

Response to Reviewer 1

Min Huang (mhuang1@engineering.uiowa.edu) on behalf of the authors

General comments

This study conducted a series of model experiments to explore the sensitivity of simulated background ozone over the western United States for one month period (June 5-July 15, 2008). I do recognize the authors' great effort to conduct lots of model experiments to address the uncertainties in estimating background ozone. However, the analysis is lack of focus and in-depth discussions. The authors need to carefully interpret the results. Some statements in the abstract are confusing as discussed below, and I do not see convincing analysis in the paper to support these conclusions.

Thanks for the comments by Reviewer 1. We have addressed all specific comments by the reviewer. In the revision, we have re-organized and heavily reworded the abstract, introduction and conclusions, and have expanded discussions for the concerns been brought up. Please see the point-by-point responses below (in blue). The original reviews are also included (in black).

Major comments:

1. In the abstract, the authors stated that "Forward sensitivity simulations show that TBG extensively affect Western US surface O₃, and can contribute to >50% of the total O₃, varying among different geographical regions and land types". This is a very big statement, how do you get the number ">50%"? Figure 4a shows that the maximum background occurs over the EPA Region 10 and the magnitude is no more than 15 ppbv, which is up to 20% of total ozone shown in Figure 2c.

Fig. 4 demonstrates the contributions to total NA background from three individual contributors: transported background (TBG) O₃ and several other precursors, biomass burning emissions and biogenic emissions. The contributions are shown for both MDA8 (4a-c) and W126 (4d-f). The purpose is to show that 1) TBG is the major contributor to background O₃, compared to the other two; 2) TBG contribution to total O₃ is large, which can be reflected from other figures as well (e.g., Fig. 5).

Fig. 4a and 4d clearly show the large magnitude in TBG contributions (based on extrapolation) and their spatial variability for both MDA8 and W126. We have explained the calculation of extrapolated TBG contributions in Fig. 4 caption: "To calculate (a) and (d), O₃ sensitivities to 75% reduction in BC O₃ and precursors were extrapolated to 100% perturbation and summed". The extrapolated TBG contributions for MDA8 (Fig. 4a) ranged from ~25 to 50 ppb, with the hotspots occurring in EPA Region 10. For W126 (Fig. 4d), the contributions ranged from ~5 to >20 ppm-h, with the hotspots shown in the Central Valley and Nevada.

In terms of relative contribution, we calculated the relative sensitivity (S₂, defined in Section 3.1.4) of surface O₃ to 50% perturbation in the boundary conditions. Fig. 5 shows that the relative sensitivities over Region 10 are the largest (i.e., over 30% and 70% for MDA8 and W126, respectively).

We have clarified and extended the explanation in Section 3.1 and replaced the fixed number in the abstract about the TBG contribution. Please also refer to our answers below.

In fact, the simulated total ozone at 60 km resolution is biased high by approximately 10-20 ppbv over the EPA region 10. Is this bias caused by excessive background influence in the model? A very important piece currently missing in the analysis is the influence of model bias on estimated background. Comparing Figure 8a and Figure 7a, the strongest sensitivity occurs on days when the model bias in total ozone is largest, e.g., June 18-22, June 24-25, and July 12-14. The model is clearly overestimating background influence on these days. How these biases affect the overall conclusion regarding the contribution of transported background to surface ozone should be clearly addressed.

It is true that the biases in the boundary conditions led to the ~10-20 ppb overprediction in total O₃ during the periods of June 18-22 and July 12-14. These have been discussed in detail in the text (Section 3.2.1). Figure 7a shows that in the 12 km/32L model grid, the model did not overpredict the June 24-25 period and has improved over the predictions in 60 km.

Lin et al. (2012) have correlated the model biases with stratospheric contributions estimated by the AM3 model on a ~50 km horizontal resolution, during the CalNex period. The r^2 value is around 0.4 for all points and decreases to ~0.2-0.3 for all points with positive biases. They concluded that “The correlation does not necessarily indicate a systematic model overestimate of stratospheric contribution to high-O₃ events as both under- and over-estimates occur.” Here we also added a quantitative exploration of the relationship between model biases and the estimated TBG contributions. The scatter plots (Fig. R1, upper panels) show the estimated TBG contributions to surface O₃ versus the model biases (modeled-observed) at surface sites, for MDA8 and W126, respectively. The correlations are very weak (i.e., $r=-0.186$ and -0.116 for MDA8 and W126, respectively). Our findings here are qualitatively similar as Lin et al. (2012).

