Comments to the Author:
Dear Authors,

I appreciate your efforts to conduct additional simulations and in-depth analyses to address the reviewers comments. I can see the manuscript is much improved with the revision.

The new reviewer report raises additional comments on the evaluations of model temperature profiles using the AIRS data, model design using the PBL schemes, and evaluation of profiles average of daytime and nighttime, etc. These comments are helpful for model diagnoses of aerosol and chemistry biases due to meteorology.

Please address these comments carefully.

Dear Editor,

Thank you for your consideration of the submitted manuscript. We have revised the manuscript in response to the reviewer's comments. The major changes are:

1. We have added the ERA-Interim reanalysis data to evaluate the vertical profiles of the model simulations.
2. Specific humidity profiles are added in the Supplementary Fig. 9.

The point-by-point response is listed below.

Sincerely,

Longtao Wu

Reviewer #4

While the other reviewers commented on the chemistry aspects, I will mostly focus on "the role of meteorology" and vertical mixing. Unfortunately, most of the discussion regarding vertical mixing is not robust (or highly questionable). Also I am surprised none of previous conclusions regarding PBL schemes (or vertical mixing treatments) in WRF are used to guide/help the investigation in this study.

Major comments

As one of the reviewers pointed out, AIRS profiles may not be appropriate to evaluate simulated profiles in the boundary layer. The fatal issue is that the AIRS profiles are not consistent with the surface observation. See LN438-439, the model overestimates surface temperature throughout the year comparing with surface observation. However, comparing with the AIRS profiles, the model only overestimates temperature near the surface in cold months while it underestimates temperature near the surface in warm months (Fig. 13). Such inconsistency may suggest AIRS profiles near the surface may not be reliable.

The vertical profiles from the ERA-Interim reanalysis dataset are added in the revised manuscript. Please see Fig. 13, Supplement Figs. 7, 8 and 9 in the revised manuscript for the comparison. Although there are some differences between AIRS and ERA-Interim, our conclusions are unchanged.

As shown in Figure 13, ACM2 predicts more stable boundary layer, how would more stable boundary layer leads to simulate lower surface NO3 and NH4 as discussed in LN473-477? Higher stability (particularly in cold season) should lead to more accumulation of pollutants near the surface.

The stability difference is quite small between ACM2 and YSU. We suspect the differences in surface NO3 and NH4 are not due to atmospheric stability changes. Figure 15 shows that more aerosols are transported above the surface in the 20km_P7 (ACM2) than in the 20km_D2 (YSU). The difference may be due to different parameterization methods of chemical transport in the PBL scheme. It is discussed in LN485-486.

The current evaluation of vertical profiles uses average of daytime and nighttime. The boundary layer structure is totally different during daytime and nighttime. I am not sure what the comparison of averaged profiles across daytime and nighttime really tell.

This study focuses on the seasonal variability of aerosols. Diurnal variability is beyond the scope of this study. The simulation of aerosol diurnal variability in California can be found in Fast et al. (2014).

Fast, J. D., Allan, J., Bahreini, R., Craven, J., Emmons, L., Ferrare, R., Hayes, P. L., Hodzic, A., Holloway, J., Hostetler, C., Jimenez, J. L., Jonsson, H., Liu, S., Liu, Y., Metcalf, A., Middlebrook, A., Nowak, J., Pekour, M., Perring, A., Russell, L., Sedlacek, A., Seinfeld, J., Setyan, A., Shilling, J., Shrivastava, M., Springston, S., Song, C., Subramanian, R., Taylor, J. W., Vinoj, V., Yang, Q., Zaveri, R. A., and Zhang, Q.: Modeling regional aerosol and aerosol precursor variability over California and its sensitivity to emissions and long-range transport during the 2010 CalNex and CARES campaigns, Atmos. Chem. Phys., 14, 10013-10060, doi:10.5194/acp-14-10013-2014, 2014.

