Response to Referee Comment 2 by Anonymous Referee #3

Review of the paper Modeling Trans-Pacific Transport using Hemispheric CMAQ during April 2010: Part 1. Model Evaluation and Air Mass Characterization for the Estimation of Stratospheric Intrusion on Tropospheric Ozone by Syuichi Itahashi and co-authors. The paper is the first in a series of two investigating CMAQ simulations of trans-pacific pollution transport for one month (April) in 2010. In contrast to what the main title promises the paper deals exclusively with ozone transport from the stratosphere and a method for a quantification of the stratospheric contribution to tropospheric ozone. The approach followed here to answer this question is quite interesting and generally well described. However, the study suffers from the extremely short period under investigation. The authors justify why they selected this period for their study. They claim that it was published earlier that trans-Pacific transport played an important role during this period (Uno et al., 2011, Lin et al., 2012a). However, they try to draw more general conclusions about the contribution of stratospheric ozone to the concentrations in the troposphere and the model performance. I believe these findings on biases and model skill are not well justified because the data set used for this type of evaluation is too small. I suggest that the authors investigate a longer simulation period in order to derive statistical parameters for the model performance and the stratospheric contribution to tropospheric O3. They can then still investigate April 2010 as a special case in more detail like it is done now.

Reply:

We thank the reviewer for providing helpful and constructive comments. We have revised our manuscript according to the reviewer's comments and suggestions. We believe that these revisions address all points raised by the reviewer. Our point-by-point responses are provided below, and revisions are indicated in blue in the revised manuscript.

First, we revised the paper to fully describe why only a one-month simulation of April 2010 was conducted in this study. This is partly reinforced by other studies, but the selection was based on our analysis of monthly variation of MD8O3 in 2010. Second, we also carefully revised our manuscript to avoid generating the conclusions drawn from the results of this limited period.

Specific major comments:

Page (P) 1, line (l) 21: You describe the bias in RH given by H-CMAQ, however RH should be a quantity given by the driving meteorology model, which in this case is WRF.

Reply:

We agree with the reviewer and removed this sentence in the revised paper.

P 1, l 28-32: These statements are based on an investigation for April 2010 but the reader gets the impression that they have a more general validity. You should extend your evaluation period for this type of conclusions.

Reply:

To remind the readers and not to overstate, we have added "during April 2010," in this sentence.

P 2, l 17: Is it still true that the emissions in East Asia increase? There are more recent publications reporting the contrary.

Reply:

The sentence was intended to convey that in recent years emissions across Asia are changing dramatically. To address the reviewer's comment, we replaced 'the recent acceleration' by 'the dramatic variation'.

P 3, l 1/2: How can a publication from 1999 say something on real trends until 2010? You need to point out that this was a model study looking into the future. In addition, it would be interesting to know if the predictions for the development of Asian emissions were correct.

Reply:

We revised to explicitly state that this estimation is based on the model simulation. The revised sentence in Section 1 is as follows:

"The global model simulation assuming the tripling of Asian anthropogenic emissions from 1985 to 2010 indicated an increase in O_3 mixing ratios by 2-6 ppbv in the western

U.S.A. and by 1-3 ppb in the eastern U.S.A. on a monthly-mean basis, with the maximum effect occurring in April-June; this increase was suggested to more than offset the benefits of 25% domestic reduction in the western U.S.A. (Jacob et al., 1999)."

Based on the EDGAR emission inventory, we checked this assumption and found that this was reasonable. This info was added in Section 1 as follows:

"Based on the Emission Database for Global Atmospheric Research (EDGAR) version 4.3.1, anthropogenic emissions of NO_x and VOCs in China are estimated to have increased by 3.2 and 2.1 times during 1985-2010, respectively (Crippa et al., 2016), which is generally consistent with the assumption by Jacob et al. (1999)."

P 3, l 8/9: When Lin et al. modelled May/June 2010 and you do April 2010: Why don't you extend your model period and put the results in perspective to their results?

Reply:

The reason to focus on April 2010 is based on the analysis of monthly mean and percentiles behavior of observed MD8O3 during 2010. To address the reviewer's comment, we have added this analysis summarized in Figure S2 in the supplemental material along with our justification (see below) to support our selection of this month for model simulation.