We have further separated the estimation of TBG contributions by the model biases, using the thresholds of ± 10 ppb and ± 5 ppm-h for MDA8 and W126, respectively, and plotted the corresponding TBG contributions in histograms (Fig. R1, middle and lower panels). It is found that for MDA8, model-estimated TBG contributions at sites that have high (>10 ppb) and low (≤ 10 ppb) errors both ranged from 26-42 ppb, with the medians around 30-35 ppb. The Region 10 points (in green) have an additional 5-10 ppb of positive biases compared to the points with low biases, and the TBG contributions ranged from ~35-42 ppb, which indicate that the additional TBG contributed to the additional biases in total O₃. The TBG addition to total O₃ over Region 10 is close to linear, due to relatively slow local O₃ photochemistry (more NO_x-limited, with more clouds and lower temperatures). The histogram of TBG contribution to W126 at sites with low model biases (≤ 5 ppm-h) shows three peaks at ~5-6, ~13-15 and 18-20 ppm-h, while the one for TBG contribution to W126 at sites with high model biases (>5 ppm-h) has a narrower span, and the first peak (for Region 10 points in green) shifts to ~10-15 ppm-h. Again, the 5-10 ppm-h of shift indicates that the overestimation in TBG contributed to the higher biases in total O₃ over Region 10.

Therefore, it is better to revise the TBG contributions to MDA8 and W126 to ~30-35 ppb for everywhere and ~5 ppm-h for Region 10/~15-20 ppm-h for Region 9, respectively. Note that the model did a better job predicting O₃ at CASTNET sites than at the AQS sites (from the Figs 2, R1 as well as the statistics in Table 4a and 4b), and the TBG contributions to O₃ at these sites mainly fall into these ranges for MDA8 and W126.

This additional discussion is based on the analysis of the biases only at AQS/CASTNET sites that had >75% of the observations during the study period. In other words, it depends on the spatial distributions of these sites—most of which are located in California.

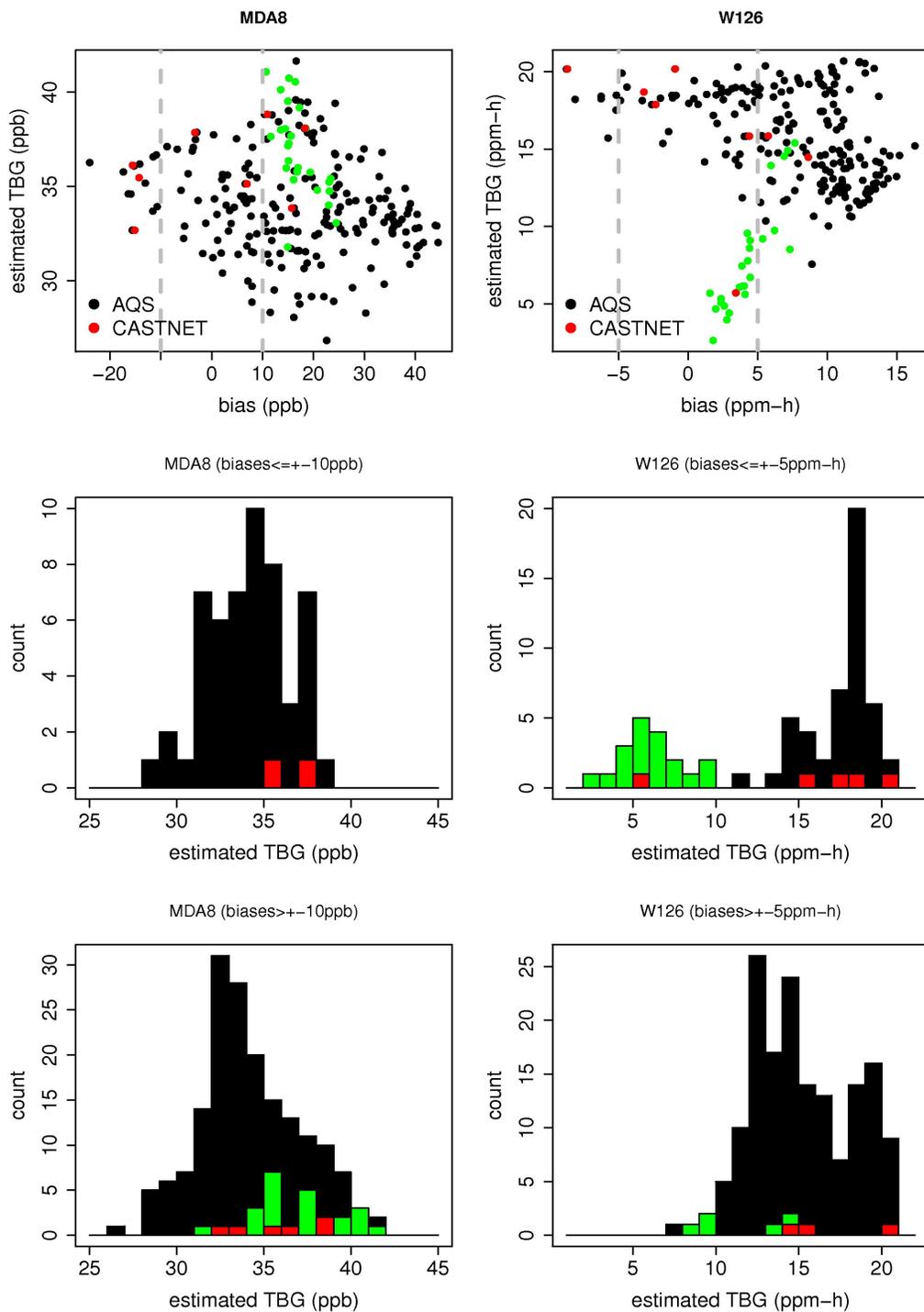


Fig. R1 Analysis of relationships between model biases and the extrapolated TBG contributions (O_3 +several precursors), based on the results from the 60 km/18L model grid and AQS/CASTNET observations. Black and red represent AQS and CASTNET sites, respectively. Green represents AQS sites over EPA Region 10 only (states of Washington, Oregon and Idaho).

