Response to Referee Comment 3 by Anonymous Referee #2

This manuscript is the first part of at least two parts of a paper series dedicated to the analysis of trans-Pacific transport. This first part is focused on the evaluation of the WRF / H-CMAQ model configuration and on the analysis of stratospheric intrusion. The thorough analysis in the manuscripts has two flaws:

First, the model simulation uses a horizontal grid spacing of 108 km, which is a very coarse resolution to realistically simulate stratospheric intrusions.

Second, unfortunately, the authors are omitting vital information about their most important diagnostic tool, the O3PV tracer. On the definition provided in the article the diagnostic method described in Sect. 3.2 seem to be not fully applicable and thus I doubt the results of Sect. 3.3.

Therefore, depending on the real definition (in contrast to my unterstanding of the description in the manuscript) of O3PV I am rating the manuscript as either reject or major revisions.

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 agree with the reviewer that a finer grid resolution is optimal to simulate STT but this is only possible when computational resources are available, which is unfortunately not the case for this work. The use of a grid resolution of 108 km in simulating STT used in this work is indeed consistent with the findings of other recent studies (i.e., Gray, 2003; Cristofanelli et al., 2003) suggested by the reviewer (see below our responses to major issues on this). Despite a coarse horizontal resolution used for this H-CMAQ modeling system, our model evaluation of O_3 concentration against surface observation network, ozonesonde, and satellite data shows an overall good model performance. Based on the good model performance and the finding from your suggested references, we think that at this time using the resolution of 108 km provides a good compromise between numerical accuracy and computational constraints. We do however acknowledge that as more computational resources become available such investigations should strive to use finer resolution in models. To address the reviewer's comment, we have added some discussion about the justification of the horizontal grid resolution used in this work along with the uncertainties associated and cited the two suggested papers.

Second, we have revised the paper to add all necessary information about O3PV. In particular, in the revised manuscript we further explain that the O3PV tracer tracks O_3 scaled to PV in the upper model layers. The O3PV tracer undergoes the same transport, scavenging, and deposition processes as O_3, but its mixing ratios are not affected by chemical production or loss processes Thus the O3PV can be used as a qualitative an indicator of O_3 of stratospheric origin as parameterized by the modeled O_3-PV correlation.

We provide below our point-by-point responses.

Major Issues

• p. 4 l. 15-20: Looking at the very coarse horizontal resolution of 108 km, it might be nice, that the 44 layer version represents STT better than the 35 layer version. However, the horizontal resolution is much too coarse to expect a good representation of the downward mixing during STT events. (e.g., Gray 2003, Cristofanelli et al., 2003). This alone compromises the usefulness of this study.

Reply:

We have carefully reviewed the two suggested references. While both indicated the models' difficulty in simulating STT at a grid resolution of >1°×1°, they show good skills when a grid resolution of 1°×1° or finer (e.g., 0.5°×0.5°) was used, which is consistent with a grid resolution of 108 km ×108 km used in this work. Based on the findings of these two papers and the overall good performance of our model application, we believe that a horizontal grid resolution of 108 km adopted in the current H-CMAQ is adequate (though not the best) to simulate STT. To address the reviewer's comment, we revised our manuscript by providing some discussion on why this grid resolution was used and what the associated uncertainties may be in Section 2.1 as follows:

"While the use of finer horizaontal grid spacing can better resolve the STT processes, it will substantially increase computational demands. Cristofanelli et al. (2003) analyzed STT by combining analysis of data from a measurement network and predictions from total of seven model simulations over Europe, and reported that three models with 1°×1° horizontal resolution were able to capture the STT whereas other models with coarser resolutions were not. Another study over Europe investigated the cross-tropopause transport in terms of resolution and diffusion coefficient using horizontal resolutions of 2°×2°, 1°×1°, and 0.5°×0.5° and showed that the simulation with the 2°×2° resolution has difficulty to capture

the tracer amount across tropopause (Gery, 2003). Based on these findings and the model evaluation results (see Section 3.1) in this work, we believe that using a grid resolution of 108 km provides a good compromise between numerical accuracy and computational constraints."

