Reviewer 1

This work investigates the impact of two-way coupling of regional chemical transport models (CTMs) with finer horizontal resolutions to a global CTM with coarser resolution. The focus is on the feedback of CO and how it changes the results of the global model, which are then compared with data from several field campaigns (HIPPO) over the Pacific. The authors show that the two-way coupling of all simulated chemicals causes a global increase in CO concentrations that are closer to observations than the results of the global model alone. They ascribe this to better representation of resolution dependent processes, as natural emissions, in the higher resolved nested models.

Investigating the resolution dependency of atmospheric chemistry is an important research topic. Global chemistry climate simulations are restricted to coarse horizontal resolutions because of computational costs and the presented newly developed two-way coupler has the potential to improve such simulations with relatively low additional CPU time needed. The results are presented in a well structured and concise way. Tables and figures are well chosen to support the understanding of the text. Applied methods are generally feasible to analyze the different simulations and compare them with the HIPPO data.

I have only one major and some minor comments:

We thank the reviewer for comments, which have been incorporated into the revised manuscript.

Major comment:

1) Section 4.3: There is a strong north-south gradient in the CO mixing ratios (Fig. 6). I recommend to analyze mean profiles (as in Fig. 8) separated for different latitudinal bands, .g., 60-90N, 30-60N, …. From Fig. 6 it seems to me that the simulated underestimation of CO over the North Pacific is partially compensated by slight overestimation over the South Pacific (at least for the coupled model), which would lead to better congruence of the simulated and measured profiles in Fig. 8 for the wrong reasons. Profiles for different latitudes would bring regional differences to light and could help understanding which processes are not well captured by the model.

In the original manuscript, we had analyzed the model biases separately for the Northern and Southern Hemisphere, for both the global model alone and the two-way coupled model (Sects. 4.2 and 4.3). In the revised manuscript, we have moved the discussion of two-way model biases to Sect. 4.2, per the reviewer's suggestion. In particular, the biases in the Southern v.s. Northern Hemisphere are specified in Sect. 4.2:

"However, the two-way coupled simulation is closer to HIPPO CO than the global model, particularly for the high values over the North Pacific. The mean bias in the

two-way coupled model is only 1.1 ppb (1.4%) below 9 km, with a bias of -1.8 ppb (-2.3%) for the North Pacific and 2.6 ppb (3.3%) for the South Pacific. The global model generally underestimates HIPPO CO with a mean bias by -7.2 ppb (-9.2%) below 9 km; the bias is much larger over the North Pacific (-10.2 ppb, -13.1%) than over the South Pacific (-1.6 ppb, -2.1%). Such a negative bias in the Northern Hemisphere is typical for coarse-resolution global CTMs (Naik et al., 2013)."

Per the reviewer's suggestion, we have added an analysis and a figure for the six latitude bands (see the details in the end of Sect. 4.3. and Fig. 9). Note that the two-way coupled model has a mean bias less than 3 ppb between 0 and 9 km at all these bands, an overall much better performance than the global model alone.

Minor comments:

1) P18962 L8ff: This is too technical for the abstract in my opinion. I recommend anything like: "We develop a two-way coupler, PKUCPL, allowing for exchange and interaction of chemical constituents between the global model (at 2.5 long.×2 lat.) and three nested models (at 0.667 long.×0.5 lat.) covering Asia, North America and Europe, respectively."

Modified.

2) P18963 L2, L3, L6: Write "global model alone". See also comment No 5).

Modified.

3) P18965 L14: Give little more information on the setup. How are the 47 layers distributed over the atmosphere? What are the time steps of the global and the regional models?

We have added the suggested info in the revised Sect. 2.1. Specifically, all models have 47 vertical layers extending from the surface up to 0.01 hPa in a hybrid pressure-sigma coordinate, and the lowest 10 layers are of ~130 m thick each (see Appendix 3 at http://acmg.seas.harvard.edu/geos/doc/man/). The chemistry time step is one hour in all models, while the transport time step is 15 minutes in the global model and 10 minutes in all nested models.

4) 18966 L7: Please note how/why these regions of the nested domains were chosen. To cover the regions with highest CO emissions?!

We have added the following statement in the revised introduction:

"These regions are main pollution source regions where small-scale nonlinear chemical and physical processes may have significant impacts on the global environment."

