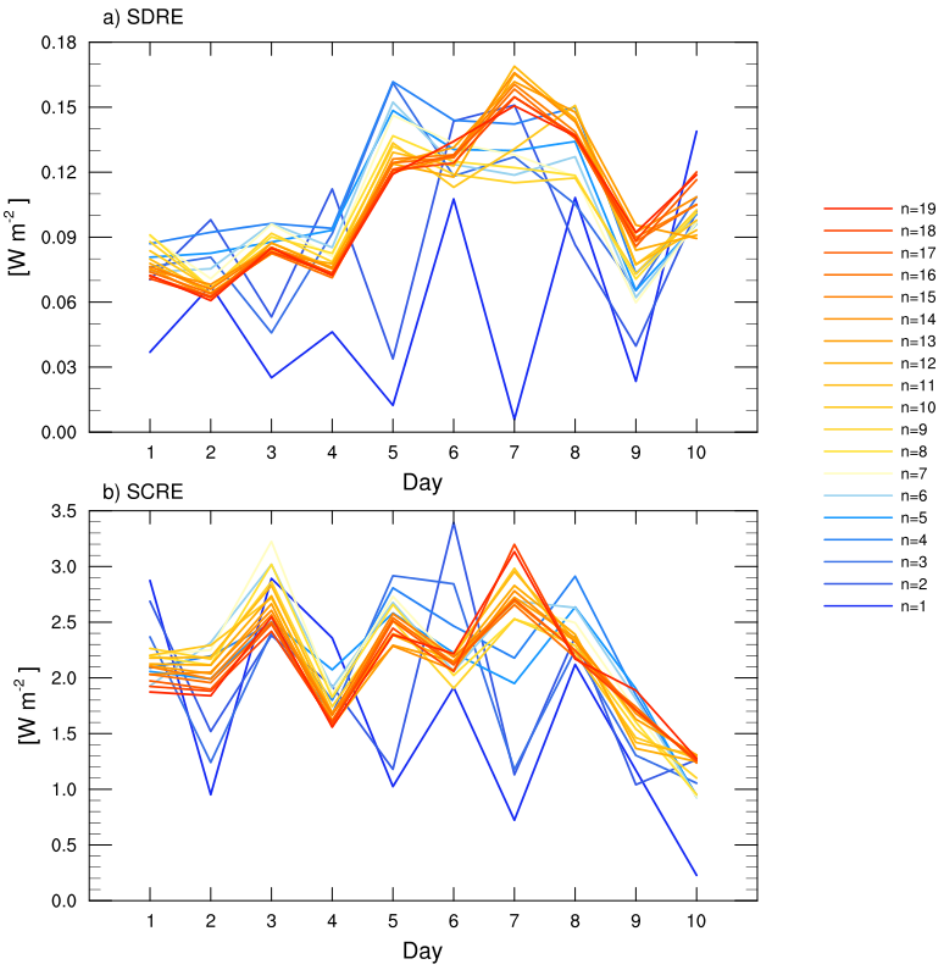

Figure S8~~5~~. Time series of ensemble spread of daily regional mean fire aerosol a) SDRE and b) SCRE in Southern Mexico during Apr. 1-10, 2009 in QFED forced ensemble simulations with varying the total number of ensemble member (n=1-20).