Also the current study compares the different performance of YSU and ACM2. This might be a very poor choice for investigation of PBL schemes in WRF. Both YSU and ACM2 are nonlocal schemes and they have similar performance in most cases (comparing to the differences between local and nonlocal schemes). It is not clear to me how the ACM2 scheme performs better than the YSU scheme in this evaluation. Also both YSU and ACM2 schemes have different treatments for daytime and nighttime respectively, a more appropriate approach should be evaluating the daytime and nighttime performance separately.

Previous studies have shown that YSU and ACM2 have good performance in WRF and WRF-Chem simulations. For example, Hu et al. (2010) showed that "the YSU and ACM2 schemes give much less bias than with the MYJ scheme". Xie et al. (2012) concluded that "It is reasonable to infer that WRF, coupled with the ACM2 PBL physics option can be a viable producer of meteorological forcing to regional air quality modeling in the Pearl River Delta (PRD) Region". Cuchiara et al. (2014) showed that "the overall results did not indicate any

preferred PBL scheme for the Huston case. However, for ozone prediction the YSU scheme showed greatest agreements with observed values". Banks and Baldasano (2016) demonstrated that "the ACM2 scheme showed the lowest mean bias with respect to surface ozone at urban stations, while the YSU scheme preformed best with simulated nitrogen dioxide. The ACM2 and BouLac schemes performed better than the YSU scheme for air quality simulations." Banks et al. (2016) concluded that "non-local PBL schemes give the most agreeable solutions when compared with observations". Chen et al. (2017) showed that "as for the PM2.5 simulation, the combination of the YSU PBL, Goddard SW and GFDL LW schemes showed the greatest consistency with the observed values". It is clarified in the revised manuscript as "Previous studies showed that both the YSU and ACM2 schemes have good performance in simulating boundary layer properties (e.g., Hu et al., 2010; Xie et al., 2012; Cuchiara et al., 2014; Banks and Baldasano, 2016; Banks et al., 2016; Chen et al., 2017)."

The goal of this study is not evaluating which PBL scheme is better. The sensitivity experiment shown here is to demonstrate that the simulation of aerosols is sensitive to the PBL scheme. We have added some of the references above in the revised manuscript.

Reference:

Hu, X. M., J. W. Nielsen-Gammon, and F. Zhang (2010), Evaluation of three planetary boundary layer schemes in the WRF model, J. Appl. Meteorol. Climatol., 49(9), 1831–1844, doi:10.1175/2010JAMC2432.1.

Xie, B., J. C. H. Fung, A. Chan, and A. Lau (2012), Evaluation of nonlocal and local planetary boundary layer schemes in the WRF model, J. Geophys. Res., 117, D12103, doi:10.1029/2011JD017080.

Cuchiara, G.C., Li, X., Carvalho, J., & Rappenglück, B. (2014). Intercomparison of planetary boundary layer parameterization and its impacts on surface ozone concentration in the WRF/Chem model for a case study in Houston/Texas. Atmospheric Environment, 96,175–185. http://dx.doi.org/10.1016/j.atmosenv.2014.07.013

Banks, R.F., Baldasano, J.M., 2016. Impact of WRF model PBL schemes on air quality simulations over Catalonia, Spain. Science of the Total Environment, 572, 98-113, http://dx.doi.org/10.1016/j.scitotenv.2016.07.167

Banks, R. F., J. Tiana-Alsina, J. M. Baldasano, F. Rocadenbosch, A. Papayannis, S. Solomos, and C. G. Tzanis (2016), Sensitivity of boundary-layer variables to PBL schemes in the WRF model based on surface meteorological observations, lidar, and radiosondes during the HygrA-CD campaign, Atmos. Res., 176, 185–201.