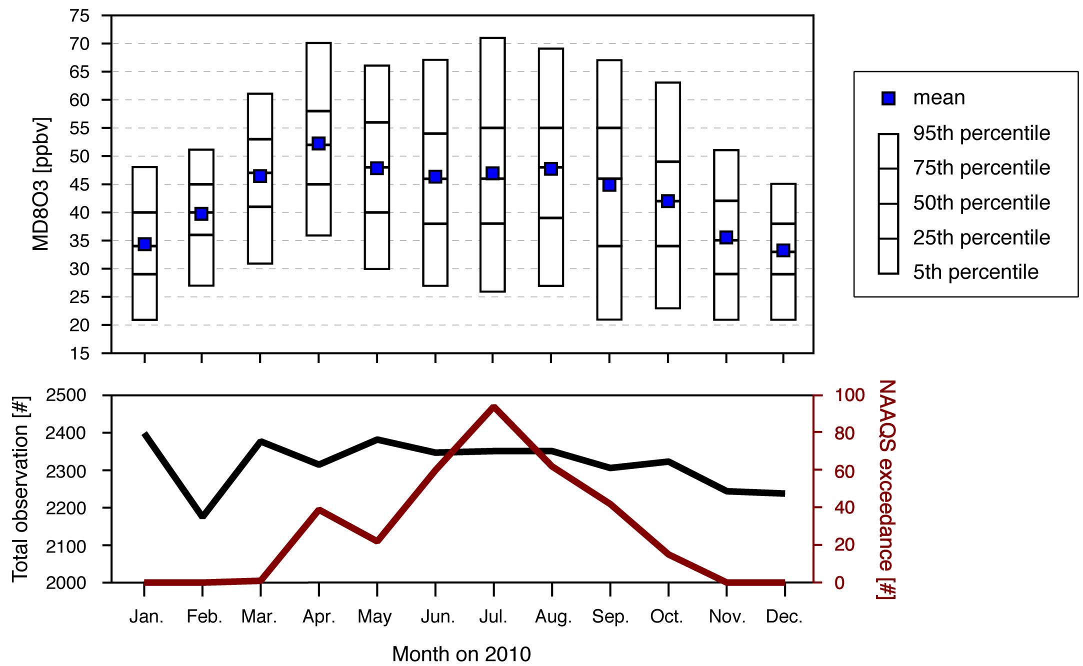

Figure S2. (Top) Monthly mean and percentiles of MD8O3 on 2010. (Bottom) Number of total observations (black color, left-axis) and exceedance of NAAQS (dark red color, right-axis; 75 ppbv is used as a criterion as 2010) on 2010.

The additional sentence in Section 1 is as follows:

"The variation in monthly mean and percentile distribution of observed MD8O3 during 2010 are shown in Fig. S2 in the supplemental material. Although high MD8O3 concentration for the 95th percentiles and the number of NAAQS exceedances were found during summer time, it is also apparent that mean MD8O3 during April 2010 was higher than any other month. Lower MD8O3 concentrations for the 5th and 25th percentiles were also noted as comparatively high during April 2010, indicating widespread enhancement of low-level O_3 further suggesting the possible impacts of trans-Pacific transport on O_3 levels across the U.S.A. during this month."

P 3, l 19-21: Because this is the case you need to cover other seasons with your model in order to evaluate it properly.

Reply:

This sentence conveys the general information, and we think the proper evaluation of stratospheric impacts is difficult even with a longer model simulation.

P 3, l 29/30: This objective is not covered in this paper at all. You should say more clearly what the objective of this paper is.

Reply:

To clarify the specific objectives of this Part 1 manuscript, we have revised the discussion in Section 1 as follows:

"The objective of this study is to better understand the relative contributions of precursor emissions from East Asia and the U.S.A. because the trans-Pacific transport has been recognized as an important factor. Previous studies primarily focused on Asian impacts on the western U.S.A., while this study investigates impacts across the entire U.S.A. In addition, some stratospheric intrusion events have been reported during spring 2010 (Lin

et al., 2012b), therefore this period is suitable to examine not only trans-Pacific transport but also stratospheric intrusion, both processes may contribute to the observed high O_3 episodes in the U.S.A. Examination of the impacts of both processes will shed light on the formation mechanisms underlying such high O_3 episodes, thus improving our understanding of their relative importance in leading to these high O_3 episodes. The results of this work will be presented in two parts. Part 1 paper focuses on characterizing the influence of stratosphere-troposphere transport on O_3 distribution in the lower to middle troposphere. A sequential Part 2 manuscript focuses on the contributions of emissions leading to higher O_3 mixing ratio through Trans-Pacific transport."

