Response to Referee #1

We truly appreciate the comments and suggestions from the referee. Incorporation of the referee's comments has greatly improved our paper. Below is a detailed response to the comments. The referee's comments are in regular font and our responses are in bold font.

General Comments

This manuscript applies five different anthropogenic emission inventories with WRF/Chem to examine surface CO and O3 for Southeast Asia. The authors compared the simulations with observation data, and concluded that none of the emission inventories are better than the others. While most current studies have focused on East and South Asia, this manuscript provides some insights about Southeast Asia. The underlying work in this manuscript will be a useful contribution to the literature. While this manuscript provides detailed comparisons and long discussion, I suggest the authors summarize the major innovations besides the study region in the introduction sections to emphasize its significance and give the readers a sense of "big picture".

We have added a few sentences summarizing the major innovations besides the study region. These sentences have been added to the Introduction in the fourth paragraph, which now reads as the following.

"While previous studies (e.g. Ohara et al., 2007) have compared different emission inventories, a comparison of simulated surface CO and O_3 mixing ratios resulting from different emission inventories, yet using the same model framework, has not been done. Here, the Weather and Forecasting Model coupled with Chemistry (WRF-Chem) is used to examine the variability of predicted O_3 and CO surface mixing ratios when five different anthropogenic emission inventories (RETRO, INTEX-B, MACCity, SEAC4RS and a modified SEAC4RS) are used as inputs. By conducting this comparison using the same meteorology and chemical mechanism, differences in results due to model meteorology and chemical mechanism are mitigated. We focus this study on Southeast Asia, an area that has received little attention, yet has substantial anthropogenic and biomass burning emissions. As part of our study, we examine the effect of biomass burning emissions on surface O_3 and CO by contrasting results from a low biomass burning period (December) with a high biomass burning period (March)."

Specific Comments

Section 3.1, for each emission inventory, it is better to state whether it includes shipping emissions, whether it includes non-combustion sources, and whether it includes four sectors (residential, industry, power and transport). I feel that such kind of clarification will make the comparison in section 3.2 more meaningful, and show the readers whether these emission inventories have consistent emission sources

This is a very good suggestion. We have added a table (below) to clarify what activities are part of each emissions inventory. Further, we have added text to the manuscript (Section 3) describing which emission sectors are missing from specific emissions inventories.

Table 1. Emission sectors used in the model simulations from each emission inventory.

RETRO	INTEX-B	MACCity	SEAC4RS
1. Power Generation	1. Power Generation	1. Energy production and distribution	1. Power Generation
2. Residential	2. Industry (combustion and non-combustion)	2. Industry (combustion and non-combustion)	2. Industry (combustion and non-combustion)
3. Industrial combustion	3. Residential	3. Land transport	3. Residential
4. Industrial processes	4. Transportation[a]	4. Maritime transport	4. Transportation[b]
5. Extraction distribution of fossil fuels		5. Aviation	
6. Solvent use		6. Residential and commercial	
7. Road transport		7. Solvents	
8. Other mobile sources[c]		8. Agriculture	
9. Waste treatment and disposal		9. Agricultural waster burning on fields	
10. Agriculture and Landuse change		10. Waste	

[a]Transportation for INTEX-B includes road, railways, aviation, and maritime transportation.
[b]Transportation for SEAC4RS includes road, railways, and aviation transportation.
[c]Other mobile sources for RETRO include aviation and maritime transportation.

Line 15 on page 9352: can the authors give more details about "what was developed for Europe"? It is not quite clear how the RETRO seasonal cycle is developed.

The final report on the RETRO emissions (Schultz et al., 2005) states that the European monthly emissions used what was developed in the LOTUS-EUROS model. The LOTUS-EUROS documentation (Schaap et al., 2005) states that the monthly emission factors are derived from a critical review of what these factors should be for each emission sector and give a table listing these factors. We have added this information to Section 3.1.

Lines 23-26 on page 9353: be specific about "other emission inventories". How can the emission inventory in 2000 be used to make the conclusion about inventory in 2010? It is better to provide more references.