The scatterplots and histogram-based analysis above give a statistically based estimation of TBG contributions to surface O₃ at observation sites. However, to completely understand the relationship between biases of predicted total O₃ and the biases introduced from boundary conditions is challenging, especially the spatial variability. We have taken great efforts to try to understand the sources of biases in the estimated TBG contributions. In this paper, these mainly include: 1) better understanding the biases from the boundary conditions (RAQMS) and 2) the quality of other model configurations (explained in the following two paragraphs).

For better understanding the biases derived from the RAQMS boundary conditions, we conducted the surface O₃ sensitivities to different extents of perturbations in boundary conditions (25%, 50% and 75%, shown in Fig. 3e-f). They demonstrate the non-linear surface O₃ sensitivity to boundary conditions and the strong extra-regional contributions. These also indicate that the relationships of surface O₃ biases and the biases in boundary conditions are complicated, and are different for MDA8 and W126.

For better understanding the biases caused by other model configurations, we show the surface O₃ sensitivity to boundary condition perturbations (by 50%) using different grid resolutions (Section 3.1.5) and original/reduced anthropogenic inventories (Section 3.1.4).

2. In the abstract, the authors stated that “The stratospheric O₃ impacts are weak”. This statement is vague. Do you actually implement a tracer to quantify the influence of stratospheric ozone? It is better to just state what you have done, i.e., the sensitivity of simulated ozone to the top boundary conditions.

We agree with this suggestion. Yes, the sensitivity itself is not sufficient to completely quantify the contribution from the stratosphere, although it is a good indicator. The top boundary conditions (and the associated biases) and the location of model top (shown below) both impact the resultant sensitivities. We have reworded the sentences in the abstract and the corresponding text in Section 3.1.2 to clarify our method and findings.

What is the time frequency of the top boundary conditions?

We have revised the manuscript (Section 2.3.1) to indicate the 6-h original time step from RAQMS and that it was linearly interpolated into 1h frequency for input to the STEM simulations.

What is the model top?

In Section 2, we have revised the text to specify “~170-210 hPa in a terrain-following sigma-z system”, corresponding to ~10-12 km AGL, with spatial and temporal variability. Fig. R2 shows the monthly mean O₃ top boundary conditions (~80-220 ppb) with the WRF pressure overlaid.

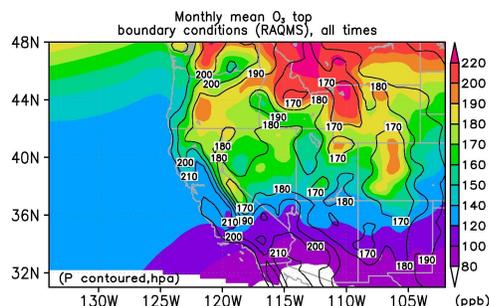


Fig. R2 Monthly mean O₃ top boundary conditions from RAQMS with the WRF pressure overlaid

Is the model vertical resolution fine enough to resolve the tropopause? This important information related to STE is currently missing in the manuscript. Figure 7 shows that the model fails to reproduce the layered structure of observed ozone at the Trinidad sonde site, which raises the concern about the credibility of the model in attributing the influence of stratospheric ozone.

The model top boundary and resolution of tropopause dynamics certainly matters. The model vertical resolution is relatively coarse in the upper troposphere/lower stratosphere (UT/LS) compared to the lower troposphere (LT), ranging from several hundred meters to >1 km (Fig. 7c-e). This is a coarser resolution than the middle atmosphere models designed to resolve strat-trop exchange, but considerably higher than in many prior global model studies of background O₃ over the western U.S. (e.g., Fiore et al., 2003; Zhang et al., 2011). However, by using a lower model top and model top boundary conditions, we avoid the pitfalls of directly simulating tropopause dynamics in a regional model (Büker et al., 2005), and focus instead on tracing the effects of the stratospheric O₃ from RAQMS assimilation in the UT and simulating the tropospheric dynamics that WRF is designed to resolve.