5) P18966 L16ff: Please mention one time – maybe in this paragraph – that "global

model" refers to the standalone global model results and "two-way coupled model" refers to the results of the global model with feedback from the regional models. Or always call it "global model alone" and "two-way coupled global model" to make it clearer which results you describe.

In the revised Sect. 2.1, we have added the following:

"Hereafter, we refer to the global model as the standalone global model and the two-way coupled model as the combined model system integrating the global model and all three nested models."

6) P18967 L21: What interpolation method is used for horizontal regridding? Conservative? Which order? Please note in the text.

We have added the following specification in the revised Sect. 2.2:

"To guarantee mass conservation in the horizontal regridding process, pollutant mass in a nested gridcell is calculated and then allocated to global gridcells based on the fraction of the area of the nested gridcell belonging to a given global gridcell."

7) P18967 L25: I would prefer "wall-clock time" instead of "computational time" in this context. Computational time I read as total CPU time, which, of course, is much higher when running the coupled system.

Modified.

8) P18970 L22f: I understand that not all test simulations can be done for a whole year or even longer. But please discuss how well January 2009 represents the annual budget and mention that there remain uncertainties because of analyzing only one month. How robust are the results? What can you say about statistical significance of the numbers in the following paragraph/Table 3?

We agree that the contributions of individual factors may be changed throughout time due to changes in emissions, meteorology and chemical reactivity, although we don't have evidence that such dependence would be very significant. We have added a statement in the revised Sect. 3.2:

"The contributions of these factors may be different in other months and years as a result of changes in emissions, meteorology and chemical reactivity; further research is needed for more systematic evaluation."

9) P18971 L4: What is meant by "global all-source emissions"? Only the online emissions can be affected by the two-way coupling.

The "global all-source emissions" means global emissions from all sources. For clarification, we have changed the sentence to "As shown in Sect. 2.1, due to changes in online emissions, global total emissions are larger in the two-way coupled model

than in the global model by 5% for NMVOC and by 1% for NOx."

10) P18971 L15ff: Please discuss briefly, what the link is between the small-scale spatial variability and the photochemical efficiency of NOx, …?

We have added a text in the revised Sect. 3.2:

"For example, the efficiency of NOx in producing ozone (and thus affecting OH and CO) highly depends on its abundance relative to NMVOC and CO. Decreasing the horizontal resolution leads to more significant artificial mixing of NOx, NMVOC and CO with a resulting effect on its photochemical efficiency."

11) P18971 L25ff: I agree that 1.5% represent the "non-emission small scale variability" and this should appear in Table 3-B2 instead of "other factors". Then the remaining 4.6% need to be "emission small scale variability". Is that what is meant by "small-scale horizontal distributions" in Table 3-B1? In my opinion, "concentrations" needs to be canceled on P18971 L29. Please make this more clear.

We agree and have revised the text as suggested.

12) P18973 L8: On this website they say "there were 787 profiles flown". Why do you use only 620?

The total 787 profiles include those over the continents. We focused on the 620 profiles over the Pacific Ocean, as stated in the original manuscript.

13) P18973 L10f: Please describe in more detail how the model data are interpolated/regridded to "ensure spatiotemporal consistency with the HIPPO data". The comparison of 2.5 °x2 ° model data with the flight tracks is not straight forward. What is done to put the data on the same vertical grid? In Fig. 5 you show profiles at 0.1 km intervals which is much finer than the vertical resolution of the model, I guess. Please clarify.

For clarification, we have added the following statement in the end of the revised Sect. 4.1:

"In particular, model results in the gridbox encompassing the location (longitude, latitude and altitude) of a given measurement are used for comparison with the observation."

14) P18974 L2ff: I would expect the numbers of the mean bias (0-9km, North/South Pacific) also for the coupled simulation here. Then these numbers do not need to be repeated at the end of section 4.3

Modified as suggested.

15) P18974 L25: Which "stratospheric influences" are this? What could be the

reasons why the models reproduce only an about 10 ppb change at 9 km height compared to the measured 20 ppb?

By analyzing the tropopause height in the GEOS-5 assimilated meteorology (used to drive GEOS-Chem here), we had found that for HIPPO-5 in August-September 2011, 70% of data between 8.5km and 10km are located above the tropopause with the remaining 30% close to the tropopause. We had thus determined that these measurements are significantly affected by stratospheric influences. Our model results also show such influences, but with a lower magnitude possibly due to overestimates in stratospheric CO. We have updated the statement accordingly in the text.