In the noted sentence, the "other emission inventories" are RETRO, INTEX-B and SEAC4RS. We have clarified this in the text. We have also added more information regarding the significant biases of MACCity emissions that was discussed in Lamarque et al., 2010. The new text is as follows.

"Lamarque et al. (2010) did not find significant biases in their comparison of 2000 MACCity emissions with published emission estimates (e.g. RETRO and EDGAR). However, they estimate that these emissions have an uncertainty of about a factor of 2 based on Bond et al. (2004, 2007) and Smith et al. (2010). The uncertainty of the 2010 emissions was not reported."

Line 18 on page 9354: be specific about ship emission. Does it include both international and domestic shipping?

For RETRO and INTEX-B emissions, only international shipping emissions are included. MACCity emissions include international shipping, domestic shipping and fishing, but exclude military vessels. We have clarified this in the text.

"In addition to the four inventories described above, we have conducted a simulation with a combined MACCity/SEAC4RS emissions inventory, which replaces MACCity with the SEAC4RS emissions over Asia yet includes the MACCity ship emissions, which include international shipping, domestic shipping and fishing. Note, that in the RETRO and INTEX-B inventories, ship emissions represents only international shipping."

Lines 24-25 on page 9354: as the authors introduced, "both the RETRO and MACCity emission inventories have monthly temporal variability", why only RETRO is chosen to estimate monthly emissions in INTEX-B and SEAC4RS. How much uncertainty would be introduced to the final simulations with the assumption that these three emission inventories follow the same monthly allocation? Besides, will this assumption affect the following discussion that different emission inventories make little variation in modeled surface mixing ratios?

When comparing the seasonality of RETRO CO emissions with MACCity emissions (Figure below), it is evident that they have very similar temporal trends. Thus, we do not expect the choice of RETRO for CO emission seasonality to contribute to the uncertainties in the results. When comparing the seasonality of RETRO NO emission with MACCity emissions, the seasonality is not as similar as it was for CO emissions. There is a bigger difference in March NO emissions compared to February, and a bigger difference in November compared to December. These differences could introduce uncertainties in the O_3 results for the region (which is mostly NO_x-limited). Therefore we have added the following sentences to the discussion on seasonality of emissions (Section 3.1).

"The MACCity seasonal variation is very similar to RETRO for CO emissions, but does differ somewhat for NO emissions. The change of NO emissions from February (when we start the simulation) to March differs between these two inventories with little change in high NO emissions for the MACCity inventory and a 5% decrease in NO emissions for the RETRO inventory. While this is a small difference, the change in NO emissions could affect O_3 production downwind of NO sources."

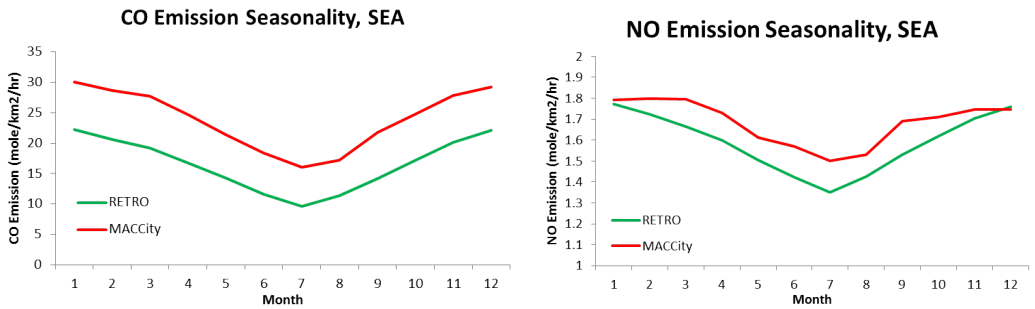

Figure. Seasonality of CO (left panel) and NO emissions (right panel) for Southeast Asia from the RETRO (green) and MACCity (red) emission inventories.

Equations (1)-(3): is the monthly estimate based on each grid or total emissions of over the entire model domain? What is the difference between "monthly emissions" defined here and "monthly-average emissions" defined in line 10, page 9366?

The monthly estimate was based on each grid. Monthly emissions are the emission data for each month for entire year. We apologize for the poor wording. We have replaced "The monthly-average" with "The monthly emission".