In this case, the difficulty in resolving the layering in the coastal profile at THD at UT is mainly due to the boundary condition biases rather than the model vertical structure. For coastal sites like THD, even if the model top O₃ biases in the UT are large, these don't affect the THD surface concentration, which is driven primarily by transport from the western boundary. As Fig. R3 demonstrates, the O₃ sensitivities at THD to changes in the top and UT lateral boundary conditions show highest values in mid/upper troposphere (>3 km). Based on analysis of the ozonesondes during CalNex period, Cooper et al. (2011) also concluded that "Polluted air masses and stratospheric intrusions that descend isentropically along the west coast likely explain the O₃ enhancement above the coast that descends from the mid-troposphere in the north to the lower troposphere in the south", which indicated that unlike on southern California, the stratospheric impact on THD barely affected LT.

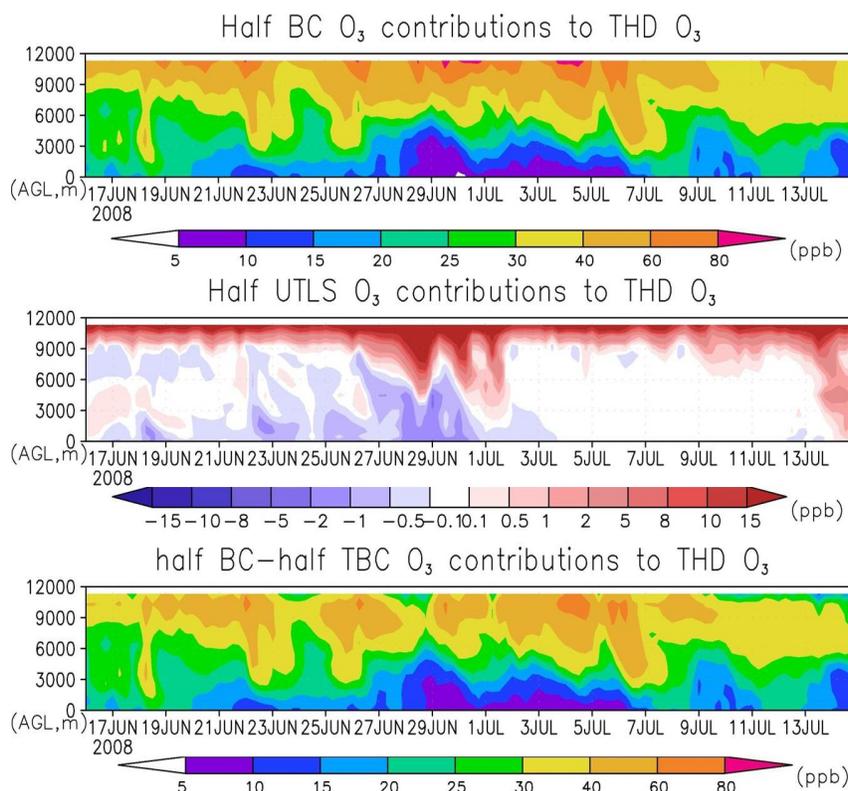


Fig. R3 Time-height curtain plots of O₃ sensitivities at THD in response to (upper) 50% reduction in model top+lateral boundary conditions; (middle) 50% reduction in model top boundary conditions; and 50% reduction in model lateral boundary conditions (calculated by subtracting upper and middle panels).

3. Abstract, Lines 19-20: “The probabilities of air masses originating from MBO (2.7 km) and THD (2.5 km) entraining into the boundary layer reach daily maxima of 66% and 34% at 3:00 p.m. PDT, respectively, and stay above 50% during 9:00 a.m. – 4:00 p.m. for those originating from SC (1.5 km)”. Using the different time frame for MBO/THD and SC is confusing. I think you want to compare the lower free troposphere (MBO and THD) versus the surface site (SC).

We calculated the impacting probability (IP, equation 11) spanning the entire daytime period in Table 5. The definition of IP and the detailed interpretations of results in the table and abstract are in Section 3.2.3. The different times are concluded from the results for the three locations. It was not that we intended to pick up different times for comparing the sites.

4. Abstract, last paragraph and section 3.3.2: This seems to disconnect with the focus of this paper regarding transported background ozone. Better if you can phase the discussion in terms of how the assimilation improves estimated background.

The main objective of this paper is to discuss the impacts of TBG (indicator of extra-regional contributions) on western US air quality during summertime. The transported plumes descend from free troposphere (FT) to impact the surface air quality. From a modeling perspective, the modeled surface O_3 is sensitive to boundary conditions, and modeled eastern Pacific FT O_3 later affects the modeled surface O_3 distributions over the western US. Therefore, demonstrating the linkages in the model between the surface O_3 and FT O_3 in the eastern Pacific is necessary, and the observed O_3 vertical profiles could be helpful for improving the simulated surface O_3 .

Data assimilation and adjoint sensitivity in this paper are included to help quantify how the surface O_3 is linked to the eastern Pacific O_3 in the FT. E.g., highest adjoint sensitivities are shown in the mid-troposphere in the eastern Pacific (Fig. 12) at earlier times. And when assimilating the available surface observations, the O_3 distributions changed over three dimensions (Fig.s 14-15). The usefulness of the vertical profiles for improving the surface O_3 predictions was also examined. However, the available vertical profiles are too sparse for significantly improving predicted O_3 , and our results suggest that improving the retrieval quality and increasing the density of observed vertical profiles will benefit future data assimilation activities and provide a means to reduce the uncertainties in surface O_3 and the estimates of TBG.

