Reply to comments from Referee #2 on "Can a coupled meteorology-chemistry model reproduce the historical trend in aerosol direct radiative effects over the Northern Hemisphere?" by Xing et al.

We thank the referee for a very thoughtful review and an overall positive assessment of our manuscript. Incorporation of the reviewer's suggestions has led to a much improved manuscript. Below we provide a point-by-point response to the reviewer's comments and also detail how we have addressed them in the revised manuscript.

General

[Comment]: *suggest the authors include a short discussion on steps they are taking to address identified shortcomings in emissions and model parameterizations to improve future similar model investigations.*

[Response]: Following the reviewer's suggestion, in the revised manuscript we have included a discussion on how we will be addressing in the future identified shortcomings in the model emissions and parameterizations, as follows:

(P27, L19-P28, L2) "An accurate temporally resolved biomass emission inventory (van der Werf et al., 2006; Shi and Yamaguchi, 2014), an improved dust emission model (Kok et al, 2014) and an advance scheme to model atmospheric organic aerosol (Koo et al., 2014; Zhao et al., 2015) are suggested for future model investigations. We are currently conducting a similar study with a finer-scale simulation and relatively better spatially and temporally resolved emission inventories over the continental U.S. domain. Further analysis of those model calculations and assessment of the impacts of the higher resolution emissions can be found in Gan et al. (2015)."

References:

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr., A. F.: Interannual variability in global biomass burning emissions from 1997 to 2004, Atmos. Chem. Phys., 6, 3423-3441, doi:10.5194/acp-6-3423-2006, 2006.

Shi, Y., and Yamaguchi, Y.: A high-resolution and multi-year emissions inventory for biomass burning in Southeast Asia during 2001–2010. *Atmospheric Environment*, *98*, 8-16, 2014.

Kok, J. F., Mahowald, N. M., Fratini, G., Gillies, J. A., Ishizuka, M., Leys, J. F., Mikami, M., Park, M.-S., Park, S.-U., Van Pelt, R. S., and Zobeck, T. M.: An improved dust emission model – Part 1: Model description and comparison against measurements, Atmos. Chem. Phys., 14, 13023-13041, doi:10.5194/acp-14-13023-2014, 2014.

Koo, B., Knipping, E., and Yarwood, G.: 1.5-Dimensional volatility basis set approach for modeling organic aerosol in CAMx and CMAQ. Atmospheric Environment, 95, 158-164, 2014.

Zhao, B., Wang, S., Donahue, N. M., Chuang, W., Hildebrandt Ruiz, L., Ng, N. L., Wang, Y., and Hao, J.: Evaluation of One-Dimensional and Two-Dimensional Volatility Basis Sets in Simulating the Aging of Secondary Organic Aerosol with Smog-Chamber Experiments. Environmental science & technology, 49(4), 2245-2254, 2015.

Gan, C.-M., Pleim, J., Mathur, R., Hogrefe, C., Long, C. N., Xing, J., Wong, D., Gilliam, R., and Wei, C.: Assessment of multi-decadal WRF-CMAQ simulations for understanding direct aerosol effects on radiation "brightening" in the United States, Atmos. Chem. Phys. Discuss., 15, 17711-17742, doi:10.5194/acpd-15-17711-2015, 2015."

Specific Comments

[Comment]: *p. 14032, line 26: I believe it should be "retrievals" instead of "retrieves".*

[Response]: The typo has been corrected in the revised manuscript (P6, L22).

[Comment]: *p. 14033, line 23: Should be "... the satellite datasets are interpolated ...".*

[Response]: The typo has been corrected in the revised manuscript (P7, L19).

Reply to comments from Referee #3 on "Can a coupled meteorology-chemistry model reproduce the historical trend in aerosol direct radiative effects over the Northern Hemisphere?" by Xing et al.

We would like to thank the reviewer for a very thoughtful and detailed review of our manuscript. Incorporation of the reviewer's suggestion has led to an improved manuscript. Detailed below is our response to the issues raised by the reviewer. We also detail the specific changes incorporated in the revised manuscript in response to the reviewer's comments.