Lines 13-14 on page 9355, how to make the conclusion that biomass burning sources dominate the emissions of NOx in March? Based on Table 1, NO emission from biomass burning is less than that from anthropogenic sources.

Thank you for pointing this out. We have now changed the sentence to "In March, the biomass burning sources dominate the emissions of CO.".

Lines 1-3 on page 9356, from RETRO in 2000 to SEAC4RS in 2012, there should be emission reductions.

Thank you for pointing this out. We have now revised the sentence to contrast RETRO emissions to MACCity/SEAC4RS emissions because of the lack of ship emissions in the SEAC4RS-only emissions. The sentence has been changed to "By comparing RETRO emissions to MACCity/SEAC4RS emissions, the total anthropogenic emissions in Southeast Asia decreased by ~30% for CO and ~13% for NO_x between 2000 and 2012 with 2010 ship emissions.

Lines 4-13 on page 9356, authors made comparison between different emission inventories, and mentioned that less NO emissions in SEAC4RS are due to lack of ship emissions. How much contribution do ship emissions make to the total emissions? Is it possible that some of the emissions have been reduced by applying emission control technologies?

On page 9354 lines 16-20, we said that the contribution from ship emission accounts for 15% of the NO emission and 0.1% of the CO emissions. Yes, it is possible that applying

emission control technologies has reduced some of the emissions. However, the SEAC4RS emission inventory does not include any ship emissions, which is the point of the sentence.

Lines 14-29 on page 9356, suggest authors show emissions in Ohara et al. (2007) and REAS v2.1 inventory in Table 1. It is hard to follow this paragraph without any numbers.

Thank you for the suggestion. We now include the emission estimates from REAS v1 in Table 1 (now Table 2). The estimate in the table comes from the Emissions of atmospheric Compounds & Compilation of Ancillary Data (ECCAD) web site (http://eccad.sedoo.fr) in order to obtain emission estimates for the same region as our model domain. We then relate those numbers to those reported in Table 6 of Ohara et al. (2007) so that we can compare the WRF-Chem emissions with the TRACE-P and EDGAR3.2 emission estimates for the Southeast Asia region. The text has been modified to include emission rates in the discussion to the following.

"The CO and NO emissions used in our study are larger than the REAS v1 emissions (Ohara et al., 2007) for our modeling domain (Table 2). The REAS v1 estimate in Table 2 comes from the Emissions of atmospheric Compounds & Compilation of Ancillary Data (ECCAD) web site (http://eccad.sedoo.fr) to obtain emission estimates for the same region as our model domain, which encompasses small regions of India and China that are not included in the Southeast Asia region denoted by Ohara et al. (2007). For our model domain the REAS v1 annual emissions are 91.4 Tg year^{-1} for CO and 4.81 Tg year^{-1} for NO_x. For the Southeast Asia region, Ohara et al. (2007) report in their Table 6 annual CO and NO_x emissions of 54.5 and 3.77 Tg year^{-1}, respectively, but these exclude international aviation, international shipping and open biomass burning. The REAS v1 emissions are even greater than the TRACE-P, EDGAR 3.2, and IIASA CO emissions (34.0, 42.6, 39.8 Tg CO year^{-1}, respectively) but are more similar to TRACE-P, EDGAR 3.2, and IIASA NOx emissions (3.06, 3.91, 3.94 Tg NO_x year^{-1}, respectively) for Southeast Asia (Ohara et al., 2007) as well as REAS v2.1 (Kurakawa et al., 2013), which were 36.2 Tg CO year^{-1}and 3.00 Tg NO_x year^{-1}. Thus, the emissions used here are larger than the REAS emissions inventories as well as other previous inventories."

Table 2. Summation of CO emissions and NO emissions (mole km^{-2} hr^{-1}) from all grids in the model domain for each month.