It is challenging to use the data assimilation to improve the background estimates, because the improvements in O_3 concentrations are for the total amounts, not a specific portion from a single contributor. However, the messages delivered from Fig.s 14 and 15 include: 1) By using data assimilation of surface observations, O_3 in FT over the eastern Pacific increased by up to 4-6 ppb, indicating the uncertainties in the boundary conditions; 2) These “optimized” fields over four dimensions can be used to estimate the uncertainties in the satellite measurements in the FT, and is consistent with previous literature (e.g., TES up to 20%).

For this specific case, a rough estimation of the improvement on TBG versus local emissions contributions through data assimilation is: Case AS-a priori along the TES orbit shows +1-6 ppb (~3 ppb) of differences at 40-42N (Fig. 14a), while at surface +5-15 ppb (~10 ppb) over northern CA and negative 1-3 ppb (-2 ppb) over Region 10. If we assume the addition of changes in offshore FT to surface O_3 is close to linear for everywhere, then we could conclude that by using the data assimilation method, the improvement of TBG contributions to surface O_3 is ~1-6 ppb, and the improvement of local emission effects for Regions 9 and 10 are ~7 ppb and ~-5 ppb, respectively.

Possible alternative ways (using data assimilation methods) to improve the estimation of TBG/extra-regional contributions include: 1) To improve the chemical fields in global models that are used as regional model boundary conditions; and 2) To improve the emissions (e.g., NO_x) in extra-regions, and then use the improved emissions to calculate the contributions from extra regions through methods such as perturbations. These are beyond the scope of this study in which a regional model is a focus, but they could be future work.

5. Different emission data (for both anthropogenic and biomass burning sources) are used for the model simulations at 60 km and 12 km horizontal resolutions. This model setup precludes the capability to isolate the influence of emissions vs. transport processes on simulated ozone. Please quantify the percentage difference for total NO_x and VOCs over the common domain in the two emission inventories.

We thank the reviewer for this question, as answering it has helped improve the article and highlight the surprisingly minimal impacts of local emissions rates on conclusions for net sensitivities to the background transport.

Fig.s R5 and R6 compared the anthropogenic and biomass burning total (sums of surface and elevated) emissions from the 60 km and 12 km model configurations, shown on a log10 scale for the emissions. For anthropogenic emissions, we show NO_x, NMVOCs and their ratios; For biomass burning emissions, we focus on CO only here, since NO_x and VOCs emissions are scaled to CO emissions in both resolutions in a similar way. Table R1 compares the differences of emissions in shared domains between the two resolutions (i.e., ratios of emissions in 60 km/12 km).

Table R1 60 km/12 km total emissions over the entire shared domain

Anthropogenic NO _x	Anthropogenic NMVOCs	Biomass burning CO
1.53	0.93	3.37

The choice of biomass burning emissions for the two resolutions considered the utility of original emission data for regional and continental-to-hemispheric modeling: Emissions at 60 km came from RAQMS at 2°×2° resolution and are not appropriate for the 12 km simulations, while the generation of 12 km emissions require local information that are not consistent in a 3600 km² (60 km) grid cell for application. Anthropogenic emissions across the two resolutions differ only over California.

In terms of the impacts of using different “local” emissions on surface O₃ sensitivity to TBG/boundary conditions, Section 3.1.4 compares the surface O₃ sensitivities to boundary conditions perturbations under original and scaled emissions in 60 km, in which the scaled amounts were closer to those used in the 12 km simulations. We found that the sensitivities are much less different compared to the impacts of model grid resolution (Section 3.1.5). Therefore, the differences of local emissions do not affect main conclusions of this study.

In page 15250, the authors attribute the ozone bias at the Trinidad Head sonde during June 28-29 in the 60 km simulation to the uncertainties in biomass burning emissions. In Figure 4b, there is a clear hotspot of biomass burning influence over Northern California. Is this hotspot still present if using BB emissions from the 12-km simulation?

Fig. 4b shows the general fire impacts for the entire study period, while for the temporal variability of the O₃ sensitivities to fires at THD, please refer to Fig. 8b.

From the ozonesondes (Fig. 7c) during June 28-29, it is shown that ~300-500 m was clean while O_3 at ~0.5-3 km ASL ranged from 60-80 ppb. THD is indirectly affected by fire emissions during this period. The NOAA smoke prediction and the GASP AOD demonstrated the fire locations in northern California (Fig. R4) and their impacts on THD during ~June 29-30. Depending on how far original the emission sources in different biomass burning emission inventories are from THD, as well as the amount of emissions, the extents of fire impact on THD could be quite different.