16) Section 4.3, Fig. 8: Please discuss the seasonal differences. What could be the reasons why the simulated profiles seem to fit better with the measured ones during NH-winter in (a) and (b) while it looks worse in (c) – (e)?

The two-way coupled model compares very well with measurements in HIPPO-1 (winter). It compares fairly well in HIPPO-2 (late fall) and HIPPO-5 (summer; expect for the positive biases above 9 km as influenced by the stratosphere). For HIPPO-3, the coupled model reproduces the mid-tropospheric peak in the HIPPO data, but with a much smaller magnitude. This is likely due to the use of monthly biomass burning emissions that do not capture the episodic emissions occurring during the HIPPO-3 times. As shown in Fig. 3b, biomass burning is very active in this spring period. Possibly for similar reasons, the coupled model also does not well capture the observed peaks in the lower and middle troposphere in HIPPO-4 (summer). For all campaigns, however, the two-way coupled model performs much better than the global model alone. We have updated Sect. 4.3 to include the above analysis.

17) P18975 L23ff: How were the numbers for emission increases determined?

We determined the emission increases first by the relative contribution of emissions (~40%) versus photochemical sources (~60%) of CO, followed by in a series of tests with CO emissions changed by a multiple of 5% (i.e., 10%, 15%, 20%, 25%, 30%, 35%, 40%, respectively. Finally, we chose the test that best matches the two-way coupled model.

18) P18977 L2ff: cf. comment No 9) 4.6% are due to "small-scale spatial variability in NOx ..." emissions. Is that right?

Yes. We have clarified the statement in the revised conclusion section.

19) P18977 L20ff: Could you draw a general conclusion on the emission increase? Would you recommend to increase the CO emissions by an averaged value of 25% for future simulations with the global model ?

Our related analysis implies that the coarse resolution of a global model is a critical limiting factor in its use for CO emission constraint. We showed this point by

universally increasing a certain amount of CO emissions globally in the global model to 'best' match the two-way coupled model. To have optimized 'effective' emissions in a global model, a more detailed test with region-dependent CO emission adjustments is necessary.

20) P18978 L5: What is meant by "multi-layer" in this context?

"multi-layer' means integrating more than 2 resolutions (e.g., 2x2.5, 0.5x0.667, 0.25x0.3125) in a coupled model system. We have clarified the statement.

21) Table 4, footnote 3: "delta air pressure" is not clear to me. Is it the thickness of layer i in hPa?

Yes. We have modified the writing accordingly.

Technical corrections:

1) P18976 L17: ...those simulated by the global model...

Modified.

2) Fig. 4: Please include the abbreviations (GM) and (NM) behind "Global/Nested Model" in the center box.

Modified.

3) Fig. 5: To make it more clearly, please include the zero line in panels b)-d) for the differences.

Modified.

4) Fig. 6: In the present form each panel looks like one single time series. To distinguish between different campaigns, I recommend to put little spaces or vertical lines between the single campaigns.

We have distinguished the individual HIPPO campaigns in the x-axis of each panel.

Reviewer 2

Review of Yan et al. acp 2014 "Tropospheric CO over the Pacific during HIPPO: two-way coupled simulation of GEOS-Chem and it multiple nested models" It's an interesting and important study that pushes to extreme GEOS-Chem's high resolution capabilities and a large aircraft data set. The focus is a rather well studied area - Pacific - and it's ever important and ever changing pollution transport. It's good to see that the authors did not stop at simply evaluating CO concentrations, an important work in itself, but also looked at OH and methyl chloroform, also comparing those

values to observations.

We thank the reviewer for comments, which have been incorporated into the revised manuscript.

It's perhaps a bit of a stretch to reason that CO concentration enhancement in the nested model is due to incorrect emissions - it seems too simplistic of an explanation.

We showed in Sect. 3.2 that the online calculation of biogenic NMVOC emissions and lightning and soil NOx emissions contribute to the improved CO simulation in the two-way coupled framework (relative to the global model alone). This factor contributes to about one third of the overall CO improvements. We did not state the correctness of emissions of CO or other species.