Emission Inventory	E_CO (mole km^{-2} hr^{-1})		E_NO (mole km^{-2} hr^{-1})	
	March	December	March	December
RETRO – 2000	410,840	496,860	30,590	39,320
INTEX-B – 2006	396,170	406,240	27,410	29,640
MACCity – 2010	436,750	454,250	27,440	28,280
MACCity/SEAC4RS	319,420	320,310	29,810	30,910

SEAC4RS – 2012	**305,542**	**300,369**	**16,610**	**17,290**
Biomass Burning – 2008	**717,940**	**58,780**	**10,220**	**700**
REAS v1[a] – 2000	**282,120**		**13,828**	

[a]REAS v1 emissions are from the ECCAD web site (http://eccad.sedoo.fr) and are the annual emissions converted to hourly emissions assuming constant emissions for the year over the WRF-Chem model domain.

Section 5.1, Figure 5 and 6, it is difficult to make comparison with current figures, I suggest the authors to add figures about the differences between modeling and observation.

Following the reviewer's suggestion, we added panels to Figure 5 and 6 to show the differences between model and observation (see Figures below). However, we are concerned that when we do this the panels get smaller and are harder to read. Further, we think that the qualitative differences between WRF and observations are pretty clear in the original Figures 5 and 6. Therefore, we have decided to use the figures that we originally submitted.

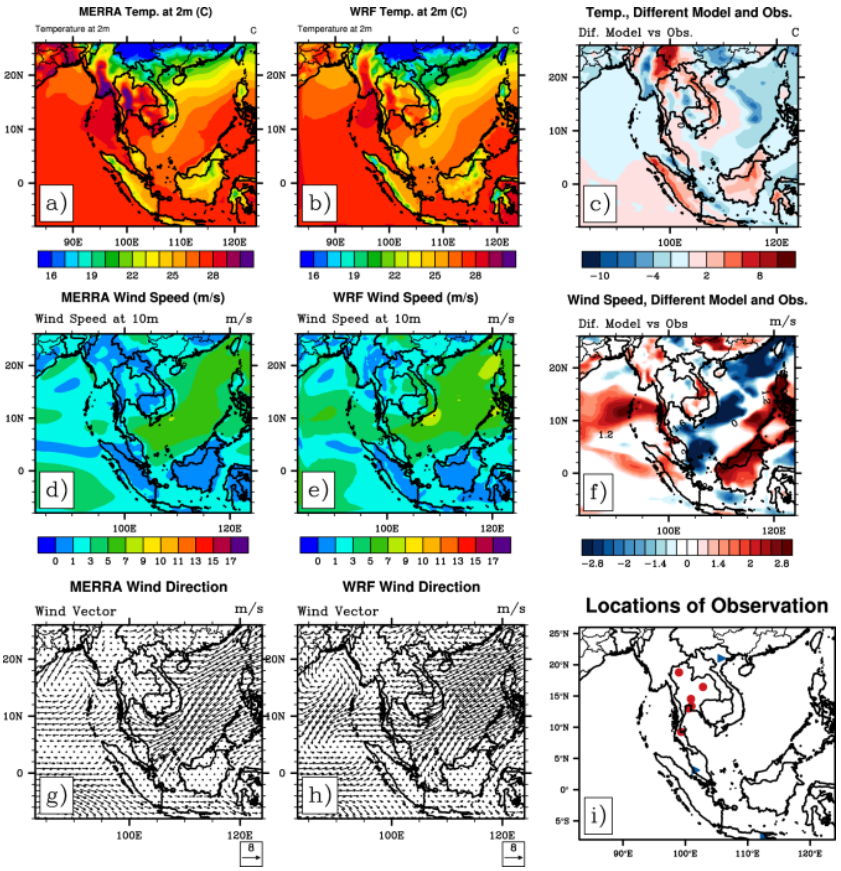

Figure 5. March 2008 monthly-averaged (a) 2 m temperature from MERRA, (b) 2 m temperature from WRF, (c) difference between (a) and (b), (d) 10 m wind speed from MERRA, (e) 10 m wind speed from WRF, (f) difference between (d) and (e), (g) 10 m wind direction from MERRA, and (h) 10 m wind direction from WRF. (i) Locations of ground-based CO and O3 measurements and ozonesonde sites.