"As indicated in Mathur et al. (2017), the 44 layer configuration employed in the H-CMAQ configuration helps better capture dynamics in the vicinity of the tropopause and reduce excessive diffusion relative to coarser vertical resolution configurations."

• p. 4 l. 25: "The value of PV generally increases with altitude...": depending on the shape of the stratospheric intrusion / the PV streamer this is precisely not necessarily the case.

Reply:

Here we discuss the general feature of PV rather than the special case of such STT events. To explicitly state that, we have revised this sentence as follows in Section 2.1:

"The value of PV itself generally increases with altitude, …"

• p. 4 l. 30 - p. 5 l. 4: The definition of the O3PV tracer is not clear. How is this tracer initialised? When (at initialisation, each step ...) and where (free tropopause, stratosphere ...) is this O3-PV relationship used to define the O3PV tracer and how (is O3PV set to O3 in respective regions)? All this is essential for the information this tracer is carrying, thus a much more detailed explanation is required here.

Reply:

Similar to the chemical concentration field, the O3PV tracer was also initialized on March 1, 2010 using the prior model simulation conducted by Hogrefe et al. (2018). The O3/PV relationship is applied to estimate O_3 in layers above 110 hPa, specifically in the three topmost layers at 58, 76, and 95 hPa described in Xing et al. (2016). The chemically-inert O3PV tracer is set to this parameterized O_3 concentration in these three layers. It undergoes the same transport, scavenging, and deposition processes as O_3, but its mixing ratios are not affected by chemical production or loss processes and do not have any other source term beyond the parameterized values in the three topmost model layers. To address the reviewer's comment, we have revised the explanation of O3/PV relationship and O3PV tracer as follows in Section 2.1:

"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 topmost vertical levels of 58, 76, and 95 hPa fitted as a 5th order polynomial function, and applicable between the range of 50 and 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. To track stratospheric air masses, the O_3 estimated using the O3/PV relationship in the three layers listed above is also added as a chemically-inert tracer species in the H-CMAQ simulations as O3PV tracer. The O3PV tracer undergoes the same transport, scavenging, and deposition processes as O_3, but its mixing ratios are not affected by chemical production or loss processes."

• p. 7 l. 2/3 What about high-PV structures in the free troposphere? Are they simply declared to be stratospheric?

Reply:

Based on the criteria to use 2 PVU as the tropopause in this study, this case is regarded as stratosphere and not used to calculate column O_3.

• p. 9 l. 2-4: "Generally, O3 and O3PV mixing ratios are very similar in the upper layers, especially above the 2.0 PVU line, indicating that O3 mixing ratio in these layers are dominated by stratospheric air mass." I thought that is the definition of the O3PV tracer, how could these tracers not be very similar?

Reply:

We have clarified the discussion by adding the following to the discussion on Figure 5:

"As shown in Fig. 4, O_3 and O3PV show similar variation in the upper model layers; however, O_3 is greater than O3PV near the tropopause indicated by 2.0 PVU, and this suggests the presence of photochemical production near the tropopause."

• p. 10 l. 9/10: more importantly the horizontal resolution needs to be increased.

Reply:

As indicated in our response to earlier comments, we agree that as computational resource constraints reduce, finer horizontal resolutions should be employed in models to better capture dynamics through the troposphere and the UTLS. To address the reviewer's comment, we have revised the manuscript to discuss the uncertainties associated with the grid resolution used in this work as follows:
"This is the hemispheric modeling system but the finer horizontal resolution will be another way to improve this."

• p. 10 l. 11-21: What do you expect? RH is a diagnostic quantity which is dependent on a bundle of prognostic variables and sensitive parametrisations. Thus RH is a very difficile variable to base further analysis on.