P 5, l 1 / 2: You state that O3 in the stratosphere is parameterized based on PV from WRF and an O3-PV function from Xing et al. (2016). Could you say a few words about how accurate this parameterization is?

Reply:

By introducing this O_3/PV parameterization, Xing et al. (2016) demonstrated that the parameterization improved O_3 model performance in the UTLS both in terms of magnitude and seasonality. The revised explanation is as follows:

"To account for the seasonal, latitudinal, and altitude dependencies in the O_3-PV relationship, a dynamic O_3/PV function was developed to consider latitude, altitude, and time based on 21-year ozonesonde records from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC) and corresponding PV values from WRF-CMAQ simulations across the northern hemisphere from 1990 to 2010 and is used in H-CMAQ (Xing et al., 2016). This parameterization of O_3/PV is constructed at three vertical levels of 58, 76, and 95 hPa fitted as a 5th order polynomial function, and applicable in the range of 50 to 100 hPa. Based on this new parameterization, it was demonstrated that UTLS O_3 agreed much better with observation in terms of its magnitude and seasonality (Xing et al., 2016). Mathur et al. (2017) further demonstrated improvements in representation of seasonal variations in surface O_3 using the parameterization"

P 5, l 17-19: Why did you simulate only such a short period? Is it computationally expensive to run the model? Which boundary conditions of those reported in the Hogrefe et al. (2018) paper did you use?

Reply:

Please refer our reply to your comment on P3, l 8/9 about the reason for the selected simulation period.

Regarding the computational burden, the part 2 paper uses the higher-order decoupled direct method (HDDM) to calculate the sensitivities. Although the HDDM is a sophisticated method to accurately derive sensitivities, the computational burden is much large. This is another reason to limit one-month simulation.

We added the description of boundary conditions as follows:

"The boundary conditions of H-CMAQ are taken from the clean tropospheric background values with updates to the physical and chemical sinks for organic nitrate species (Mathur et al., 2017)."

P 8, l 13-30: Again, given these significant deviations between model results and observations, it would be beneficial for your interpretation if you extend the modelled time period.

Reply:

We agree that a longer simulations period would enable more robust conclusions applicable over broader periods. However, as we mentioned above, this study focused on April 2010 when the increase of widespread MD8O3 was observed. For this one-month simulation, we prepared a suite of available data for model evaluations, including surface, vertical profile, and satellite observations.

P 10, l 8/10: Doesn't this suggest that the scaling approach is not accurate enough for modelling ozone concentrations in the upper troposphere. So isn't there a need for adding a model component that covers stratospheric ozone with its entire chemistry and dynamics?

Reply:

As we have already stated, the extension of model top layer beyond 50 hPa may be needed. We have revised this sentence to explicitly convey this meaning as follows:

“Using a finer vertical resolution for the upper layers and extending the model top beyond 50 hPa to cover larger portions of the stratosphere could be potential strategies to address this need.”

P 10, l 19-21: What could be the reason for this positive bias if it occurs despite nudging of RH from reanalysis data?

Reply:

Such positive bias was also found in AQMEII project despite the nudging on reanalysis data. Because RH is a diagnostic quantity which is dependent on a number of prognostic variables and sensitive parameters. Within our best knowledge, we cannot determine the critical reason to this.

P11, l 10-20: The comparison of the ozone profiles to the model values (Table 3) suggest that the model gives too low O3 concentrations, in particular in the free troposphere. You also state this in l 9/10 on P 11. However, in Figure 7 we see a mostly positive bias with too high modeled column values, in particular over the continents, where the O3 soundings were performed. Could you explain this? Does it tell me that satellite observations deviate quite much from ozone soundings? And does it mean that your findings whether the model is too high or too low depends on the observations you compare it with?

Reply:

Regarding satellite observation, Ziemke et al. (2006) reported scattered correspondence of satellite derived column ozone to ozonesonde dataset, with slight positive bias on satellite data. We have added this statement as follows:

“In addition, the model underestimation especially in the free-troposphere is noted through comparison with ozonesonde measurements (Table 3); however, this comparison showed model overestimation. The evaluation of satellite data compared to ozonesonde exhibited scattered correspondence and slight overestimation by satellite derived column O_3. Therefore, the model performance could differ from that for column O_3.”