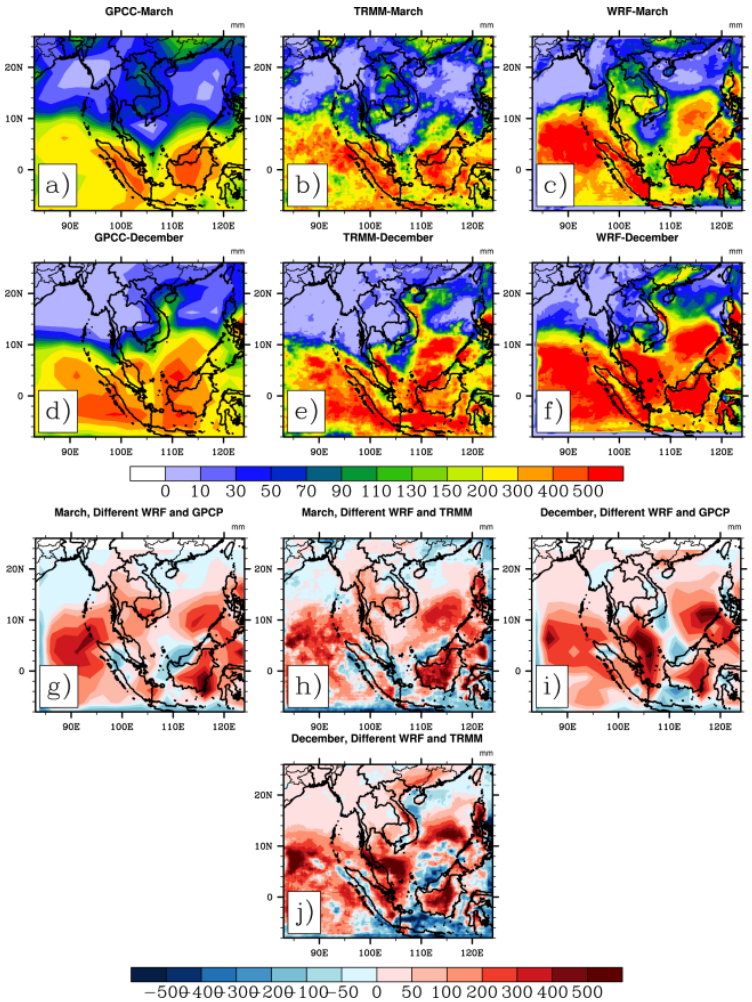

Figure 6. Accumulated precipitation (a) GPCC, March, (b) TRMM, March, (c) WRF, March, (d) GPCC, December, (e) TRMM, December, (f) WRF, December. Differences between (g) WRF and GPCC for March, (h) WRF and TRMM for March, (i) WRF and TRMM for December, and (j) WRF and TRMM for December.

Lines 3-5 on page 9362, how to separate the influence of biomass burning and anthropogenic emissions?

We determine the influence of biomass burning versus anthropogenic emissions by comparing model results from March (high biomass burning emissions) to those from December (low biomass burning emissions). We have reworded the text at these lines to clarify this approach to the following.

"By comparing the model results from March (high biomass burning emissions) to those from December (low biomass burning emissions), the influence of biomass burning emissions can be seen for all three species. CO mixing ratios are > 500 ppbv over Burma and northern Thailand during March compared to 200-500 ppbv during December."

Lines 11-19 on page 9362, what is the definition of "variation" in the predicted monthly-average surface mixing ratios across the five simulations? The authors listed ranges of variation for difference pollutants, but what kind of information we are supposed get from the values of variation? What cause the variations? While the high variations of NOx were explained by differences in ship emissions, what are the reasons for high variations over land?

We apologize for not defining "variation" in the manuscript. It is the standard deviation of the 5 simulations. This is now defined in the text when we introduce the variation discussion. Because the meteorology and biomass burning emissions are the same among the 5 simulations, the primary cause for the variations are the differences in the anthropogenic emissions. As part of this, missing sectors in the emission inventories or different emissions for different years contribute to the variations. We have modified the text to bring out these causes to the following.