Reply:

Yes, we understand that RH is a diagnostic variable. It, however, can be used as another indicator for the stratospheric air mass, as demonstrated in our discussion. To address the reviewer's comment, we added the following on Section 3.1 before the discussion on RH:

"Although RH is a diagnostic variable, it may also provide an indication of stratospheric air masses and is thus included in the model evaluation."

We reconsidered the air mass characterization technique, and RH is not used to diagnose the stratospheric air mass.

• p. 10 l. 22-30: You show here that RH is far from realistic in the model but still the new analysis method in 3.2 is based on this diagnosed quantity?

Reply:

Based on the reviewer's comments and the short comment by Heini Wernli, we have removed the use of the RH-PV relationship to estimate the stratospheric air mass in the

reworked analysis presented in the revised manuscript.

• p. 11 l. 10-20 / Fig. 7: I can not agree, that the model captures the observation well. The only thing that is correct is the location of the maximum over the Pacific Ocean.

Reply:

To avoid the overstatement, we have revised this paragraph as follows:

"The observed and modeled tropospheric column O_3 are compared in Fig. 7. The observed latitudinal gradients in tropospheric column O_3 with values greater than 40 D.U. over mid-latitudes, column values around 30 D.U. over high- and low-latitudes, and values below 20 D.U. over the Pacific Ocean near the equator are captured well by H-CMAQ. To illustrate the differences between observations and simulations, the normalized bias is also shown in Fig. 7. This normalized bias map shows model tropospheric column O_3 overestimation over Russia and Africa and a slight underestimation over the Pacific Ocean. While the comparison with surface observations from WDCGG shows model underestimation at four sites over eastern Europe, the model slightly overestimates tropospheric column O_3 in this region. In addition, the model underestimation especially over free-troposphere is found through the model evaluation with ozonesonde (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. 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)."

• Sect. 3.2

– p. 11/12 / Fig. 8 / Table 5: From the data provided here, I can not agree to the method how the relationship between PV and RH is established. There is no proof, that the exponential fit is the best one. Table 5 does not provide any statistical measures to assess the quality of this fit. Maybe an elephant might have been an option too?

– I can think of low humidity conditions without stratospheric influence (e.g. above deserts).

– p. 12, ll. 9ff.: How do you deal with high-PV structures in the troposphere. Where is the tropopause diagnosted in these cases?

Reply:

In the revised manuscript, we have revised the air mass characterization technique, and no longer use relationship between PV and RH is not used anymore. The revised technique classifies a stratospheric intrusion in the case of weak photochemistry (calculated by O3PV/O_3) as shown in the flowchart of Figure 8.

– p. 12/13: too understand this method it is essential to understand how the O3PV tracer is initialised. As explained above, the description provided in this manuscript is not self-explanatory. I assume: the O3PV tracer is set every time step to O3 where PV is higher than 2 PVU (this might include high-PV structures in in the troposphere) and might blur the signal of "real" stratospheric air.:

Reply:

At initialization, 3-dimensional fields of the O3PV were derived from prior H-CMAQ simulations described in Hogrefe et al. (2018). Since the O3PV tracer was included to qualitatively track O_3 from the model upper layers (nominally representative stratospheric origin) as parameterized by the PV-O_3 correlation, the tracer's mixing ratio in layers only above 110hPa is set every time step based on the PV-O_3 scaling. More details on the definition of the O3PV tracer are provided in our response to the comment on "p. 4 l. 30 - p. 5 l. 4"

– Additionally, as the O3PV tracer is transported and depositioned due to its own gradients many deviations between the Ozone and the O3PV tracer might be caused by differences in transport and sinks and not in photochemistry.

Reply:

As explained in the manuscript, the O3PV tracer undergoes the same transport (advection, turbulent mixing, and clouds), scavenging and deposition processes as O_3. The only difference is that it does not undergo any chemical transformation, and thus the differences between modeled O_3 and O3PV fields can be attributed to the photochemistry which impacts the simulated O_3 but not O3PV.