From Table S1 in supplemental material, the negative biases are found at > 60°N sites, but satellite data are not available over this latitude.

P 13, l 16-18: This is another example for the main problem of this study: You investigated April 2010, only, but you give the impression that you could draw more general conclusions out of it. You should extend the modeled time series in order to give these conclusions.

Reply:

To address the reviewer's comment, we have caveated our conclusions by adding "during April 2010," to this sentence.

P 15, l 21-24: Which measure did you use for saying the model has good skill for representing the main hemispheric O3 distribution? The model is obviously too high over Africa in the equator region and it shows higher values over continents and lower values over oceans compared to the satellite observations.

Reply:

This was based on the statistical analysis summarized in Table 3. To avoid the overstatement, we have revised this sentence as follows in Section 4:

"The results of the statistical analysis for tropospheric column O_3 are also listed in Table 3. The mean of observed and modeled tropospheric column O_3 across Northern Hemisphere is close on average, with an R of 0.65, an NMB of 4.7%, and an NME of 13.5%. The performance of tropospheric column O_3 judged based on the evaluation protocol developed for mixing ratios, suggests that the model satisfies the performance criteria proposed by Emery et al. (2017)."

Minor comments:

P 2, l 16: do you mean that the number of low O3 days increased or the concentrations on the low O3 days? These are very different things and it is not clear, here.

Reply:

We have revised this sentence as follows:

"O_3 concentrations on low O_3 days have increased"

P 14, l 31: shows

Reply:

We have corrected this.

P 15, l 31/32 and P 16, l 8: which emissions lead to high surface O3 mixing ratios?

Reply:

The part 2 paper addresses this question. To address the reviewer's comment, we have added the sentence as follows in Section 4:

"The Part 2 paper will focus on other factors that affect surface O_3 mixing ratio, namely emissions, and also examine the relative importance of NO_x and VOCs."

P 17 – 22: The references need to be revised with respect to formatting and initials.

Reply:

We have rechecked reference style and revised.

Figures:

Figure 2: It is impossible to judge the distribution of the red and the green points when the blue squares are plotted in this way.

Reply:

We have revised this figure to show each surface observation as follows.

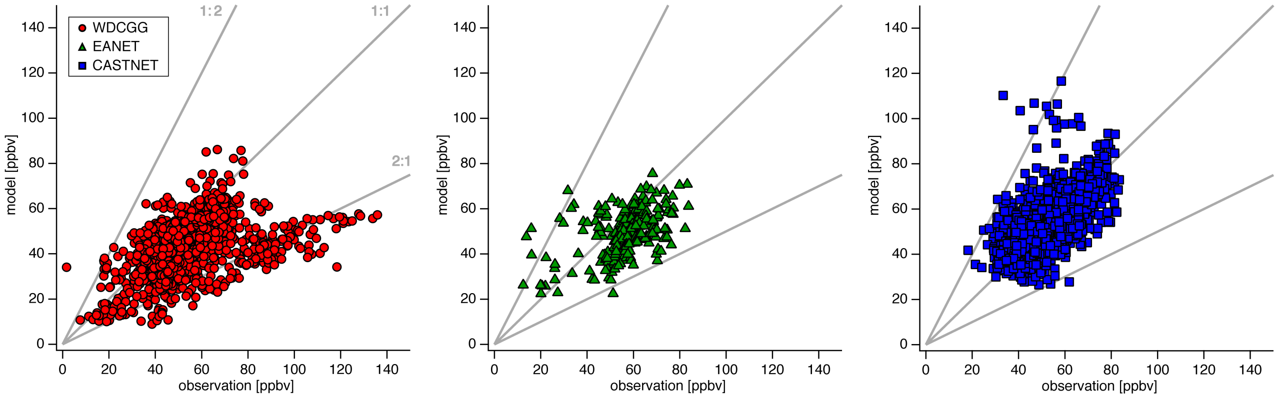

Figure 4 and Figure 5: These plots look nice but I think not all of them are needed. You may put some of them into the supplemental material.

Reply:

To address this comment, we have divided six sites shown in Figures 4 and 5 into 3 sites (Trinidad Head, Boulder, and Huntsville) in the main text and 3 sites (Hilo, Wallops Island, and Rhode Island) in the supplemental information.

Figure 8: It is not clear to me what the exponential fit stands for. Is it used somewhere else?

Reply:

We have revised the discussion using RH-PV, and now removed this figure.