General

[Comment]: *I would suggest the authors take a more rigorous approach to assessing the trend estimates, particularly the uncertainties in the trends, and avoid the desire to jump to conclusions. I would suggest assessing the trends at both p=0.05 and p=0.1 (perhaps even p=0.2) levels of confidence and including a discussion of the differences in the significance of the trends in the text. The limitations of the data must be recognized and my hope here is that the clear negative trends in EUS and EUR can be assessed at the level of significance allowed by the data and that these regional trends can be contrasted against the puzzling lack of any positive trends in AOD over ECH.*

[Response]: As the reviewer suggested, we have revised Table 2. The significance at three levels, i.e., p=0.05, 0.1 and 0.2 is calculated for observed and simulated trends. The SeaWiFS retrieved AOD exhibits a positive trend (significant at p=0.1 level) in AOD over ECH. Declining trends shown in both observed and simulated AODs in EUS and EUR are significant (at least at p =0.2 level). We have included additional description about the trend and its significance in the revised manuscript, as below:

(P9, L19-21) The linear least square fit method was employed, and significance of trends was examined with a Student t test at the 80%, 90% and 95 % confidence levels (p = 0.2, 0.1 and 0.05).

(P11, L17-19) The SeaWiFS retrieved AOD presents a more significant (at p=0.1 level) increasing trend compared to other satellite products, though the growth rate (+0.004 yr^{-1}) is still lower than that of simulated AOD trend (+0.014 yr^{-1}).

(P11, L22-23) Additionally, declining trends in both observed and simulated AODs in EUS and EUR are significant (at least at p =0.2 level).

(P12, L2-7) The opposite trend (though not significant at p=0.2 level) estimated from both AVHRR and TOMS might be explained by limited number of grid values available for calculation (refer to Figure 3). Both simulation and satellite retrievals (except TOMS in EUR which has no trend) show declining trends in EUS and EUR before 2000 with magnitudes comparable to those for the period post-2000, and the declining trend in EUR is significant (at p=0.2 level) in both simulated and AVHRR retrieved AOD.

(P12, L15-19) However, after 2000, the increase of transport from ECH results in an increasing trend shown in both simulated and most satellite-retrieved AOD (significant at p=0.2 level for SeaWiFS and MISR) in NPA of +0.003 (+2%) yr^{-1} and +0.001 to +0.002 (+0.3 to +1.7%) yr^{-1} respectively, except in AVHRR retrieved AOD which gives opposite trend.

[Comment]: *My other concern is the use of the term 'non-feedback' to describe the second model experiment. It is stated that the distinction between the 'feedback' and 'non-feedback' experiments is that only the feedback case has 'aerosol direct radiative effects updated in the rapid radiative transfer model...'. Does this mean that the non-feedback case uses climatological aerosol fields in the model radiation code? Even more fundamentally, if the comparison of quantities like TOA shortwave radiation (SWR) shown in Figure 7 is taken from the output of the model and the output results from the input of a constant aerosol climatology, what is being compared? I'll note that the 'no feedback' case shown in Figure 7 produces an almost constant TOA SWR for regions where the feedback case shows significant trends. My concern here is that in the literature 'feedback' and 'no feedback' is used to distinguish between setups where the internal model chemical fields (gas-phase and/or aerosols) are, or are not, allowed to affect the dynamical fields, but in both cases the aerosol fields are calculated internally within the model. One may find systematic effects by allowing these feedbacks, but the model aerosol fields are still largely the same. Here there seems to be two different representations of the aerosols producing two fundamentally different effects on aerosol quantities like AOD. I would suggest a less ambiguous term than 'feedback' to differentiate the experiments along with a more complete description of how the 'non-feedback' experiment is setup. Further, if the non-feedback experiment uses a climatological aerosol, is there any validity in comparing the estimates of direct radiative efficiency and correlations with observations as shown in Table 4?*
[Response]: we agree with the reviewer that our original statement need to be clarified. In the case defined "non-feedback" in this study, no default or climatological aerosol profiles were provided to the RRTMG model employed in WRF (Hogrefe et al., 2015). Therefore, the difference between feedback and non-feedback simulations are suggestive of the aerosol radiative effects. In addition, this study used a coupled model which allows the DRE to affect the dynamical fields and subsequently influence on aerosol quantities associated with the "updated" dynamical fields, such as soil dust and photolysis which are calculated online with the dynamical model. To help clarify the reviewer's concern, we have provided additional description about these aspects in the revised manuscript, as below:

(P5, L18-P6, L1) "Similar to Hogrefe et al (2015), in the non-feedback case, no default or climatological aerosol profiles were provided to the RRTMG model employed in WRF. Therefore, there are no aerosol effects on the radiation calculations in the non-feedback case. DRE is thus estimated as the difference between feedback and non-feedback simulations. In addition, the coupled model used in this study allows the DRE to affect the dynamical fields and also represent the subsequent modulation of aerosol quantities associated with the "updated" dynamical fields, such as soil dust emission flux and photolysis rates which are calculated online with the dynamical model."