"The variation, which is defined as the standard deviation of the five simulations, in the predicted monthly-averaged surface mixing ratios of CO, O_3, and NO_x across the five simulations is highlighted in Figure 8. Because we conducted each simulation with the same meteorology and biomass burning emissions, the primary cause for the variations are the differences in the anthropogenic emissions. CO mixing ratios vary across simulations by <20%, but variations of ~30-60% are found near Bangladesh and Indonesia for both March and December. O_3 mixing ratios have up to 30% variation near the tip of the Malaysian peninsula and near Indonesia, but have much smaller variability elsewhere. Mixing ratios of NO_x have the most variation among the simulations. The 70-100% variations for NO_x, especially over the South China Sea, are from the differences in ship emissions from each inventory. There are also high NO_x variations in several cities as seen by the locally high values in Fig. 8e,f due to different emission strengths in each inventory and to missing emission sectors in some inventories (e.g. shipping emissions in the SEAC4RS inventory)"

Lines 23-24 on page 9362, Saraburi is missing in this sentence.

Thank you for pointing this out. We have now modified the sentence to "The 6-hourly daytime (00:00, 06:00, 12:00 UTC) CO mixing ratios from WRF-Chem with each of the five inventories are compared to observations from the six ground-site measurements: Chiang Mai (CM) in northwest Thailand, Khonkaen (KK) in eastern Thailand, Nonthaburi (NTB) in the Bangkok metropolitan region, Sarabuti (SRB) just north of Bangkok, Chonburi (CB) southeast of Bangkok, and Suratthani (SRT) in the southern peninsula (Fig. 5g)."

Fig. 9 (g)-(l) and Table 3, if I understand it correctly, figures show that the model underpredicts most of December CO except some points at Chiang Mai and Saraburi. But Table 3 only shows positive bias at Chonburi, and the authors stated that higher emissions are modeled at Chonburi and Suratthani (line 8 page 9363). The authors are better to clarify these statements and explain the underestimates of CO (in both March and December) by model simulations.

We have now changed the sentence and elaborated on explaining the underprediction of CO to "In December, the predicted 6-hourly daytime surface CO for all simulations is

much less than the observations, with the exception of the Chonburi site. The large underprediction is reflected by the bias calculation (Table 4). Part of the underprediction is a result of the coarse model resolution (36 km), which cannot capture the highly variable emissions and high CO concentrations in an urban setting where the measurement site is located. However, the underprediction of CO could also be due to low anthropogenic emissions (discussed further in Section 5), a high planetary boundary layer height, which would cause dilution of surface mixing ratios, and/or missing chemistry in the model such as heterogeneous chemistry (Mao et al., 2013)."

Fig. 10 and 11 (also Fig. 14 and 15), I suggest the authors plot the ratios between observations and model simulations, instead of absolute values.

We think differences in the spatial patterns between observations and model are sufficiently depicted in the original figures, which allow us to use the same spatial resolution provided in the original data or model output. Thus, we prefer using the submitted figures for our paper.

Lines 3-8 on page 9366, I am curious why the model predict different peak values and regions from observations? Also why all five simulations predict relative low NO2 column (lines 15-17) over Burma in December?

For March, the WRF-Chem NO_2 column mostly reflects the biomass burning emissions pattern (Figure 3), while for December WRF-Chem is more similar to the anthropogenic emissions (Figure 4). The OMI NO_2 column does not show the high NO_2 over northern Thailand and Burma where the model has high biomass burning emissions in March. To explain this difference, WRF-Chem fire emissions could be too high, or OMI may miss high NO_2 because of cloud interfering with the instrument's view. In situ measurements would allow us to evaluate better the performance of the model. The anthropogenic emissions in Burma are lower than the surrounding regions and are possibly too low. We have added this discussion to the text.

Lines 8-11 on page 9366, it seems that shipping emissions explain a lot of variations among model simulations with different inventories (also lines 1 and 10 on page 9356, line 18 on page 9362, and line 22 on page 9365). I am wondering whether it is necessary to separate shipping emissions in each inventory and show them in Table 1.

The shipping emissions (listed in the Table below) for our model domain from each inventory do vary, especially between RETRO and the other two inventories. Shipping emissions are important in near coastal regions. For example, Huszar et al. (2010) ACP found that the contribution of ship-induced surface NO_x to the total NO_x is 10-30% in the North Atlantic coastal regions. In addition to adding ship emissions to Table 1, we have modified the text in the NO_2 evaluation section to the following.