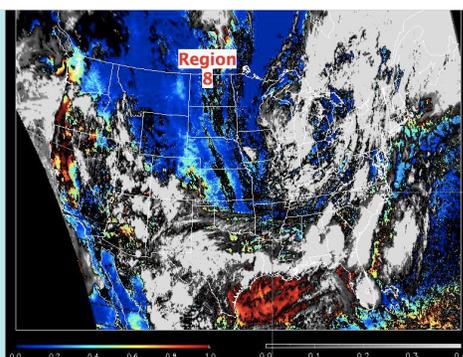
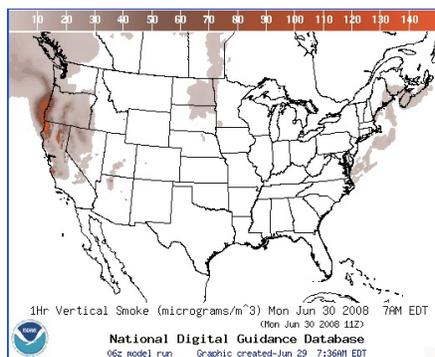


Fig. R4 NOAA smoke prediction (left) and GASP AOD (right) for ~June 29-30, 2008 (figure sources: smog blog, http://alg.umbc.edu/u/asaq/archives/2008_06.html).

The biomass burning emissions in the two model grids are shown in Fig. R6 (upper panels) for June 29, 00 UTC. Both emission inventories indicate the fire locations in northern CA with similar magnitudes. In 60 km, due to the coarse resolution, the fire locations/emissions covered THD directly and may have caused the overprediction. In 12 km, the emissions indicate fire occurrence there, and the impacts on THD O_3 may be underestimated due to various factors such as the meteorological fields and the amounts and diurnal viabilities of the emissions, etc.

6. Figure 4, please clarify whether the w126 weighting function is applied before or after the subtraction of two sensitivity simulations. I believe the results are likely very different. How does the calculation method affect the attribution?

The sensitivity calculation for W126 in this paper consistently followed this method: Assuming we have O_3 time series from model simulations 1 and 2:

- a) Calculate W126 based on O_3 from simulation 1, W126_S1
- b) Calculate W126 based on O_3 from simulation 2, W126_S2
- c) Calculate W126 sensitivity: W126_S1 - W126_S2

This method is more practical than the calculation that applies weighing functions to the sensitivity of O_3 time series themselves, because W126_S1 and W126_S2 are based on the simulation cases that represent possible scenarios.

7. Figure 1: Is this the only flight available during the study period? Why not show the overall statistics (observed vs. modeled mean and standard deviation for each 1 km altitude bin) using data from all flights available?

There were several flights during the ARCTAS-CARB field campaign period (June 18-24, 2008), which covered expanded areas in California. The comparisons between model and aircraft observations are for evaluating the model boundary conditions in the context (described in text in Section 2.3.1), so we selected the “boundary flight” on June 22 and used the samples over the eastern Pacific on that flight, since the O_3 vertical structures are significantly different over offshore and southern California urban regions (e.g., LA) and the Central Valley where most of the other samples were taken. The evaluations between model and other available aircraft observations have been described in detail in previous papers (Huang et al., 2010, 2011) to address a range of science

questions. We now briefly added in the discussion how they differ from the single flight shown in Fig. 1.

Specific comments:

Page 15236, Line 20: what is the resolution (thickness) near the transition layers, e.g. near the PBL top and near the tropopause? I think this can explain why less background influence is estimated if the model include more vertical layers.

We have descriptions in Section 2.3 about the grid configurations: “The 18 layer grid had ~7 layers below 1 km and ~10–11 layers below 4 km, and the 32 layer settings had ~11 layers below 1 km and ~20–21 layers below 4 km.” We have added some descriptions in this section to describe the upper troposphere. The vertical resolution can also be read from Fig. 7c-e.

Page 15237: Line 12: Time-varying BCs at what frequency?

The original boundary conditions from RAQMS had a 6-h temporal resolution, and are interpolated to 1-h interval to drive the STEM model.