Specific comments:

-The work on two-way coupler is commendable and its thorough testing much appreciated. As resolution increases, surely this will be of use in the future as well.

We thank the reviewer for positive reviews.

-I don't see any comment on addressing vertical resolution. Can that remain unchanged?

In the first paragraph of the revised Sect. 2.1, we have added that:

"All models have 47 vertical layers extending from the surface up to 0.01 hPa in a hybrid pressure-sigma coordinate, and the lowest 10 layers are of ~130 m thick each (see Appendix 3 at http://acmg.seas.harvard.edu/geos/doc/man/)."

Global anthropogenic emissions from EDGAR are quite a bit outdated, at least its reference. It would be useful to read a comment on what is expected of these emission inventories given that the observations are much more recent.

We admit that EDGAR 3.2-FT2000 data used here are relatively old. We had however replaced this global dataset with regional (and presumably more accurate) inventories in all three nested domains.

We have also compared EDGAR 3.2-FT2000 with the updated EDGAR v4.2. For areas not replaced by regional inventories (mostly in the Southern Hemisphere), EDGAR 3.2-FT2000 CO emissions are 131 Tg/yr for 2000, slightly larger than 123 Tg/yr EDGAR v4.2 for 2008 (the latest year in EDGAR v4.2); the difference is about 0.3% of total CO sources (emissions + photochemical production) shown in Table 2. Therefore the use of EDGAR 3.2-FT2000 should not affect our simulations significantly. We have added the analysis in the revised Sect. 2.1.

-I couldn't understand the title of section 3.1 - please simplify. In that section, I am

concerned about the last sentence. It seems the issue of differences between different nesting should be explored, but the authors are not reassuring the reader that this is not an issue in their work

We have changed the title to "Accumulation of small differences differentiates the two-way coupled simulation from the global model".

Here we showed that through the two-way coupled framework, the improved simulation via higher resolution in one area can greatly affect the simulation for other areas, at whichever resolution other areas is simulated. Such effect has not been documented previously, is briefly shown here, and should be explored further by later studies.

-p1897, line 25: "flied" should be "flew"

Modified.

-Section 4.4 I wish the authors elaborated on the implication of higher resolution on decreased emissions estimates. Do we see a trend that the lower resolution we use the higher our CO emission constraints tend to be? It would be nice to see that supported by comparison with previous emission studies that used coarse resolution.

We had stated the implications of resolution on CO emission constraints. We are working on even finer resolution models to test the impacts on model simulation and emission constraint. We expect that there is a resolution threshold, beyond which a finer resolution does not improve the simulation significantly; we don't know the threshold yet.

We have done additional simulations with the global model alone at the horizontal resolution of 4 °lat. x 5 °lon. We indeed see a trend that the lower resolution we use the lower our simulated CO concentrations: A global model at 2 °lat. x 2.5 °lon. leads to global tropospheric mean CO lower by 10.4% than the two-way coupled simulation, and a 4 °lat. x 5 °lon. global model leads to an additional 3% reduction in CO. Consistently, the global OH content decreases and global mean MCF lifetime increases when the resolution increases. We have included the above analysis in the revised Sect. 4.4.

We had cited several top-down CO constraint studies in Sect. 4.4 of the orignial manuscript. We have speficied an exmaple in the revised manuscript as follows: "Global models are used often to constrain CO emissions, where the mean model bias relative to measurements is attributed to emission errors (e.g., Stavrakou and Müller, 2006; Kopacz et al., 2010; Hooghiemstra et al., 2011). These studies tend to suggest CO emissions, as constrained by modeling and measurements, to be higher than those in emission inventories. For exmaple, Kopacz et al. (2010) used satellite CO measurments and a 5° long. x 4° lat. global GEOS-Chem model to constrain CO emissions for 2004-2005. Their results suggested a total of global CO emissions from

combustion (fossil fuel, biofuel and biomass burning combined) at 1350 Tg yr-1, much larger than current bottom-up emission inventories (48% larger than our value at 913 Tg yr-1 shown in Table 2).”

-p.18976, line 16 "show" should be "shows", "CO" should be "CO concentrations"

Modified.

-I look forward to further applications of the two-way nested grid that the authors mention in the last sentence of Conclusion section

We thank the reviewer again for positive reviews.