Chen, D., X. Xie, Y. Zhou, J. Lang, T. Xu, N. Yang, Y. Zhao and X. Liu (2017) Performance Evaluation of the WRF-Chem Model with Different Physical Parameterization Schemes during an Extremely High PM2.5 Pollution Episode in Beijing. Aerosol and Air Quality Research, 17:262-277. doi: 10.4209/aaqr.2015.10.0610

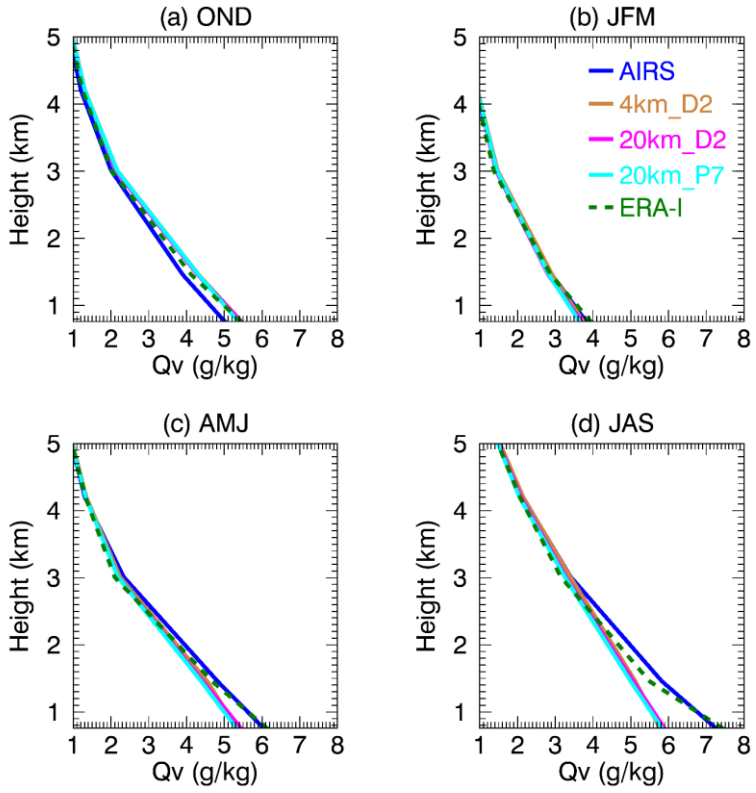
LN454-456, The discussion may be wrong. Fig. 13c indeed shows a neutral (or slightly stable) boundary layer below 3 km AGL. This does not mean the model cannot capture the well-mixed boundary layer. Actually in the convective boundary layer, the observed profile of potential temperature is indeed slightly stable [Deardorff, 1972], that is why some PBL schemes (e.g.,

YSU) added the countergradient term to make the simulated profiles in the convective boundary layer slightly stable [Frech and Mahrt, 1995]. Again, I am surprised none of the previous efforts in terms of PBL scheme evaluation is surveyed before the numerical experiments and during the writing of the manuscript.

Figure 12c shows that the 4km_D2 experiment doesn't capture the well-mixed boundary layer of aerosols observed by CALIOP. While the AIRS observation and ERA-interim data show conditionally unstable lower troposphere which favors upward displacement of surface aerosols, the simulated stable boundary layer would limit the uplifting of aerosols from the surface, contributing to the low biases in the aerosols in the boundary layer. We think the misrepresentation of boundary layer stability is one source of errors for the discrepancies in the simulated aerosol profiles. We have added discussions of previous studies on PBL schemes in LN 221-224.

Vertical profile of RH is not a good choice to evaluate different vertical mixing treatments. Instead, specific humidity should be used for PBL evaluation.

Specific humidity is included in the Supplementary Fig. 9 in the revised manuscript. All the simulations show dry biases near the surface in the warm season comparing to ERA-Interim. However, it cannot explain the low bias of dust above the surface (0.3 – 3 km) relative to CALIOP measurements.



Supplementary Figure 9. Vertical profile of seasonal mean specific humidity (g/kg) in the WRF-Chem simulations, AIRS and ERA-Interim. The 20km run (not shown) is similar to the 20km_D2 run while the 4km run (not shown) is similar to the 4km_D2 run.

In summary, the current design of numerical experiments in terms of role of vertical mixing treatments and analysis of the results are not adequate to diagnose the model errors associated with predicted chemical species.

Other specific comments:
LN438-439, warm and dry biases near the surface usually mean too strong vertical mixing rather than "not enough convective vertical mixing".