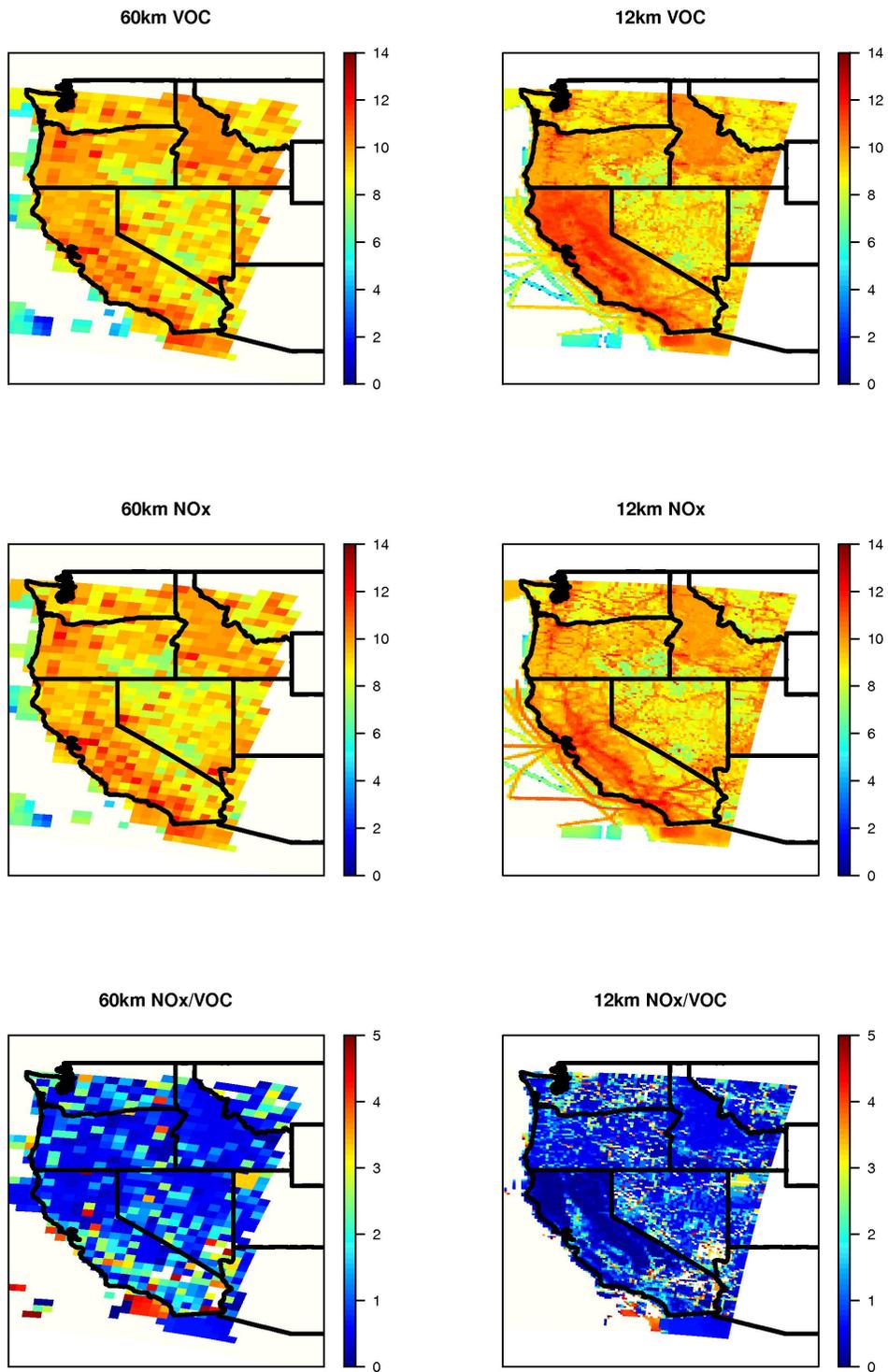


Fig. R5 Mean 60 km and 12 km anthropogenic NO_x and NMVOCs emissions (on a log₁₀ scale, molec./cm²/s) and the NO_x/NMVOC ratios (dimensionless)

