Interactive comment on "Trans-pacific transport and evolution of aerosols: Evaluation of quasi global WRF-Chem simulation with multiple observations" *by* Zhiyuan Hu et al.

Anonymous Referee #4

This study evaluated a fully coupled meteorology-chemistry model (WRF-Chem) configured to conduct quasi-global simulation for years of 2010-2014 using multiple observation datasets. The evaluation has been focused on the simulation over the trans-Pacific transport region.

After going through the manuscript, though a lot of analysis and comparison have been done between model and observation, I still have some concerns regarding the simulation and results, especially the relative old anthropogenic emission inventory and the way to including biomass burning emissions. Also, as an evaluation paper, I did not see to much quantitative analysis and conclusions when compare the differences between model and observation. In some places, the scientific points are not well presented,

and some presentations and conclusions are more or less like an assumption description or arbitrary statement, which need to provide enough evidences or references to convince the reader.

I recommend this paper for publication in GMD after major revision if the authors satisfactorily address all the comments and questions.

Specific Comments:

1. Section 2.3, the emissions data is a very important input data for the part of chemical transport model. I am surprised that that the author did not use the recently updated anthropogenic emission inventory, e.g. the HTAP v2.1, which has been widely used from last year in a lot of model, including WRF-Chem. Even for the emissions over Asia and US, the MEIC (http://www.meicmodel.org/dataset-mix.html) and NEI 2010 are also the updated version compare to the emission inventory used in the manuscript. When the evaluated results show big differences between model and observations, how do the authors quantify how much is due to emission uncertainty and how much is due to model performance in simulating the long-range transport? So I strongly recommend the authors to use the recent anthropogenic emissions or the updated version.

2. Section 2.3: I am wondering how did the authors include the GFED3 biomass burning emission into the model? Normally, the standard WRF-Chem code uses PREP-CHEM-SOURCE to generate and include the GFED3 biomass burning emission in the forecast. However, it can only generate the emission rate of the first simulation day (month) if the authors did not cycle the chemistry in the simulation everyday (every month). Otherwise, the authors should modify the code update (read) the GFED3 biomass burning every day (every month). Actually, the GFED3 has daily data, why did the author only include the monthly data when compared with the daily observation (IMPROVE data).

3. Section 4, Figure 2, 3, 4: Please show the difference between model and observation, especially with quantitative presentation (e.g. percentage differences).

4. P18, L398: The MODIS overestimation is compared to AERONET? Please quantify their differences.

5. P19, L403: Are there any references about the domination of sea-salt aerosol?

6. P19, Figure 7: Please explain the underestimate in in July and August over the West US.

7. P20, L442-444: Any explanations about this conclusion?

8. L444-447: I don't think so. I did not see that the model is able to reproduce the seasonal variation well. Again, please provide quantitative value using statistic method to convince the readers.

9. L468-L470, the seasonal variation of anthropogenic BC emission overs Asian is not such significant to make this big differences, which can be found in either INTEX-B or MEIC emission inventory.

10. L471-474: I don't understand the point.

11. Figure 11: there are big differences if the authors quantify them, especially under 1km. It is not such subjective to get this conclusions.

12. L496: Is it Figure 11?

13. Figure 12: Any references to show similar vertical distribution?

14. L502-507: Which is the major factor, the retrieval bias in observation data of CALIPSO or the emission uncertainty in the inventory? This conclusion looks like assumption without any strong evidence support.

15. Section 4.5: Please use the general model evaluation method or figure (Taylor diagram) to provide quantitatively values before getting the conclusions.

16. L570: It is better to compare the GFED3 with other biomass burning emission, e.g. FINN, or refer the published results to get this conclusion. Also, about the underesti-

mate of BC and OC, please see my comment 2 about including biomass burning into the model.