– Fig. 10 and corresponding text: The description of you results reads as if stratospheric ozone would be inert and only tropospheric ozone would take place in photochemistry. The fastest process of all are the autocatalytic cycles of ozone production and destruction. Therefore, the amount of stratospheric ozone influences directly the photochemisty. How is this stratospheric ozone mass calculated? Is it the integral over O3PV? In that case, I would say that the assessment is wrong as you miss its photochemical sink. (provide more details about the calculation p.13, ll.12-14)

– From the current knowledge about the method I would say, that a continuously initialised stratospheric tracer could be a diagnostic tool to diagnose stratospheric influence. But the quantification diagnostic introduced in Sect. 3.2. does not work, unless the authors omitted to provide a lot of vital information about their method.

• Sect. 3.3: As I question the diagnostic method explained in Sect. 3.2, I have to doubt the results of this section as well. Of course you can say, whether the air is influenced by stratospheric air, but the percentages provided in Fig. 11 mean nothing.

• p. 15 l. 30-33: Due to the coarse horizontal resolution of the model it was not to be expected that stratospheric ozone is transported downward efficiently enough to reach the surface.

Reply:

Chemistry is still active across the modeled vertical extent. As in other studies utilizing scaling of O_3 based on PV values, here too O_3 in the model's UTLS is scaled to the dynamically evolving PV fields. The O3PV is an additional diagnostic tracer added to qualitatively track the influence of this O_3 originating in the model upper layers (representing stratospheric O_3 in the absence of a complete representation of the stratosphere and it's chemistry) through the modeled vertical extent, with special interest on the amounts in the boundary layer. We hope that this along with the changes related to the tracer description described in response to earlier comments helps better explain the configuration, processed modeled, and the interpretation of the modeled fields of O3 and O3PV.

To address these four comments and questions, and also to take into account the short comment raised by Heini Wernli, we have revised the air mass characterization technique for stratospheric air mass characterization.

The flowchart depicted in Fig. 8 is revised to exclude the judgement based on PV-RH relation according to the comment by Heini Wernli. The way to judge stratospheric air mass is limited to O3PV/O_3 near 1.0 (range between 0.9 and 1.1), which indicates the weak photochemistry and possible impacts by stratospheric air mass. Then, the top layer (z=44)

is defined as stratospheric air mass, and if the above grid cell is judged as stratospheric air mass under the weak photochemistry, the grid is determined as stratospheric air mass. Figs. 9, 10, and 11 have been updated according to this revised method to estimate the stratospheric air mass. Fig. 8 and Table 5 have been removed because now we do not need PV-RH relationship to judge the stratospheric air mass.

Minor Issues

• title: should contain the model version, as evaluations are always specific for the used model version. Additionally, the title is misleading as the authors miss to point out the interdependencies between STT and trans-Pacific transport.

Reply:

We think the model version in the title may not be needed, but we have revised the title to explicitly indicate both trans-Pacific transport and STT.
We however, provide model versions and details on specific addition in the manuscript description.

• p. 1 l. 29/30: not clear what the message is. Where else could STT impacts come from?

Reply:

To address the reviewer's concerns we have clarified the discussion in the abstract as follows:
"Over the U.S.A., STT impacts show large day-to-day variations, and STT impacts can either originate from the same air mass over the entire U.S.A. with an eastward movement found during early April, or stem from different air masses at different locations indicated during late April."

• p. 2 l. 1: as STT is event based I doubt that the impact is near constant.

Reply:

We have revised previous Figure 12 (now Figure 10) to use monthly-mean data, and explicitly mention that the analysis is based on monthly means. Then we showed Figures 11 and 12 to show the daily variation of stratospheric intrusion over the U.S.A.

• p. 2 l. 17: “acceleration of anthropogenic emissions” ? emissions are not accelerated. They might increase and their increase might be accelerated ...

Reply:

We have revised this sentence as follows in Section 1:
“the dramatic variation of anthropogenic emissions in East Asia”

• p. 4 l. 21: What is cb05e51? A GIT tag ?