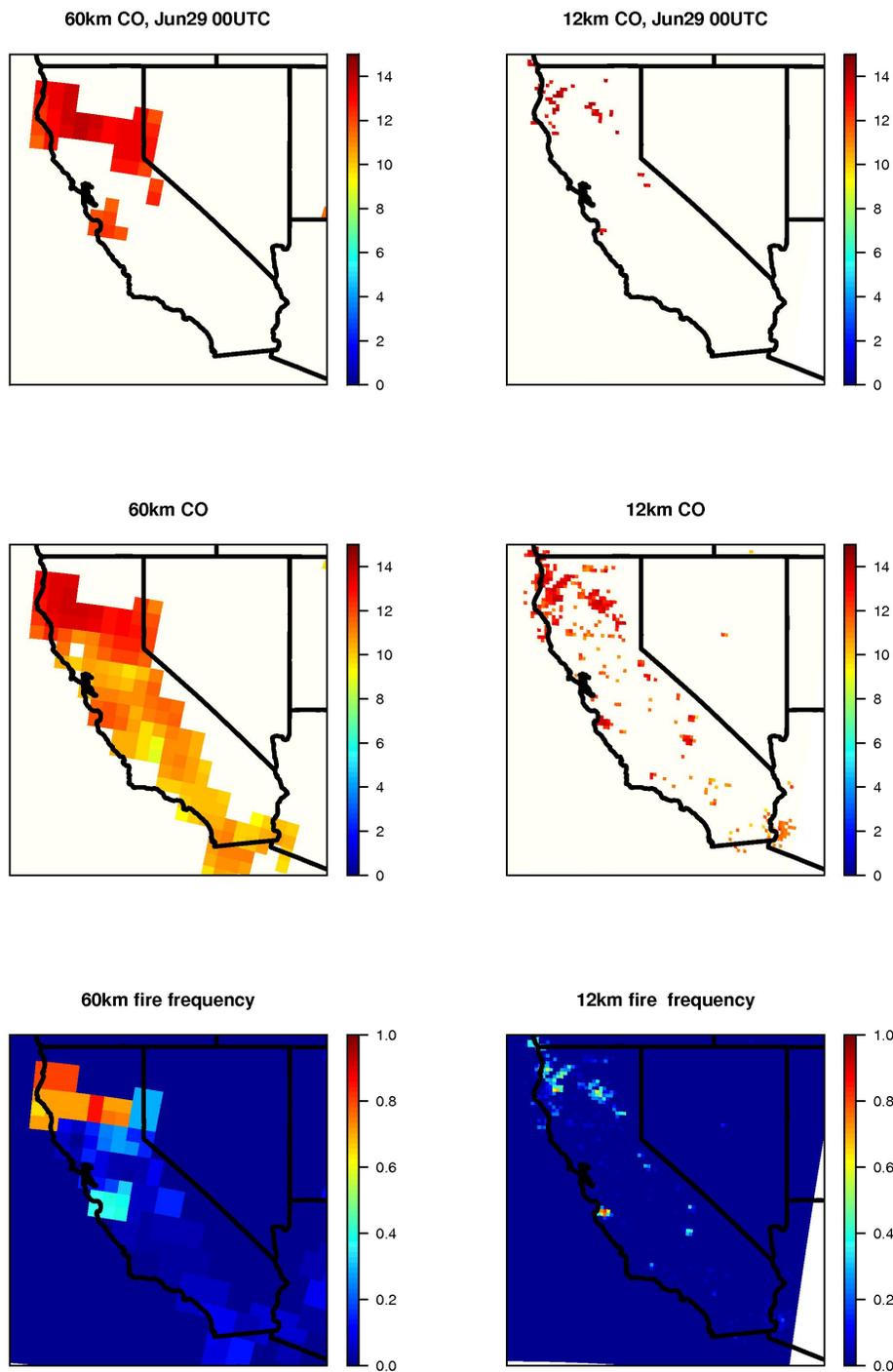


Fig. R6 June 29, 00UTC and monthly mean 60 km and 12 km CO fire emissions (on a log10 scale, molec./cm²/s) and frequency (dimensionless). The fire frequency for each grid is defined as the number of times that fire emissions>0/all times through the study period.

References

- Büker, M. L., M. H. Hitchman, G. J. Tripoli, R. B. Pierce, E. V. Browell, and M. A. Avery (2005), Resolution dependence of cross-tropopause ozone transport over east Asia, *Journal of Geophysical Research*, 110, D03107, doi: 10.1029/2004JD004739.
- Cooper, O. R., et al. (2011), Measurement of western U.S. baseline ozone from the surface to the tropopause and assessment of downwind impact regions, *Journal of Geophysical Research*, 116, D00V03, doi: 10.1029/2011JD016095.
- Huang, M., G. R. Carmichael, B. Adhikary, S. N. Spak, S. Kulkarni, Y. F. Cheng, C. Wei, Y. Tang, D. D. Parrish, S. J. Oltmans, A. D'Allura, A. Kaduwela, C. Cai, A. J. Weinheimer, M. Wong, R. B. Pierce, J. A. Al-Saadi, D. G. Streets and Q. Zhang, Impacts of transported back-ground ozone on California air quality during the ARCTAS-CARB period - a multi-scale modeling study, *Atmospheric Chemistry and Physics*, 10, 6947-6968, doi: 10.5194/acp-10-6947-2010, 2010.
- Huang, M., G. R. Carmichael, S. N. Spak, B. Adhikary, S. Kulkarni, Y. F. Cheng, C. Wei, Y. Tang, A. D'Allura, P. O. Wennberg, G. L. Huey, J. E. Dibb, J. L. Jimenez, M. J. Cubison, A. J. Weinheimer, A. Kaduwela, C. Cai, M. Wong, R. B. Pierce, J. A. Al-Saadi, D. G. Streets and Q. Zhang, Multi-scale modeling study of the source contributions to near-surface ozone and sulfur oxides levels over California during the ARCTAS-CARB period, *Atmospheric Chemistry and Physics*, 11, 3173-3194, doi: 10.5194/acp-11-3173-2011, 2011.
- Lin, M., A. M. Fiore, O. R. Cooper, L. W. Horowitz, A. O. Langford, Hiram Levy II, B. J. Johnson, V. Naik, S. J. Oltmans, C. Senff (2012), Springtime high surface ozone events over the western United States: Quantifying the role of stratospheric intrusions, submitted to *Journal of Geophysical Research*, May 21, 2012.
- Zhang, L., Jacob, D. J., Downey, N. V., Wood, D. A., D. Blewitt, Carouge, C. C., van Donkelaar, A., Jones, D. B. A., Murray, L. T., and Wang, Y.: Improved estimate of the policy relevant background ozone in the United States using the GEOS-Chem global model with $1/2^\circ \times 2/3^\circ$ horizontal resolution over North America, *Atmospheric Environment*, 45, 6769-6776, doi: 10.1016/j.atmosenv.2011.07.054, 2011.