Specific Comments

[Comment]: *Page 14031, Line 25 to Page 14032, Line3: As stated just above, a more complete description of how the non-feedback experiment is designed would help the reader.*
[Response]: As detailed earlier, following the reviewer suggestion, in the revised manuscript we have modified the discussion (P5, L18-P6, L1) to provided additional description on how the feedback and non-feedback runs were designed and what the difference between the two

simulations imply.

[Comment]: *Page 14035, Lines 14-17: That trends were estimated separately over the 1990s and 2000s eventually becomes clear, but the process of separating the decadal trends should be introduced clearly.*
[Response]: We have clarified it in the revised manuscript as below:
(P9, L21-23) "To be consistent with the available period of satellite products mostly covering the period of the 2000s, the two decades were separated into two time periods spanning 1990s (1990-2010) and 2000s (2000-2010) for analysis."

[Comment]: *Page 14035, Line 23: It is here that Figure 2 is introduced. Can I suggest that regions without data be shaded grey in the panels showing trends to allow differentiation from regions with small trends? And are the trends shown in Figure 2 calculated over the full length of the record for each satellite? There is an oblique reference at Page 14036, Line 21 that suggests the trends are all calculated for the post-2000 period, but it should be clearly stated in the figure caption.*
[Response]: As the reviewer suggested, we have revised Figure 2 in which the regions without data are shaded grey. The trends shown in Figure 2 are calculated over the full length of the record for each satellite. We have clarified this aspect in Figure 2 in the revised manuscript.

[Comment]: *Page 14036, Line 27 to Page 14037, Line 4: It is the JJA season that is analysed here and I understand the biomass burning in south-east Asia (Myanmar, Thailand, Laos, Vietnam) peaks earlier in the year – February through April (see van der Werf et al., Global Change Biology, 9, 547-562, 2003). Are the authors certain that problems with biomass burning are the cause for the discrepancy between the modelled and observed trends over south-east China?.*
[Response]: We agree with the reviewer that the biomass burning in south-east Asia during summer (time period of analysis in this study) is not the peak season for fire activity in the region. Nevertheless, the effects of biomass burning are still noticeable in the satellite retrieved AOD trend (as see in Figure 2) in south-east Asia. Additionally, the model failed to capture the negative trend in the south-east Asia. Some studies (e.g., Lin et al., 2010) also suggest that the reduction in satellite observed AOD in northeastern China since late 2008 is associated with the recession which might be not well represented in the emission inventory. We have clarified this aspect in the revised manuscript as below:
(P11, L12-17) "As displayed in Figure 2, the model failed to capture the negative trend of the satellite retrieved AODs in the south-east Asia. Variations in biomass burning activity in Southeast Asia is difficult to capture in the model without an accurate temporally resolved biomass emission inventory, currently not available. The declining trends shown in the observed AOD may be also associated with the recession since late 2008 (Lin et al., 2010) which may be not well represented in the emission inventory"

Reference:
Lin, J., Nielsen, C. P., Zhao, Y., Lei, Y., Liu, Y., and McElroy, M. B.: Recent changes in particulate air pollution over China observed from space and the ground: effectiveness of emission control.

Environmental science & technology, 44(20), 7771-7776, 2010.

[Comment]: *Page 14037, Lines 16-19: The discussion here ties the declining AOD found in the north Pacific (NPA) with the declining AOD in the eastern US (EUS). Given the west to east transport of pollutants, is it realistic to think that declining emissions (assuming the decline in AOD is directly related to emissions over the source region) in EUS contribute to a decline in aerosol amounts three-quarters of the way around the world? There has been some very interesting work tying decadal-scale trends in ozone at Mauna Loa with shifts in the large-scale circulation. I would suggest this as a more probable, but still speculative, possibility.*

[Response]: We agree with the reviewer that our original statement was vague and perhaps misleading. We further investigated on the AOD trend in NPA during 1990s, and found the negative trend is more likely associated with the declining trend in the global background AOD from biomass burning in southeastern Asia, dust variations in Sahara desert and anthropogenic emission reduction in Japan, Europe and North America. Quantification of those impacts on such interesting trends in NAP needs further investigation, however, it is beyond of the scope of the current study. We've clarified it in the revised manuscript, as below:

(P5, L10-12) "Out of the four ocean regions, the North Pacific (NPA, 30N-50N, 150E-130W) is in between ECH and EUS, the North Atlantic (NAT, 35N-50N, 60W-15W) is in between EUR and EUS.