Reply:

This statement stands for EPA modifications implemented in CMAQ version 5.1. This is the actual chemical mechanisms available in CMAQ version 5.1 and later, and is documented in the release notes and model documentation web page (https://www.airqualitymodeling.org/index.php/Cb05e51_species_table). We would like to keep this expression.

• unify “O3/PV” vs. “O3-PV” relationship.

Reply:

We have unified this wording in “O3/PV” throughout the manuscript.

• Sect. 2.1: Are these (WRF and H-CMAQ) continuous simulations or are they re-initialised?

Reply:

WRF is newly simulated from March 2009 to have more than one-year spin-up time. H-CMAQ is initialized on March 1 based on concentration and tracer fields archived from a longer H-CMAQ simulation described in Hogrefe et al. (2018). In the current study, the H-CMAQ simulations start on March 1; we use the entire month of March as additional spin-up and then, April is used for the analysis period.

• Sect. 2.2.3 It is really necessary to talk about un-used flight data?

Reply:

This information was provided for completeness as we only used such data at one site. Readers may be curious why we did not use the data at the other two sites.

• Fig. 1: Usage of lighter colors would make it easier to see the symbols. The grey aircraft symbol is hard to distinguish from the grey map lines.

Reply:

The gray color has changed into light-blue color in Figure 1.

• longitude / latitude information is missing in all maps

Reply:

We have prepared the map indicating longitude and latitude of H-CMAQ modeling domain in Figure S3 in the supplemental material.
We believe including longitude and latitude would lead to busy figures; therefore, we would like to omit longitude/latitude information in Figures 1, 3, 7, and 11.

• Fig. 4: thick line not identifyable, black lines are distinguishable only at 300 % zoom and more.

Reply:

To be consistent with Figure 5, we have changed the color of thick lines (2.0 PVU) from black to red, and we also enhanced the thickness of other black lines.

• Table 1: The tables content is not understandable without providing more details,
e.g.:
– What does "ranged" and "zero-out" mean?

Reply:

To be consistent with another expression, we have revised "ranged 10-25 ppbv" into "10-

25 ppbv".
We also have added the footnote to explain "zero-out" as follows:
"[a]: Estimate the impact from the difference between the standard simulation and a simulation with eastern Asian anthropogenic sources shut off."

– "tagged O3": which tagging method?

Reply:

We have added the footnote for detail information of "tagged O_3" as follows:
"[b]: This tagged method divides simulated O_3 into individual O_3 tracer to track O_3 produced in different region."

– "tropopause tracer": How defined / initialised?

Reply:

We have added the footnote for detail information of "tropopause tracer" as follows:
"[c]: This tracer method accounts for STT contribution to O_3 using e90 tracer, which differentiates tropospheric air mass based on the globally uniform surface source and 90-day folding lifetime; both have been spun-up for three years."

– Table 1: the descriptions of the "Estimated impacts" are completely messed, e.g., "5-7 ppbv (17 April- 15 May 2006; INTEX-B), increased by 1-2 ppbv from April-May 2000" What does this mean? The estimate stems from a measurement in 2006 during the INTEX-B compaign and is compared to a 2000 value, where we do not know anything about? And do you mean that it impact increased by 5-7ppbv?

Reply:

The original table includes two types of information derived by different methods, hence we have revised this row into two rows to clearly indicate the method description.

Literature:

• Cristofanelli, P., Bonasoni, P., Collins, W., Feichter, J., Forster, C., James, P., Kentarchos, A., Kubik, P., Land, C., Meloen, J., Roelofs, G., Siegmund, P., Sprenger, M., Schnabel, C., Stohl, A., Tobler, L.,

Tositti, L., Trickl, T., and Zanis, P.: Stratosphere-to-troposphere transport: A model and method evaluation, J. Geophys. Res., 108, 8525, doi:10.1029/2002JD002600, 2003.
• Gray, S.: A case study of stratosphere to troposphere transport: The role of convective transport and the sensitivity to model resolution, J. Geophys. Res., 108, 4590, doi:10.1029/2002JD003317, 2003.