"The largest variation … a result of both low NO_2 mixing ratios from the MOZART boundary conditions and different estimates for shipping emissions among the different inventories (Table 2, Figs. 3 and 4). Both the RETRO ship emissions, which are 75-80% smaller than INTEX-B and MACCity ship emissions, and the SEAC4RS only simulation,

which does not have ship emissions contribute to the variation. When the MACCity ship emissions are combined with the SEAC4RS emissions (MACCity/SEAC4RS), the agreement with OMI NO_2 column is much better than the SEAC4RS only simulation.

Table. Total and shipping CO and NO emissions for the model domain.

	E_CO (mole km^{-2} hr^{-1})				E_NO (mole km^{-2} hr^{-1})			
	March		December		March		December	
Emission Inventory	Total	Ship	Total	Ship	Total	Ship	Total	Ship
RETRO – 2000	410,840	3,404	496,860	3,364	30,590	5,097	39,320	5,186
INTEX-B – 2006	396,170	5,888	406,240	5,785	27,410	3,273	29,640	3,301
MACCity – 2010	436,750	3,569	454,250	3,717	27,440	3,980	28,280	5,138

Section 5.2.2 and 5.2.3, the authors compared monthly-average correlation coefficients and biases across model results with five different emission inventories, and then concluded that "none of the anthropogenic emission inventories are better than the others" (lines 25-26 on page 9346). I am wondering whether it is possible to do some paired difference tests and show the significance level if the modeled results are independent on the type of emission inventory?

Thank you for the suggestion. Doing paired difference tests will strengthen our conclusions. We apply Kruskal-Wallis non-parametric test, which is used for comparing more than two samples that are independent, or not related, to analysis whether there is significance level for each model simulation or not. The p-values from each simulation from different emission inventory were shown in table below. The p-values of ozone are greater than 0.05, so there is no statistical significant difference between each data set for both March and December (Table below). There is a statistically significant difference between model and observations for CO at some sites in both March and December (Table below). For example Khonkaen, Saraburi, Nonthaburi and Chonburi have p-values < 0.05, while Chiang Mai CO has a statistically significant difference only in December. Suratthani, which is further from urban emissions, is the only site where there is no statistical difference for CO.

After doing this statistical analysis, we have modified our conclusion and abstract to state the following. "None of the anthropogenic emission inventories are better than the others for predicting O_3 surface mixing ratios. However, the simulations with different anthropogenic emission inventories do differ in their predictions of CO surface mixing ratios producing variations of ~30% for March and 10-20% for December at Thai surface monitoring sites.

Table. p-values at the Thai monitoring stations using the Kruskal-Wallis non-parametric test.

Species	CM		KK		SRB		NTB		CBR		SRT	
	Mar	Dec	Mar	Dec	Mar	Dec	Mar	Dec	Mar	Dec	Mar	Dec
Ozone	0.846	0.409	0.526	0.557	0.726	0.416	0.576	0.576	0.783	0.77	0.03	0.397
CO	0.593	0.0007	0.024	0.023	0.0001	0.0001	0.002	0.039	0.0001	0.072	0.272	0.307

Section 5.2.4, how do the model results compare to each other when different emission inventories are used to predict NO2 columns?

In the original manuscript, we included a sentence (page 9366, lines 8-11) stating, "The largest variation among the model simulations occurs in this region near Indonesia and is a result of both low NO_2 mixing ratios from the MOZART boundary conditions and different estimates for shipping emissions among the different inventories." We now add to this information that the NO_2 columns from the different WRF-Chem simulations have a similar pattern and magnitude, and have put this information in a separate paragraph.

<Editorial Comments>

1. Use carbon monoxide or CO (ozone or O3) consistently
2. Is the abbreviation for Chonburi CB (line 23 page 9362) or CBR (Fig. 9)? Be consistent.
3. Line 9 on page 9364, Table 4, instead of Table 5, shows correlation coefficients.
4. Line 10 on page 9364, Table 5, instead of Table 6, shows O3 biases.
5. Show units in Tables 3 and 5.

Thank you for the editorial comments. We have made these changes.