(P12, L10-18) "The AOD trends in North Pacific (NPA) before-2000 could possibly be associated with declining trend in global background AOD from biomass burning in southeastern Asia, dust variations in Sahara desert and anthropogenic emission reduction in Japan, Europe and North America. For the period before 2000, both simulated and satellite-retrieved AOD in NPA exhibit declining trends of -0.003 (-2.08%) yr^{-1} and -0.001 to -0.002 (-0.5 to -0.6%) yr^{-1} respectively. However, after 2000, the increase of transport from ECH results in an increasing trend shown in both simulated and most satellite-retrieved AOD (significant at p=0.2 level for SeaWiFS and MISR) in NPA of +0.003 (+2%) yr^{-1} and +0.001 to +0.002 (+0.3 to +1.7%) yr^{-1} respectively, except in AVHRR retrieved AOD which exhibits opposite trend."

[Comment]: *Page 14040, Lines 1-3: Account must be made of the fact that the Mauna Loa AERONET station sits at an elevation of 3400m. This would seem to be the reason for the very low AOD shown in Figure 5.*

[Response]: As the reviewer suggested, we have added the following discussion in the revised manuscript:

(P15, L10-12) " The AERONET monitored AOD is even lower than the satellite retrievals, which might be explained by the high elevation (3400m) of this AERONET site."

[Comment]: *Page 14040, Line 26 – The text says the SWR at TOA shown in Figure 6 is for 2000* *and 2010, though the text states 1990 and 2010.*

[Response]: We have fixed the typo in the revised manuscript.

(P16, L9-10) "The spatial distributions of observed and simulated clear-sky SWR at TOA are presented in Figure 6 for 2000 and 2010"

[Comment]: *Page 14047, discussion beginning at line 17: I had tremendous difficult with the discussion of how Et* becomes smaller than Et when AOD levels are higher. Given that Et and Et* are measures of the radiative efficiency and the radiative efficiency per unit AOD measured by Et decreases at higher AOD, then Et* should be larger than Et. But I believe the reason the discussion is structured the way it is, is because the DRE is negative and so larger negative values are smaller? For example, for the SHR region shown in panel (a) of Figure 10, the Et* is -18*AOD, but at values of AOD=1, the individual data points have DRE/AOD values of -12. While Et* is smaller than Et numerically, the estimate of the physical effect of aerosols on DRE given by Et* is larger than that given by Et. This is my interpretation of the use of the word 'smaller', but perhaps I am not correctly following the argument? In any case, I would suggest a bit of clarification in the discussion to make it easier to follow. Perhaps defining Et and Et* using the absolute magnitude of DRE so that the numerical relationship between Et and Et* is consistent with the physical interpretation?.*

[Response]: The Eτ* defined in this study is the sensitivity of DRE to the change in AOD, so the Eτ* is calculated from the slope of the linear regression between DRE and AOD. Suggested by Figure 10, the magnitude of Eτ* (the slope) becomes smaller at higher AOD levels. We agree with the reviewer that the original statement is confusing. We have revised this section as below:

(P23, L22-P24, L7) "However, the E_τ^* (estimated as the slope of the linear regression between DRE and AOD) is smaller in magnitude (i.e., absolute value) than E_τ at high AOD levels. This is illustrated in Figure 10 which displays the relationship between average daytime clear-sky DRE and AOD (note that each point in these figures represents a daily average value). It is noticeable that the response of DRE to AOD becomes non-linear at high AOD, more apparent in the relationships for SHR region where the AOD is higher. Consequently, in this regime of high AOD the estimated $E_\tau^* < E_\tau$. Additionally, Yu et al (2006) noted smaller radiative efficiencies for anthropogenic aerosols with higher absorbing components. This may also result in $E_\tau^* < E_\tau$ in regions of anthropogenic aerosols (such as ECH, EUS and EUR)."

[Comment]: *Page 14048, Line 8 and Line 10 – note the inconsistent use of E2t and Et2.*

[Response]: We have fixed the typo in the revised manuscript (P24, L15). We thank the reviewer for catching this typo.

[Comment]: *Page 14048, Lines 13-17: This argument is difficult to understand, particularly the phrase '...though the values of TOA-Et* are smaller, but still slightly larger...'*

[Response]: We have clarified it in the revised manuscript.

"(P24, L19-22) Over land regions where anthropogenic aerosols are dominant, the simulated E_τ at TOA in ECH, EUS and EUR is about -45.4, -49.6 and -57.2 W m^{-2} τ^{-1} respectively, which are slightly larger in magnitude than the measurement-derived TOA-E_τ which is about -9 to -33, -24 to -37 and -11 to -34 W m^{-2} τ^{-1}, respectively."