A warm bias near the surface promotes convective vertical mixing, but a dry bias prohibits it. Combining the effects of temperature and humidity, we compute the equivalent potential temperature (θ_e) (Fig. 13) that shows the convective instability of the atmospheric profiles. The discrepancy in θ_e between the observations and model simulations is quite clear. We think the neutral or slightly stable profiles in the model limit the uplifting of aerosols from the surface, contributing to the low biases of simulated aerosols in the boundary layer. We have clarified this in the revised manuscript. Text revised to clarify.

LN454, In Fig. 13c, I actually see unstable layer below 3km rather than "1.5km"

The ERA-Interim shows a neutral layer between 1.5 and 3 km.

LN457-458, The logic is confusing. "not enough convective vertical mixing" should result in high biases of simulated pollutants near the surface, rather than "low biases"

The simulated pollutants near the surface are biased high in the 4km_D2 experiment. The low biases are between 0.3 km and 3 km. It is clarified in the revised manuscript.

LN461-462, it is not clear, how vertical redistribution of pollutants by vertical mixing could change the column-integrated AOD.

Column-integrated AOD is the integral of aerosol extinction at each layer over the atmospheric column. As shown in the aerosol extinction profiles in Fig. 12, the simulations have low biases above the surface (0.3-3.0 km) compared to CALIOP. We think the simulated stable lower troposphere limits the vertical displacement of pollutant transport above the surface, which contributes to the low biases of aerosol extinction and AOD in the simulations. Text revised to clarify.

LN468-469, In terms of vertical gradient of equivalent potential temperature, I don't think the model did better in cold season than warm season. In cold season, particularly OND, the model significantly underestimates stability (Fig. 13a).

Comparing to ERA-Interim, the vertical gradient of equivalent potential temperature is well simulated in the cold season. In the warm season, both ERA-I and AIRS observed unstable environment while the model simulations produce neutral or stable lower troposphere.

LN478-480, I don't think you have proved vertical mixing of ACM2 is stronger than YSU. Fig. 13 actually shows the opposite.

The stability difference is quite small between ACM2 and YSU. We suspect the differences in surface NO3 and NH4 are not due to atmospheric stability changes. Figure 15 shows that more aerosols are transported above the surface in the 20km_P7 (ACM2) than in the 20km_D2 (YSU). The difference may be due to different parameterization methods of chemical transport in the PBL scheme. It is discussed in LN485-486.

LN484-486, how would "more conducive convective vertical transport in the PBL scheme" could lead to increase of aerosol in the boundary layer?

We mean "aerosol above the surface", as shown in Fig. 15 that more aerosols are uplifted above the surface in the 20km_P7 (ACM2) than in the 20km_D2 (YSU). It is clarified in the revised manuscript. Text revised to clarify.

LN492-493, why would "stable environment" lead to low biases of aerosol in the boundary layer and column-integrated AOD? "stable environment" should lead to accumulation of aerosol near the surface. Again, it is not clear how vertical re-distribution of aerosol would change column-integrated AOD.

You are right that stable environment leads to accumulation of aerosol immediately near the surface, resulting in low biases of aerosol above the surface between 0.3 and 3 km. The low bias in the boundary layer thus contributes to the low bias in the column-integrated AOD as the concentrations of aerosols above 3 km are very small. We have clarified this in the manuscript.

LN498-499, in Fig. 13, I don't think ACM2 performed better in terms of boundary layer structure than YSU.

Figure 15 shows more aerosols are transported above the surface in ACM2 than in YSU.

I saw three routine soundings in CA (http://weather.uwyo.edu/upperair/sounding.html), are you sure they are not in your domain?

The three sites are Edwards (34.90°N, 117.92°W), Vandenberg (34.75°N, 120.57°W), Oakland INT AP (37.75°N, 122.22°W). None of them are located in the region discussed in this paper (the SJV or the red dashed box in Fig. 1a).

References:

Deardorff, J. W. (1972), Theoretical expression for the counter gradient vertical heat flux, Journal of Geophysical Research, 77(30), 5900-5904.

Frech, M., and L. Mahrt (1995), A 2-Scale Mixing Formulation for the Atmospheric Boundary-Layer, Bound-Lay Meteorol, 73(1-2), 91-104.